Björn Wiemer

Hearsay in European languages:
toward an integrative account
of grammatical and lexical marking

1. Introduction

The purpose of the present article is to give a survey and classification of means with reportive function used in European languages. Reportivity is a subdomain of evidentiality. Hearsay is understood as a synonym, it is a cover term for different kinds of functions subsumable under reportive evidentiality. These functions, however specific they may be, all indicate that the speaker of the actual utterance bases his/her assertion on previous utterances, usually made by another person (other persons), but not necessarily so (see 3.5). I am not going to discuss some of the notorious principal issues that have lately been raised, such as the relationship of evidentiality to epistemic modality, or the question what counts as a conventionalized, holistic construction in this or that language. For the purposes of this paper suffice it say that I advocate a strict separation of evidentiality and epistemic modality on the conceptual and analytic level[1] and am aware of the fact that many units marking evidential functions carry epistemic components ("overtones"), either additionally or even as a feature more salient than the evidential function. On the one hand, I have been eager to take into account as many types of reportive marking devices as we can find (with one regular exception, see section 2) by including diverse kinds of distinct units for which a hearsay meaning can be considered as a stable meaning component, regardless of genre or register variation and frequency, and regardless of whether other components, e.g. epistemic ones, co-occur with the hearsay meaning. However the range of meanings other than reportive will be accounted for, too. On the other hand, I have been anxious to include only those devices for which a reportive component can be accepted firmly enough, i.e., considered as a conventionalized part of its meaning. There are some borderline cases which will be looked at critically during the discussion of data in section 3.

I conceive of evidentiality as a c o n c e p t u a l domain. My understanding actually fits the definition given by Aikhenvald (2003: 1): "Evidentiality proper is understood as stating the existence of a source of evidence for some information; that includes stating that there is some evidence, and also specifying what type of evidence there is." However, contrary to Aikhenvald, I apply this definition not only to markers that are considered to be grammatical in the strict sense, but no less to lexical means, more properly speaking: to various classes of function words (see section 2), whose semantics contains a stable, non-detachable reference to hearsay. This broader functional notion corresponds to Aikhenvald's recently coined term 'information source' (Aikhenvald 2007). From this extension of the range of linguistic devices for marking functions from a given conceptual domain we should distinguish the question to which degree a linguistic device can be regarded as sufficiently conventionalized (and not only as an evidential strategy), i.e., with a stable, non-detachable indication of an evidential, more specifically: a hearsay function.

Therefore, the approach taken here is based on a functional-onomasiological perspective, but the classification grid to be introduced in section 2 gives scrutiny to the formal properties of devices which serve as indicators of hearsay. These formal properties will, as far as possible at the present stage of our knowledge, be correlated with (i) subfunctions of reportive evidentiality and with (ii) possibly interfering inferential functions and functions from other domains (see section 3). I will group marking devices according to their morphological format and syntactic behaviour; the principles will be argued for in section 2, the classification along

---

[1] For arguments in favour of such a separation cf., inter alia, Aikhenvald (2004), Anderson (1986), de Haan (1999a), Ehrich (2001), Wiemer (2006b), to name but a few.

these principles will be carried out in section 3, which constitutes the main body of this paper. As regards that part of the overview in section 3 which deals with grammatical marking *sensu stricto* (bound morphology, extensions of TMA grams), it can be read as a critical assessment of the state of the arts, insofar as it concerns grammatical evidentiality marking that has been being studied already for a honourable span of time.

The chosen procedure should allow describing and comparing these devices to one another independently from decisions as for whether the given marker has undergone a process of grammaticalization or lexicalization. I thereby want to avoid an involvement into still ongoing disputes concerning the delimitation of grammaticalization from lexicalization (and other phenomena of language change and their results). From the perspective taken and the purpose of a morphosyntactically oriented classification of marking devices, such an involvement is unnecessary, we need not wait until such disputes might eventually been abandoned or solved. Thus, although in section 2 I will briefly "out" my stand regarding the difference between grammaticalization and lexicalization (viz. the relation between grammar and lexicon), presently the prime purpose is a descriptive and exploratory one, namely: to work out a representative picture of form:function-correlations of contemporary languages in a delimited geographic area without any bias toward theories concerning specific processes of linguistic change.

Here Europe is conceived of as the whole geographically defined subcontinent, including its eastern edges, since I want to integrate by one better known language from the Turkic family (Turkish in Turkey), the Indo-European family in the Caucasus (Armenian), from Kartvelian (Georgian), and from Finnic (Estonian). The survey is short of data on Scandinavian (except of Swedish), it also lacks information on Celtic, Basque, Hungarian and practically all languages with minor numbers of speakers, such as numerous Turkic, Finnic and Indo-European varieties on the Balkans and the European part of Russia. Greek has not been accounted for either, as relevant insights have been presented by Stathi (this volume).

As far as I see, the present study is the first attempt at a comprehensive account of a subdomain of evidentiality (but see the preliminary study based on questionnaire data reported in Ramat 1996). I do not intend to draw any more far-reaching conclusions on areal patterns of formal classes of hearsay marking. An areal assessment would be premature given the fact that we are only beginning to understand along which parameters, both functional and structural, evidential marking ought to be estimated and classified, let alone the question what can count as a sufficiently conventionalized structural or lexical indicator of, or functional extension into, evidentiality, more specifically into hearsay. Only a few salient features from among an enormously intricate network of factors can be illustrated, others can only be "scratched" upon, many others will remain unmentioned.

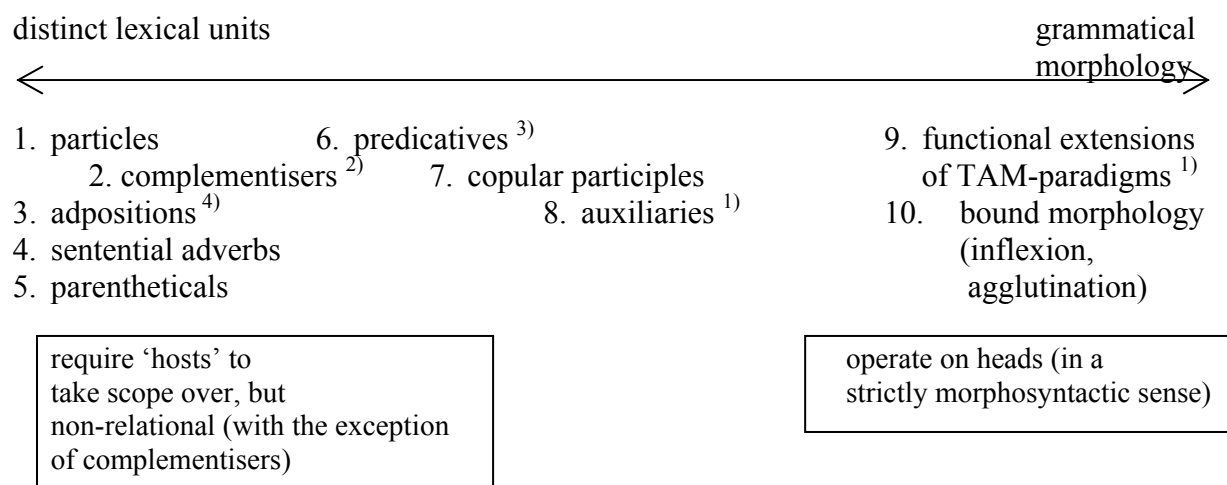## 2. Hearsay markers on a lexicon—grammar cline

What we are looking for now is a viable classification of marking devices for reportive functions comprising not only the traditionally investigated grammatical means, but also distinct classes of words carrying hearsay meaning. By "viable" I mean that there should be a unified and theoretically justifiable basis applicable to differently structured (ideally to all) languages. From the start we can exclude "trivial" means of marking hearsay like illocutive verbs, i.e., verbs with a 'say'-component in their semantic description, with an "ordinary" argument structure implying a sentential complement and various subject NPs (e.g., *People say / Somebody said that P*) as well as steady collocations, or phraseological syntagms, like *Legend has it / There are rumours (that P)* etc.

Among the first to formulate criteria for defining evidential units in both semantic and structural terms was Anderson (1986). His semantic criterion corresponds neatly to the approach advocated here: "Evidentials show the kind of justification for a factual claim which

is available to the person making that claim", as well as the following caveat: "Evidentials have the indication of evidence as their primary meaning, not only as a pragmatic inference" (1986: 274). In fact, Anderson restricted the notion of 'evidentials' to grammatical units (1986: 275): „Morphologically, evidentials are inflections, clitics, or other free syntactic elements (not compounds or derivational forms)." Whether there might be a lexical—grammatical distinction for other "free syntactic elements", he did not consider. The same holds true for de Haan, who, in refining Anderson's criteria, added that evidentials do not show agreement with the speaker and they cannot be in the scope of negation[2] (de Haan 1997: 147-150; 2000: 75f.).

Semantic (or functional) criteria like the one given by Anderson have hitherto been used to find out evidential markers from among grammatical morphemes in the narrow sense; such criteria have been generally accepted. Since they are semantic ones, nothing hinders us to apply them to units other than grammatical morphology, too, first of all to morphologically independent lexical units. Given this, we can "browse" through all sorts of morphemes in languages and then arrive at a cline of the following shape:

Figure 1:   **Types of evidential markers on a morphosyntactic cline**

distinct lexical units   grammatical
                                                                                          morphology

$\longleftarrow$ ————————————————————————————————— $\longrightarrow$

1. particles	6. predicatives [3]	9. functional extensions
2. complementisers [2]	7. copular participles	of TAM-paradigms [1]
3. adpositions [4]	8. auxiliaries [1]	10. bound morphology
4. sentential adverbs		(inflexion,
5. parentheticals		agglutination)

require 'hosts' to take scope over, but non-relational (with the exception of complementisers)	operate on heads (in a strictly morphosyntactic sense)

Remarks:
[1]   Often hearsay functions of modal auxiliaries show up saliently only in marked mood forms (see Romance modals discussed in 3.3.2).
[2]   This includes conjunctions.
[3]   Function as heads of PPs or of sentential dependents; in the latter case they combine with complementisers.
[4]   Heads of NPs (= PPs), PPs headed by them have propositional scope.

Presenting the types of units on a cline does not imply that units on one pole are the input of diachronic processes yielding an output at the other pole. The cline is meant simply as a very rough (and preliminary) representation of the fact that some morphemes (those approaching the left side) have independent word status and thus belong to the lexical stock of a language, whereas other units (closer to the right end) are bound morphemes and function as operators on heads (in a strictly morphosyntactic sense). In-between are units with an independent morphological status (in this respect, separate lexical units), which however function only as components of complex predicates (auxiliaries, copular participles) or are themselves heads of

---

[2] This restriction can be violated only under very specific discourse conditions (de Haan 1997: 154f.), which however do not bear on the argument developed in this article.

sentential arguments (predicatives). Finally, among functional units with distinct word status (on the left end) complementisers are the only class that takes part in narrowly understood dependency relations: they relate superordinate predicates to their propositional arguments which, as a rule, have a predicate-argument structure of their own. The other functional classes are used either as adjuncts (PPs headed by adpositions, sentential adverbs), or they are outside clausal syntax (particles, parentheticals). For this reason complementisers are put slightly more to the right than the other distinct function words. These distinctions should become clear in the course of the data discussion in section 3. Between different morphosyntactic classes of possible evidential markers manifold relations show up, arising mostly as the result of parallel or subsequent diachronic development from an earlier unit. Here is no place to dwell on this issue systematically, but see the remarks on 'heterosemy' in 3.9.

On the cline there is no obvious place for monoclausal constructions like NcI-constructions with certain epistemic or illocutive verbs (*suppose*, *claim* etc.) or with perception verbs (like *seem*, *appear*), which for different languages have been claimed to develop reportive meanings (as for SEEM/APPEAR-verbs see 3.6). I have deliberately excluded monoclausal (or, following generative terminology, 'raising') constructions not because they were not worth deliberating, but because even more space would be required to clarify the relation of these constructions to particles and "ordinary" sentential complements of the respective verbs (as well as the significance of constructions for evidentiality marking).

Returning to the cline, one might object that what appears as 'lexical units' on the left pole can, at least partly, be re-termed as 'function words'. In fact, usually grammars and textbooks refer to adpositions, conjunctions (complementisers) and particles under this rubric. However, what would be gained by saying that bound morphology, TMA paradigms and auxiliary constructions belong to a language's grammar as well as function words (adpositions, conjunctions etc.)? Such a qualification amounts to saying that all functional units of a language belong to its grammar. This sounds like a trivial, if not an analytic assertion. It would not supply us with a basis useful for a discrimination of different kinds of functional elements (morphemes, word classes). Instead, a useful discriminatory basis should allow for operative and verifiable decisions on whether a given item is (rather) a distinct unit of its language's lexicon (= inventory of lexemes) or (rather) a formative of more complex units on which it depends[3]. Certainly, there may be cases in which a definite and clear decision is difficult to take; but this is just a corollary of clines with focal points as the one introduced above.

I have now come to the point where I have to take a concise position as for what shall be considered as grammatical(ization) vs. lexical(ization). The one does not exclude the other, the relationship between both has to be seen as the relation of a holistic unit (morphologically either simple or complex) to constituents of a larger syntactic format. I adhere to Lehmann (2002: 1), who formulated this relationship in the following condensed way: "A sign is lexicalized if it is withdrawn from analytic access and inventorized. On the other hand, for a sign to be grammaticalized means for it to acquire functions in the analytic formation of more comprehensive signs."[4] Exactly such a withdrawal from analytic access is what happens when, e.g., a particle or a preposition evolves from petrified verb forms, regardless of whether it underlies erosion or not; compare particles like Cz. *prý* < *praví*.3.PL.PRS.IND 'they say',

---

[3] Apart from this, if we do not accept a lexicon—grammar cline based on an opposition between an inventory of lexical units (functional or autosemous) and an inventory of morphemes that are highly dependent on their immediate morphosyntactic context with which they form analytically accessible, productive complex units, other methodological problems arise: (a) How are we to distinguish whether a one-word or a multiple-word form is conventionalized holistically? (b) How are we to distinguish units of different "formats" with identical functions (in particular, in discourse)? Both problems arise with respect to particles and parentheticals.
[4] For a very similar approach cf. Himmelmann (2004).

Russ. *mol* < †*molvilъ*.3.SG.PAST.IND 'he said', various forms of the descendents of Latin *dicere* in Romance, but also particles deriving from SEEM/APPEAR-verbs (see 3.5-3.6), or the Lithuanian hearsay preposition *pasak* 'according to' < *pasakyti* 'to say' (see 3.8). Since, for instance, hearsay particles have scope over whole propositions, in order to claim that they are "grammaticalized" we must first show how the more comprehensive signs look like, which format they have and in which sense the particles can be considered as inextricable parts of these comprehensive signs. Do we want to say that since particles are operators on whole sentences (or conjunctions on clauses), they are grammatical formatives (or components) of the latter? To my mind, to do so would mean to overuse the notion of 'grammatical marker'. Independently from this, we still would have to describe every particle (conjunction, preposition, etc.) as a distinct unit of the lexical inventory of the language. This is what usually is neglected or overlooked in recent research. The following remark from Lehmann (2002: 16) applies to all functional word classes distinguished in this paper: "Those numerous complex prepositions and conjunctions which constantly come and go do not indicate incomplete grammaticalization processes, but are simply p r o d u c t s o f l e x i c a l c h a n g e." (emphasis mine, BW). This makes them no less worth thorough investigation, and it is time to start doing this research. Consequently, the following survey and remarks concerning the classification of hearsay marking devices are also meant as a step toward the creation of a cross-linguistic data base establishing and describing the inventory of evidentiality markers. From this angle, this is a problem paper which certainly digs out more questions than it supplies answers; hopefully, it at least helps to find the way toward solutions of at least some nagging problems.

3. Distribution of markers across European languages

In a sense, bound morphemes indicating hearsay, either as a central function or as a functional extension, are common in at least some parts and/or language families of Europe. If we disregard tense-aspect and mood grams which had acquired the status of inflectional endings long before they underwent evidential extensions, the number of languages with bound morphology specialising in evidential functions diminishes drastically. Within Europe we find them only on its south-eastern periphery (Balkan, Southern and Eastern shores of the Black Sea region). But again, even here in practically no case do these markers specialise in hearsay, instead, the reportive function is only a contextually conditioned reading of a broad evidential meaning potential. All bound morphemes of this sort represent agglutinative morphology rather than inflectional affixes, and almost all derive from the (present) perfect. A remarkable exception, both as for etymology and the range of evidential functions is the Estonian suffix -*vat*, which, like the predicative active participles used in Estonian and the two Baltic languages, is restricted to hearsay (see 3.1 and 3.2.1).

This overview starts with hearsay marked on, or by, the predicate (verb phrase). Bound morphology which has acquired evidential functions will be looked at first (3.1) and separately from cases in which TMA paradigms have experienced functional extensions into evidentiality, but not necessarily turned unequivocally into evidential markers (3.2).

Clitics will be dealt with in 3.5 and 3.9. I am aware that in many cases it is arguable whether a clitic has turned into an agglutinated affix. As long as the unit in question does not attach in a fixed position to a specific kind of word form (without any "material" intervening its connection with the stem), I will regard it as a clitic and not as a word-form internal morpheme.

3.1. Bound morphology: inflection and agglutination

The only example of an inflectional gram used for evidential purposes in European languages is the Georgian perfect. It forms part of a threefold series of stems (Boeder 2000: 278). In the

standard language evidential readings (hearsay or inferential) are not marked by any additional morphology (Boeder 2000: 277). It is only in dialects of Western Georgia and adjacent Kartvelian languages where we find "a richer system which provides special evidential verb forms as counterparts not only for the non-evidential aorist but also for the present, imperfect, and future" (Boeder 2000: 309).[5] If used with values of indirect evidentiality, the standard Georgian perfect allows for definite time adverbials and can thus principally be used as "propulsive tense" of a narrative (Boeder 2000: 296 and 317, f. 39). An analogous release of restrictions in evidential narratives can be observed with the Turkish *mIş*-suffix (Johanson 2000) as well as in Baltic, where perfect participles are encountered as the basic tense of narrative texts in folklore (cf. Wälchli 2000).

This suffix is probably the most frequently mentioned case of an agglutinated evidential marker in Europe. In fact, cognates of *-mIş* are attested in practically all Turkic languages. But even if we restrict ourselves to standard Turkish (being spoken on the edge of Europe), it is essential to distinguish the agglutinated suffix from the (nearly) homonymous (and cognate) copula particle *ImIş*. These morphemes differ both in their meaning potential and their distributional properties, including their impact on word prosody (cf. Csató 2000: 36-39; Johanson 2000: 80f.). A short comparison of both units will be given in 3.9. At the moment it is important to realize that the suffix *-mIş* can be added only to primary verb stems, i.e., stems that are not modified by other grammatical affixes. It is itself an anteriority marker able to focus on resultative states and therefore tightly associated to core meanings of perfect grams; this obviously gave rise to its contemporary meaning of 'indirectivity' (according to Johanson 2000), among which hearsay and inferential interpretations figure prominently. The connection with the perfect is never dissociated. Since *-mIş* can also be used for different kinds of inferential judgements, we can qualify it as a general marker of indirect evidentiality. In this respect the meaning potential of *-mIş* closely resembles the Georgian perfect (see above) as well as Armenian, Balkan Slavic (see 3.2.1) and Albanian (see below).[6]

Another agglutinated suffix exists in standard Estonian, namely *-vat*. In contrast to Turk. *-mIş* (and its cognates in other Turkic languages), Est. *-vat* is totally unrelated to the perfect system; it originates from a merger of the marker *-v(a)* of the present active participle and the partitive ending *-t* (Kehayov 2002: 129); compare, e.g., *Ta tule-vat* 'S/he has come, they say' (against the non-reportive past *Ta tule-b* 'S/he has come/came'). Another difference in comparison to Turkic (and Albanian *kam* 'have', see below) is the restriction of *-vat* to hearsay, i.e., it is not used in inferential function. This parallels the outstanding functional specialization of active participle constructions in Estonian and the Baltic languages (see 3.2.1).

The last language to be mentioned in this subsection is Albanian. Obviously, it is the only European language in which the functional extension from perfect to evidentiality has to some degree been accompanied by coalescence (in the sense of Lehmann 1995): the inflected HAVE-auxiliary agglutinates to the abridged participle, and it does so in reverse order compared to the perfect proper[7], which is still productively used as an analytical tense:

---

[5] Further analyses on such a system in Kartvelian and neighbouring Caucasian languages can be found in Friedman (2000: 353-355) and Topadze (2008).

[6] Actually, the functional similarity of grammatical evidentiality marking and its tight connection to the perfect in the languages of this large area are obviously a result of areal convergence. It is very likely that the Turkic system rendered the model on which evidential extensions of the perfect spread via contact into adjacent language varieties.

[7] Megleno-Romanian, which is spoken in the vicinity of Albanian (in the Greek-Macedonian border region), also uses an inverted perfect with an inflected (agglutinated or cliticized ?) HAVE-auxiliary as an admirative (Friedman 2003: 190). As for Aromanian, however, only the Frasheriote dialect spoken in southwestern Macedonia appears to have developed an admirative marked by the invariant "particle" *-ka*, borrowed from Albanian *kam* '(I) have' and attached to the inherited Romance HAVE-auxiliary or the inflected verb in the present tense (Friedman 2000: 347f.).

(1)   **Albanian**

   <u>perfect</u>   <u>admirative</u>

   kam   *hapur* 'I have opened'   *>*   *hap-kam* 'I have opened' (ADM/EVID)[8]

   have.PRS.1.SG open.PART   open.have.PRS.1.SG

Albanian is outstanding as it is the only language of Europe (including its "edges" in the Black Sea region) in which the original perfect auxiliary incorporated so intimately with the verb stem (the participle) that this morpheme combination has been reinterpreted into a marked form of the p r e s e n t tense and given rise to a whole set of admirative (and evidential; see f. 8) paradigms combinable with any tense and mood except the aorist (cf. Breu 2009: 3.3 and Friedman 2000: 342f. for details). Neither Balkan Slavic, nor Turkish, nor standard Georgian, nor Armenian show verb morphology marked for evidentiality in the present tense. This contrasts with Georgian dialects (see above), Estonian and Baltic (see 3.2.1).

3.2. Functional extensions of TMA-paradigms

Since the TMA domain comprises an agglomerate of all kinds of grams with different combinations of tense, aspect and mood distinctions, I will divide these devices of hearsay marking into subtypes: first come grams dominated by tense-aspect components, then follow grams for which mood distinctions are central. The tense-aspect part further splits into perfect grams, future grams and past grams; the last mentioned are the most problematic ones (see 3.2.3). A decision as to whether these extensions have become new meanings of the paradigm or whether they are to be treated just as evidential strategies in Aikhenvald's sense, cannot be taken in a wholesale manner, but must be considered for every single case.

3.2.1.   Extensions from analytic perfects

Beside agglutinative perfect grams (Albanian and Turkish) and the inflectional perfect of Georgian, treated in 3.1, evidential extensions are attested for languages with perfect grams of an analytical structure based on participles. These are: (a) the extant Baltic languages (Lithuanian, Latvian) and some Finnic languages, first of all Estonian (but not Finnish!), (b) Balkan Slavic (Bulgarian, Macedonian), (c) West Armenian. The arrangement of these languages under (a-c) is meant to indicate that they belong to three areal clusters, noticeably all of them on the Eastern periphery of Europe. At least Balkan Slavic and Armenian form part of a much larger convergence area (see f. 6).

   Armenian has to be treated with some reservation, insofar as Modern East Armenian has lost its core perfect character, since it can be used also "in narration as a nonevidential past tense" (Kozinceva 2007: 81; cf. also Kozintseva 2000: 407f.). This does not apply to Modern West Armenian, in which the perfect is not used as a narrative tense and closely resembles both the paradigmatic and functional distinctions of Balkan Slavic: their perfects have not lost their character of typical perfects (Donabédian 1996: 88f.).[9]

   As concerns Balkan Slavic, Bulgarian and Macedonian conform to each other in almost every respect; their "dialects show considerable morphological variation in their treatments of

---

[8] For the classification pursued it is of no importance whether the Albanian evidential forms should better be qualified as 'admiratives' (or 'non-confirmatives', in Friedman's terms) and whether the (ad)mirative function has to be regarded as an evidential one. For different views on the relation between evidentiality and (ad)mirativity in Albanian cf. Breu (2009: 4.3) vs. Friedman (2000: 342f.).

[9] The main difference with regard to Lithuanian (and Latvian) lies in the fact that, contrary to Armenian, Baltic has no aorist-imperfect system, only a general past. The paradigmatic and functional link between perfect and evidential usage is however the same, as is the "ban" on the perfect as a narrative tense (unless in re-narration, when hearsay applies).

evidentials [as do varieties of other Balkan languages; BW] (…), but the underlying semantics are fairly consistent" (Friedman 2003: 192). The Balkan evidential system has sometimes been characterized as a highly "epistemicized" one, i.e., that an epistemic judgment usually accompanies the evidential function (Plungian 2001: 354); others, like Friedman (2000; 2003), maintain that it is the "speaker's attitude" toward the information conveyed, not reference to the "source", that is the decisive motif behind these forms. The problem however is that time and again investigators have come up with examples and context types in which no necessary epistemic load has to be imputed into the Bulgarian (or Macedonian) forms (see, for instance, ex. 2, where epistemic connotations are absent). This notorious bone of contention has remained extremely confusing. Whatever stance one might take towards this issue, hearsay is only one possible interpretation of the paradigmatic forms referred in Bulgarian grammar as 'preizkaznieto naklonenie' ('renarrative mood'), in Macedonian grammar as 'prekažuvanje' ('renarration'). Thus, by an utterance like the following one the speaker indicates only that s/he has not witnessed the described event him/herself; out of context the hearer is free to interpret the form *peel* 'sung' as either inferential or reportive:

(2) *Djado*   *mi*   *(e)*   **pe-e-l**   *v* *cărkva-ta*.
      grandpa   me.DAT   COP.PRS.3.SG   sing.IMP.PPA:SG.M   in church.DEF
   i. '(My) Grandpa **is said to have sung** in the church (choir).'   → hearsay
   ii. '**Obviously**, (my) grandpa **sang** in the church (choir).'
                 → inferential, or perfect ('indefinite past')

The form of the predicate *peel* goes back to the Common Slavic *l*-participle, originally the active participle of the indefinite past used as a perfect. In most Slavic languages this form has ousted the synthetic preterits (aorist, imperfect) and changed into a general past tense, but in Bulgarian and Macedonian the paradigmatic opposition between perfect, aorist and imperfect has survived without restrictions. The meaning of the (indicative) perfect often interferes with evidential readings (plus a possible mirative one); absence vs. presence of the copula (*e* in ex. 2) is no reliable criterion (see below).

There is however a point concerning the formal expression of the functionally extended perfect, which makes it clearly distinguishable and which can likewise be illustrated from ex. (2). In contrast to Georgian, Armenian and the Baltic languages, Bulgarian and Macedonian have introduced a new combination of the perfect marker (the *l*-suffix) with a derived stem of the imperfect. The "evidential" participle *peel* is made up of the imperfect stem *pe-e-* plus the *l*-participle. Both suffixes (i.e., an allomorphic *'a/'e*-alternation for the imperfect expanded from the infinitive stem preceding the *l*-suffix) belong to common Slavic heritage, the c o m b i n a t i o n of both does not. This innovation occurred relatively late and was conditioned by the establishment of a series of evidential forms (for a simplified, but instructive table showing the paradigm structure cf. Friedman 2000: 330).

The two main differences between Bulgarian and Macedonian are the following: (i) in the analytic tense Macedonian never uses a copula in the third person, and it uses the imperfect forms of the copula (*bev*.1.SG, *beše*.2.SG, etc.) in the evidential ('non-confirmative') series, whereas Bulgarian uses the present tense forms (*săm*.1.SG, *si*.2.SG, etc.); (ii) Macedonian has developed a resultative perfect with the HAVE-verb (*imam*.PRS.1.SG + indeclinable participle) whose *l*-participle (i.e., actually the perfect) forms the basis of an additional evidential paradigm (e.g., *Si imal napraveno* '[Apparently] you did / have done (it)'). The rise of the HAVE-perfect began in the Southwestern area (Ohrid, Struga) and was obviously triggered by the model of the Albanian HAVE-perfect (see 3.1; cf. also Breu 1987: 293f.).

In sum, a functional characteristic concerning most languages mentioned so far is that the reportive function of the perfect can be qualified only as a subfunction within indirect evidentiality (including admirativity). Exceptions are the northern-most languages, Latvian

and Estonian as well as Lithuanian, as long as active participles are concerned (see below). Furthermore, in all these languages most (in standard Georgian even all) forms of the indicative tenses are "homonymous" with the 'evidential series'. Neither in Baltic, nor in Balkan Slavic, nor in Georgian the "copula criterion" stated by normative grammars stands the test of empirical distribution in discourse. This fact aggravates the weak paradigmatic distinction.[10] Nevertheless the distinction from the respective perfect paradigms is greater than in the case of mood or aspect-tense paradigms with evidential extensions in other European language groups (see the following subsections). Finally, in every case paradigmatic weakness corresponds to functional weakness: the forms are not obligatory (cf. also Friedman 2003: 193, 209f. on Balkan Slavic and Albanian).

Now let us have a closer look at the Baltic languages and Estonian. All three languages possess series of active participles from the present and past – in the Baltic languages also from the future – stem, with which the copula is normally absent or itself coded in a non-finite (participle) form (see, for instance, f. 13, also Wiemer 2007a: 207-210). In this sense we may say that these languages display greater "paradigmatic consistency" over tenses than Balkan Slavic, Georgian, Armenian and Turkish, in which participle-based evidentiality marking is restricted to forms deriving from past tense and present perfect. Diachronically, in Baltic the present-tense based verb forms with reportive functions do not derive from reinterpretations of the perfect; the Baltic and Estonian present tense evidentials should rather be explained by a converging effect of two factors: (i) analogy with an evidential reinterpretation of the perfect (which is formally identical with the Balkan Slavic one) plus (ii) logophoric constructions which by time acquired independent syntactic status (for the details cf. Wälchli 2000: 191-197; Wiemer 1998: 233-239; 2007a: 228-234). As a consequence, the forms used in Baltic and Estonian vary quite considerably in the non-past domain: Latvian and Estonian are actually more similar to each other than the genetically closely related Latvian and Lithuanian (cf. Holvoet 2007: 81-91 and Wälchli 2000: 188-190 for details). This, by the way, clearly indicates that language contact has played a considerable role in the spread of evidential marking by grammatical means in the eastern part of the Circum Baltic Area. Compare the following examples which illustrate the past evidential, whose form is closest to the perfect indicative:

(3) **Lithuanian** *Tėvas   sugrįžęs   iš miško.*
father.NOM.SG.M   return.PPA:NOM.SG.M
'Father apparently (they say) has returned from the forest.'

(4) **Latvian** *Viņš   dzīvojis.*
he.NOM   live.PPA:NOM.SG.M
'He has lived, they say.'

(5) **Estonian** *Ta   elanud.*
s/he.NOM   live.PPA
'S/he has lived, they say.'

There is only one basic formal difference of Estonian as compared to Latvian and Lithuanian: the active participle functioning as evidential marker is uninflected for those nominal categories which are usually marked on Estonian verbs.

Different kinds of reinforcement by a doubling of participles on the copula occur (as they do in Balkan languages, too); cf. (Holvoet 2007: 97-99; Wiemer 2007a: 215). Forms based on a pluperfect are also created, but their time reference quite often does not differ clearly from

---

[10] As for Bulgarian cf. Friedman (2000: 334-338), and Kehayov (2002: 127f.), as for Lithuanian Wiemer (1998: 232f.; 2006a; 2007a: 201-210), as for Georgian Boeder (2000: 285-294).

the time reference of reportive forms based on the p r e s e n t perfect (cf. Kehayov 2002: 129f., 136f. on Estonian).

Lithuanian (more precisely, standard Lithuanian) is specific insofar as it has two different participial constructions which tend to be distributed nearly complementarily in terms of evidential subdomains. It thus differs from the two other languages both functionally and with respect to concurring constructions. All participle constructions illustrated in (3-5) are basically restricted to hearsay. But Lithuanian has an alternative construction with an invariant participle for the present and the past (or perfect) tense, ending in -*ma* (simultaneous to reference time) and -*ta* (anterior to reference time), respectively:

(6) *Krūmuose   vištų.   **perėta** / **perima**.*
     bush.LOC.PL hen.GEN.PL brood.TA / brood.MA
     'In the bushes (some) hens **(must) have brooded** / are [**must be**] **brooding** their eggs.'

Sentence (6) could be uttered after the speaker has found a couple of eggs in the bushes, but does not see the hens; s/he thus concludes that hens must have sat there. This construction is used almost exclusively for inferential evidentiality, it is therefore in (almost) complementary distribution with the agreeing participle construction illustrated in (3). From the structural point of view this construction does not have any agreeing NP, instead the highest-ranking (= most agent-like) argument must be coded with the genitive (*vištų*).[11] Latvian and Estonian do not have a specialized construction for inferential evidentiality, nor do the other European languages with evidential extensions of perfect paradigms.

There is however another detail worth mentioning. Latvian and Estonian can make reportive marking more explicit by using an uninflected form of the copula BE.[12] Actually, this is just an expansion of the constructions in (4-5): Latv. *Viņš **esot** dzīvojis*, Est. *Ta **olevat** elanud* 'He lived, they say'. This construction likewise is restricted to hearsay function. In Lithuanian *esą*, the lexical and morphological cognate of Latv. *esot*, is used for this purpose, too:

(7) *Jis   **esą**   gyvenęs   Vilniuje dešimt metų.*
     he.NOM   COP.INDECL live.PPA:NOM.SG.M
     'In Vilnius he (has) lived, they say, 10 years.'[13]

In this case, Lith. *esą* sort of reinforces (or stresses) the reportive value (whatever that may mean), as do its Latvian and Estonian „cousins". But, contrary to them, *esą* can occur alone and behaves like a particle, sometimes even like a conjunction or the predicate (cf. Holvoet 2007: 85f.; Wiemer 2007a: 217-220; see 3.6, 3.7 and 3.9). This makes it differ from Latv. *esot* and Est. *olevat* in grammatical and lexicographic terms. We may say that the particle and complementiser uses of Lith. *esą* are the result of the loss of paradigmatic relations with the inflected forms of the copula proper (including its inflected participles!), whereas in Latvian and Estonian this relation obviously has not been cut off.

---

[11] For a systematic account cf. Wiemer (2006a; 2007a: 213-216).

[12] In Latvian, the future form of this uninflected participle (ending in -*šot*) is used, too (cf. Holvoet 2007: 82f.).

[13] The temporal value of such a sentence varies, as it corresponds either to the past or to the present perfect. Informants confirm that *esą* combines with a finite verb or participle in any tense. Perfect evidentials (reportives) are characterized by the declinable (agreeing) form *esąs*.M, *esanti*.F etc. (e.g., *Jis **esąs** gyvenęs Vilniuje* 'He **has lived** in Vilnius, people say'), but these are very uncommon in colloqial speech (informants reject them) and seem to be restricted to some dialects.

3.2.2. Extension from future gram

Future markers, in particular grams that can be classified as future perfect (*futurum exactum*), as indicators of inferential evidentiality are widespread in Europe (as elsewhere). They are commonplace in Romance, Germanic, Baltic and South Slavic. But practically nowhere do they expand into the domain of hearsay, the only exception being Portuguese. As Squartini (2001: 319), from whom the following example is quoted, writes, this is „a case which (...) is unique in Romance and has never been pointed out before". Compare one of his examples:

(8)   *Segundo fontes que lhe são próximas Soares **terá dito** a Gomes Motas e a Carlos Monjardino que as criticas à liderança de Guterres foram "pura locura".*
       'According to sources which are quite close to him S. **said** [lit. *will have said*] to G.M. and C.M. that the criticism of G.'s leadership was pure nonsense.'

Obviously, this case is unique not only in Romance, but even on a European background.

3.2.3. Extensions of past tense grams

Squartini (2001: 308-314) gives a conscientious account of usage types of the Italian imperfect (*Imparfetto*) which have repeatedly been considered to carry reportive meaning: "the evidential value expressed by the Imperfect is a form of report of what the speaker knew was due to happen and about which confirmation is lacking". An important condition necessary for the Imparfetto to acquire this function is that "the reported information refers to a scheduled situation" (Squartini 2001: 311). Examples which corroborate this conclusion are utterances in which the speaker presupposes knowledge shared with the interlocutor. It is important that aspectual values usually ascribed to the Imparfetto can be cancelled, as in (9), where the adverbial *fino alle sei del pomeriggio* 'until six o'clock in the evening' delimits the time span and thereby normally requires the Passato Remoto (= aorist) or Passato Prossimo (= perfect or generalized past):

(9)   *Paolo quel giorno **lavorava** fino alle sei del pomeriggio.*
       'On that day Paolo **was to work** [lit *worked*.IMP] until six o'clock in the evening.'

Here the Imparfetto can but convey reportive meaning, which arises by implicature from the fact that the speaker possesses only indirect knowledge about the described state of affairs and that s/he could obtain this knowledge only via hearsay. What further corroborates Squartini's conclusion is the fact that the assumed reportive use of the Imparfetto patterns exactly like the conditional, which in modern Italian conveys only a reportive, but no inferential function (see 3.2.4). However, despite this clear parallel and the fact that the aspectual core meaning of the imperfect can be overridden, there remains the question whether we can ascribe to the imperfect really a new conventionalized meaning, and not just a contextual implicature.[14]

This question can be answered in the affirmative for the Estonian modal auxiliary *pidama*,[15] whose past tense form can acquire reportive meaning (see 10a). According to Erelt

---

[14] Such an implicature can be detected in other languages as well, regardless of their tense-aspect system. See, for instance, utterances characteristic of the German past tense (*Präteritum*, aspectually indifferent) formulated as questions like those adduced by Squartini: *Wer **bekam** das Schnitzel?* 'Who is to get the escalope?' (lit. 'Who **got** the escalope?'), which covertly refers to a speech act (ordering of a meal). A reportive implicature is also possible with Russian (or Polish) imperfective aspect (both in present and past tense), e.g. Russ. *(Ja znaju, čto) ix poezd **otpravljalsja** v vosem' časov* '(I know that) their train **was to leave** at eigth o'clock' (lit. *left*.IPF.PAST). Like the Imparfetto, this function of imperfective verbs (often called 'prospective') is typical of scheduled situations, but, contrary to the Imparfetto, it is compatible with knowledge deriving from speaker's eyewitnessing (*...i ja videl, kak poezd uxodil* '...and I saw how the train was leaving').
[15] The stem is *pida-*. The form used here represents the so-called *ma*-infinitive.

(2001: 16), this functional extension is to be observed from the fact that past tense[16] forms of *pidama* (+ infinitive) "can be replaced by the form of the morphological quotative" (= reportive), as in (10b):

(10a) *Praegugi* **pid-i** *neid* *mitme-s* *koha-s* *maa-s* **vedele-ma**.
     now_too must-PST.3SG they.PART many-INE place-INE ground-INE lie-INF.

(10b) *Praegugi* **vedele-vat** *neid* *mitme-s* *koha-s* *maa-s*.
     now_too lie-PRS.REP they.PART many-INE place-INE ground-INE.

    'Even now some of them [little slips of papers] **are reported to be lying** on the ground in various places.'

As Erelt (2001) stresses, only the past tense of *pidama* can be used with hearsay meaning, at least in contemporary standard Estonian. Usage patterns typical of the older written language and some Northern dialects show also indicative present forms in this function, which have been related to German influence (compare the modal *sollen*; see 3.3.2). *Pidama* can occur together with other grammatical forms that are associated with reportive meaning, first of all with the *vat*-suffix (see 3.1); cf. Erelt (2001: 16f.), Kehayov (2002: 136). Notably, the past tense form *pidi* evokes an imminence effect ('was about to / on the verge of'; cf. Erelt 2001: 12-15). This effect can probably be linked with the reportive function on the basis of intentional stage and presupposed knowledge, which have been shown to be crucial components for a reportive interpretation of the Italian *Imparfetto*, too.

3.2.4. Extensions from moods

Let us now look at hearsay extensions of mood grams. A clear case in point is German, which uses the subjunctive (*Konjunktiv I* and *II*) for this purpose. Typically it occurs in the news, as in the following example with the *Konjunktiv I* from a TV report (January, 13th, 2007):

(11) *Angela Merkel hielt sich bedeckt // die **arbeite** gut und gerne mit Stoiber zusammen //*
     'Angela Merkel remained guarded//she **is said to cooperate** well with Stoiber.'

Here the subjunctive form *arbeite* (vs. indicative *arbeitet*) serves as an indicator that the author of this utterance did not produce the assertion 'AM likes to cooperate with St.' himself. *Konjunktiv II* works the same way; actually the only difference between both subjunctives lies in (relative) time reference: *Konjunktiv II* is based on the pluperfect and thus can always be recognized from the auxiliary (*hätte* vs. *hatte* and *wäre* vs. *war*), whereas the forms of *Konjunktiv I* are to a large part homonymous with the indicative (present and perfect);[17] cf. Roels et al. (2007).

    The hearsay function of the German subjunctives most clearly arises in unembedded clauses (see ex. 11), where it developed from its original, merely syntactic function, namely: as an indicator of indirect speech. This diachronic background certainly explains why the evidential extension of the subjunctive is restricted to hearsay and does not include inferentiality. The same holds for some dialects of Retoromance, in particular for Surselvan, where the subjunctive is exploited as a means of marking hearsay in main and subordinate

---

[16] Traditionally, and confusingly, this tense has been called "imperfect"; notice that it does not bear any specific aspectual value.

[17] This formal (and diachronically motivated) difference is usually neutralized, since in real discourse the time reference of these forms is indistinguishable (cf. Thieroff 1992: 223-227). In this respect, the form:function-relations of the subjunctive paradigms in contemporary German very much resemble the situation in Estonian and Baltic (see 3.2.1).

clauses (Grünert 2003: 536-540). This behaviour clearly departs from the usual pattern in Romance (see below) and can most probably be explained by contact with German.

In Dutch there is no formal equivalent of the German situation, since Dutch does not have a subjunctive, or conditional, as an inflectional category of the verb. Instead, an analytical device is exploited, for which see 3.3.1.

In Romance the situation differs from German in some significant respects. In these languages it is the conditional which is the basis for evidential extensions; and in most Romance languages these extensions are not restricted to hearsay, but result in an undifferentiated indirect evidential function (see section 4). See the following examples (quoted from Squartini 2001: 318, 306):

(12) **Portuguese**
*Segundo certas vozes, teu pai **teria reunido** a esta altura um bom pecúlio.*
'According to some rumors, your father **had saved** [lit. *would have saved*] a lot of money at that time.'

(13) **French**
*Aux dernières informations, les concurrents **auraient franchi** le Cap Horn.*
'According to the latest information, the competitors **have round**ed [lit. *would have rounded*] Cape Horn.'

Among the contemporary Romance languages Italian stands out, insofar as its conditional can no longer be used with inferential function, being restricted to reportive evidence. However, this situation must have obtained only recently, because during the 19th century the Italian conditional could still mark inferential evidentiality. Modern Italian appears to be almost unique among contemporary Romance, since it no longer shows an overlap between the evidential functions of the future tense (→ inferential) and the conditional (→ reportive); cf. Squartini (2001: 307f., 314f., 324-327). It is only in Italian where we observe a clear-cut complementary distribution between mood and tense grams: the French and Portuguese conditionals allow for both interpretations, whereas in Spanish (and Catalan) the conditional as a means of marking evidentiality seems in general to be stigmatized, and the Romance future is generally restricted to inferential evidentiality (with the remarkable exception of Portuguese; see above); cf. the systematic and comprehensive study Squartini (2001).

Only Romanian comes close to the Italian distributional pattern (Squartini 2005: 252), provided we account for the additional complication that in Romanian both conditional and future can combine with the presumptive. This is a basically modal category which in Romanian can be exploited for evidential purposes, but obviously only as a strategy that allows for either reportive or inferential interpretations. Its formal expression is based on the combination of a gerund of the lexical verb with an indeclinable (petrified) form of an auxiliary verb usually called a (modal) particle; the gerund usually derives from present tense (*-înd*), although many researchers admit that past gerunds are possible, too (cf. Friedman 2000: 350f., Squartini 2005: 252-263). This formal structure bears similarity to the structure of the Baltic and Estonian active participle constructions, which can be based on the present tense, too (see 3.2.1).

3.3. Auxiliaries
Similarly to TAM-paradigms, this subsection divides into tense-aspect auxiliaries, mood auxiliaries, modal auxiliaries (briefly: modals) and volitional verbs. However, within the domain of tense and aspect the only auxiliaries relevant would be copulae as components of analytic perfects. These were discussed already in 3.3.1, where we saw that the distinction between perfect (indicative) and evidential (reportive) paradigms tends (or, is claimed) to

hinge on the presence (→ perfect) vs. absence (→ evidential) of the copula. We may therefore immediately turn to mood auxiliaries.

3.3.1. Extensions from mood auxiliaries

Cases in point exist in Germanic. Germ. *würde* (+ infinitive) is used as an analytical substitute of the subjunctive (see 3.2.4). Historically this is the morphological past subjunctive of *werden* 'to become' (Roels et al. 2007: 189; Smirnova 2006). We can regard *würde* as an auxiliary, because it has a defective paradigm (no infinitive or other infinite forms, only present tense forms) and combines only with the infinitive of the lexical verb. This combination can be considered as an analytical mood, which has more and more been replacing the simple subjunctive (discussed in 3.2.2) both in sentential arguments of illocutive verbs (indirect speech) and in syntactically independent sentences, in which it clearly marks hearsay.[18] Compare the following example, in which the *würde*-construction cooccurs with the *Konjunktiv I* (underlined):

(14)   *Vor allem der Kommunikationsgedanke stehe im Vordergrund, betonte auch Jeanette Klee. Schließlich <u>gebe es</u> einige ältere Menschen, die zum Teil auch <u>nicht mehr in der Lage seien</u>, für sich selbst zu kochen, oder Gesellschaft während des Essens **bevorzugen würden**.*
'Above all the idea of communication is in focus, also Jeanette Klee emphasized. <u>It is said that</u> after all <u>there are</u> some elderly people who <u>are no longer capable</u> of doing the cooking for themselves, or who **would rather prefer** to take their meals in company.'
(Mannheimer Morgen, 17.12.2004)

In Dutch we find a structural equivalent which functions in a similar, though not identical manner. The auxiliary *zou(den)*[19] is the usual device to form an analytical conditional; morphologically these are the past tense forms of the auxiliary *zullen*, which is etymologically related to Germ. *sollen* and Eng. *shall*. As a conditional auxiliary *zou(den)* is used both in main and subordinate clauses. Thus syntactically it behaves like Germ. *sollen*. *Zou(den)* is also used as a marker of reported speech, again as is Germ. *sollen*, which however is not a mood marker, but a modal auxiliary (see 3.3.2). As a mood marker *zou(den)* parallels Germ. *würde*, and it also serves as a reportive marker; but, in contrast to *würde* (and *sollen*), *zou(den)* indicates reported speech only in main, not in embedded clauses; cf.[20]

(15)   *Hij **zou** de moordenaar van het meisje zijn.*
        'He **is supposed to** be the murderer of the girl.'

In sum, as reportive marker German *würde* + infinitive is less restricted as for its admissible grammatical contexts and, to this extent, can be considered as more grammaticalized than *zou(den)*, which, thus, shares with Germ. *würde* and *sollen* only their semantic restrictions in

---

[18] Since *würde* + infinitive is connected not only with the subjunctive, but also with the future (*werden* + infinitive), its paradigmatic status in the German tense-mood system has been an issue of constant debates, as has the question to which degree this auxiliary construction can be considered a reportive marker (as opposed to the mere syntactic function of a subjunctive). The most recent overview of research is given in Smirnova (2006: 19-40). In conclusion of her primarily diachronic investigation Smirnova subsumes that the evidential (i.e., reportive) reading of this construction is not as grammaticalized as its subjunctive function, but should nonetheless be figured out as a distinct function (2006: 334).

[19] The form *zou* is for the singular, *zouden* for the plural.

[20] The only case when *zou(den)* can be imagined in indirect speech are sentences in which the content of speech refers to an interval posterior to the original speech act. Thus, a German sentence like *Er sagte, er **würde** nie mehr wiederkommen* lit. 'He said he **would** never return again' can be translated with *zou(den)*: *Hij zei dat hij nooit nog **zou** terugkomen*. This however is only a manifestation of the regular grammatical correspondence between future tense (in direct speech: *Ik **zal** nooit terugkomen* 'I **shall** not [= don't want to] come back') and past conditional/subjunctive in indirect speech (T. Mortelmans, p.c.).

the domain of evidentiality (i.e., only hearsay). Its categorial function as mood auxiliary differs from its German cognate *sollen* (which is a typical modal), but overlaps considerably with *würde* (cf. the systematic Dutch-German comparison in Roels et al. 2007).

3.3.2. Extensions from modal auxiliaries

This pattern is fairly well attested in Germanic and Romance. The modals concerned in Germanic all belong to the field of deontic necessity (MUST, SHOULD), whereas in Romance modals of possibility (CAN) are also encountered. Furthermore, evidential extensions of German (but not Dutch) modals are clearly restricted to hearsay, whereas in Romance they throughout lead to undifferentiated indirect evidentiality. Another difference between German and Romance lies in the fact that in Romance evidential meanings of modal auxiliaries are very intimately bound to grammatical mood, since it is first and foremost their conditional forms that carry such meanings (B. Staib, p.c.), in particular a reportive one. In German, by contrast, reportive extensions of mood and modal auxiliaries can be kept apart neatly. This will be illustrated in the following.

Germ. *sollen* is used not only as a deontic modal ('should, ought to'), but also as a marker of hearsay. This use however is restricted to the present tense forms[21] (see 16-20). The forms *sollte(n)* – morphologically the past tense of *soll(en)* – can be used only to mark a conjecture (or assumption), apart from a deontic use.[22] Potential ambiguities with a deontic reading are dissolved[23] in accordance with the (un)controllable character of the denoted state of affairs: if the intended illocution refers to a controllable situation, the deontic meaning prevails, if the situation is conceived of as uncontrolled (i.e., as a simple description), the reportive reading comes to the fore (for a detailed analysis of context factors cf. Diewald 1999: 278-282). Compare (17a-b): without context, (17a) is ambiguous, but the slight change in (17b) forces a controlled action, so that the hearsay reading is ousted (or at least heavily backgrounded):

(16) *In China **sollen** sie ganz andere Feiertage haben.*
    '**It is said that** in China they have totally different feasts.'

(17a) *Er **soll** im Bett liegen.*   (i)   'He **is said to** be lying in bed.' → hearsay
                        (ii)   'He **has to** lie [= stay] in bed.' → demand

(17b) *Er **soll** im Bett liegen <u>bleiben</u>.*
    'He **has to** stay [lit. <u>stay lying</u>] in bed.'   → demand (deontic) / ~~hearsay~~

Conversely, if *sollen* combines with an anteriority ("perfect") infinitive, which marks an event prior to, or a resultant state simultaneous to, reference time, a reportive reading is triggered, a deontic blocked (see 18). Notice, therefore, that this distribution is not restricted to telic verbs, since it does not hold solely for resultative states (as in 18), but applies to any situation anterior to the time of reference, as shown by (19-20):

(18) *Er **soll** liegen geblieben sein / hingefallen sein / ein schönes Bild gemalt haben.*
    'He **is said to** have remained lying / to have fallen / to have painted a beautiful picture.'

---

[21] From a diachronic viewpoint the contemporary present tense of modals derives from the paradigm of former *praeteritopraesentia*. However for the present classification this is irrelevant, since these forms were reinterpreted as a special present tense conjugation a long time ago.

[22] Corpus examples with *sollte* like those adduced in Faller (2006: 17) are no proof against, since here *sollte* goes together with the sentential adverb *angeblich* 'allegedly'; without the latter a reportive reading is hardly available. Actually, the difference in functional distribution yields as an argument for considering *sollte(n)* not as the past tense of *soll(en)*, but as a separate lexeme.

[23] Disambiguation is possible with negation, because the deontic reading requires external negation (i.e., negation includes the modal), whereas with evidential (reportive) reading internal negation holds (i.e., negation is in the scope of the "evidential modal"). Cf. the thorough analysis on this matter by Ehrich (2001).

(19)   *Er **soll** (schon einmal) Schach gespielt haben.*
       'He **is said to** have played chess (once before).'

(20)   *In früheren Jahrhunderten **sollen** die Menschen mehr auf dem Land gearbeitet haben.*
       'In earlier centuries, people **are said to** have worked more in the country.'

Sollen never implies an inferential meaning, i.e., it patterns complementarily in comparison to its superficial past tense form (see f. 22). *Sollen* can be used with reference to anonymous, generalized or contextually unspecified original speakers as well as to specified speakers mentioned in the closest context (Mortelmans 2000: 134). In this respect, it is slightly more restrictive than the subjunctive, which can be used even if the actual speaker refers to an earlier assertion of him/herself (Diewald 1999: 229f.).

A deontic-reportive polysemy characteristic of Germ. *sollen* is observed also for the Polish, Czech and Slovak verb 'to have' (Pol. *mieć*, Cz. *mít*, Slov. *mat'*) used as an auxiliary. We encounter the same restriction to reportive (and exclusion of inferential) meaning; most probably the functions of HAVE were calqued from Germ. *sollen*. Curiously, in all three West Slavic languages the HAVE-verb can be used as hearsay device also in its past tense forms (21), in contrast to German. Cf. an example from Polish (cf. also Hansen 2001: 137f. as well as Rytel 1982: 49f. on Polish and Czech):

(21)   ***Miał** zapomnieć swój parasol na dachu samochodu.*
       'He **is said to have** forgotten his umbrella on the roof of the car.'

However, the Upper Sorbian HAVE-verb (*měć*), which also "copies" the deontic-reportive polysemy of Germ. *sollen*, shows the same restriction to present tense, if used as a hearsay device (Faßke 1981: 77f.). Here the parallel with the German modal is even stronger and corroborates the assumption that the reportive function of West Slavic HAVE-modal has been calqued from German *sollen*.

Let us now come back to Germanic. On the background of the data presented here it is worth remarking that in modern English no single modal belonging to the former series of *praeteritopraesentia* (see f. 21) shows hearsay extensions. Notably, Old English did know *sculan* 'should' as a reportive marker, although this usage was predominantly bound to subordinate clauses embedded in contexts of overt speech acts (Traugott 1989: 41f.). Since in this paper we are not concerned with diachronic issues, I am not going to dwell further on the relation of OE. *sculan* to Eng. *should*. There is however a secondary modal, with a different etymological background, which in modern English has been gaining hearsay function, namely: *be supposed to*; cf. a short remark in Chafe (1986: 268) and the corpus-driven analysis by Moore (2007). Her analysis makes one understand that a possible reportive reading of utterances like

(22)   *The Brownings **were supposed to** be in Italy at that time.*

arises from the speaker's (writer's) wish to evoke „an indeterminate but external source for the belief or expectation" (2007: 119).

In 3.3.1 I mentioned the hearsay use of the Dutch mood auxiliary *zullen*, a cognate of Germ. *sollen*. Notice that as for Dutch only the past tense forms *zou(den)* are used reportively, whereas with Germ. *sollen* reportive use is restricted to the present tense forms. These cognates are therefore diametrically opposed in terms of their categorial distribution. From this perspective the Swedish cognate modal *ska* occupies an intermediate position, behaving in a way slightly closer to its German than to its Dutch cousin. *Ska* is the present tense form and used as a reportive marker. According to Kronning (2007: 300-303) its past tense form *skulle* can likewise be used for the purpose of hearsay, this reading proves however to be

difficult to get – contrary to *ska* – if there is neither a lexical specification of hearsay in the context nor any such hint in the speech situation (2007: 302f.). Evidently the past tense form does not by itself mark hearsay, but *skulle* is compatible with such a reading, similarly to German *sollte(n)* (see f. 22). Following Kronning's exposition, both tense forms, *ska* and *skulle*, are likely to have developed as an extension either from deontic or from alethic (viz. dynamic) modality.

Apart from *ska*, Swedish knows another modal used as hearsay marker, namely: *lär*. Etymologically it derives from *lära* 'to seize, apprehend, learn' and has become void of all its paradigmatic forms except the present indicative. Like *ska*, its evidential function must have evolved from modal meanings, either deontic or epistemic. The reportive interpretation of both *ska* and *lär*, if used alone, is easily available for contexts in which the reported event is anterior or simultaneous to the actual speech event; only quite specific contexts trigger a reportive interpretation of *ska* and *lär* if the reported situation is posterior to the speech event. In this respect these Swedish auxiliaries are similar to Fr. *devoir* (cf. Squartini 2004: 874-879), albeit not wholly identical to it (Kronning 2007: 293-296; cf. also Dooley/de Haan 2006). However, *lär* and *ska* can be combined to one complex unit whose unequivocally reportive meaning does not depend on the linguistic context, also with regard to situations posterior to the speech event.

Other modals of NECESSITY seem to be less well represented. Dutch *moeten* 'must' is a potential case in point; its hearsay reading seems to be triggered only by favourable context conditions, whereas the evidential default interpretation is the inferential one (F. de Haan, p.c.). Cf. the following example from de Haan (1999b: 76):

(23) *Het **moet** een goede film zijn*.   i. 'This **must** be a good film.'
      ii. 'This **is said to** be a good film.'

Neither Germ. *müssen* nor Engl. *must* show any remarkable signs of an extension from inferential to reportive function. Mortelmans (2000: 137f.), who claims that this happens, does not give any English example; nor is the single German corpus example she adduces convincing, although Mortelman's implicit intuition concerning discourse conditions favouring a shift from inferential to reportive meaning is probably correct. An analogous remark deserves the example constructed by Ehrich (2001: 150): *Das Testament **muß** unterschrieben sein (habe ich gehört)* 'The testament must have been signed (as I heard)'. Here *müssen* acquires a reportive interpretation only by virtue of the added comment *habe ich gehört* '(as) I have heard'. What the speaker is primarily saying in such a case is that s/he infers that the testament has been signed; as a kind of "evidential corroboration" s/he can back this inference by reference to hearsay, which is not inherent to *müssen* as such.

This indifferent evidential meaning corresponds to what we encounter in some geographically close Romance languages. Fr. *devoir* 'must' has been evolving a hearsay function via favourable contexts of report; its Italian cognate *dovere* has done so to a more restricted extent, even more restrictive are the conditions under which Sp. *deber* can acquire an air of hearsay. With both the Italian and especially with the Spanish modal these conditions are characterized by the non-factual status of the state of affairs which is being conveyed as (having been) bound to occur. Their status as reportive markers is thus much more arguable than that of Fr. *devoir* (cf. Squartini 2004).

In contrast to German and English, but in accordance with Dutch we find in Romance the auxiliary for MUST with hearsay function. Again, this function seems to have arisen on the background of favourable context conditions out of a broader meaning of undifferentiated indirect evidentiality (cf. Squartini 2004). Compare an Italian example which corresponds exactly to the Dutch one in (23) (same for Fr. *devoir*):

(24) ***Dev'essere un buon film.***   i. 'It **is said to** be a good film.'
                                        ii. 'It **appears to** be a good film.'

Outside Germanic, Romance and West Slavic modal auxiliaries hardly ever seem to have acquired a reportive function. The only case known to me is the past tense of Est. *pidama* 'must', discussed already in 3.2.3. Remarkably, like the Germanic and Romance modals mentioned above *pidama* belongs to the field of (originally deontic) necessity. Apart from the reportive past tense (*pidi*), the present tense form *pidama* if affixed by the reportive *vat*-suffix can also convey a reportive function, but it then will evoke epistemic overtones (Kehayov 2002: 136).

3.3.3.   Extensions from volitional verbs

The only indisputable case seems to be Germ. *wollen* 'to want'. The conditions of its use are identical to those of *sollen* (cf. Diewald 1999: 283f.; Remberger, to appear), as is the restriction to the present tense forms (see 3.3.2). A reportive reading is excluded if agentive control is obvious. Thus, an utterance like *Sie will liegen* will be interpreted volitively ('She wants to lie (down).'), a reportive reading is practically excluded. In comparison to *sollen* there is only one crucial additional restriction: with *wollen* the speaker of the original utterance must be identical with the referent of the subject NP of the actual utterance, whereas it is exactly this referential identity which is excluded for *sollen*. We can easily see this from a comparison of the following examples. If in the made-up examples (19-20) *wollen* replaces *sollen*, only (19') can be given a plausible interpretation, whereas (20') is utmostly odd. The reason is that the identity condition just formulated for *wollen* requires there to be a real referent of the subject NP existing at the time of utterance:[24]

(19') *Er **will** (schon einmal) Schach gespielt haben.*
        'He **is claiming to** have played chess (once before).'
(20') ?? *In früheren Jahrhunderten **wollen** die Menschen mehr auf dem Land gearbeitet haben.*
        ?? 'In earlier centuries, people **claim to** have worked more in the country.'

Furthermore, *wollen* is hardly imaginable as a reportive marker in the first person. First-person use would yield a volitional reading. In minimal pair conditions the substitution of *wollen* for *sollen* would render the sentence unintelligible or force a very complex interpretation due to which the actual speaker verbalizes what somebody else has claimed him/herself (i.e., the actual speaker) to have told about him/herself; simultaneously the actual speaker denies to have uttered the alleged assertion; see (25b). Occasionally, first-person *wollen* could also lead to the pragmatically undesirable effect of unmasking oneself; see (26b). This contrasts with first-person usage of *sollen*, which is invariably interpreted as a simple reference to other people's assertions (utterances) about the actual speaker (25a, 26a) (# means pragmatic unintelligibility or inappropriateness):

(25a) *Ich **soll** im Bett liegen geblieben sein.*
        'I **am said to** (*am obliged to, *should) have stayed in bed.'
(25b) ? *Ich **will** im Bett liegen geblieben sein.*
        'I **am said to have said that** I have stayed in bed.'
        #?'I **want to** have stayed in bed.'

(26a) *Ich **soll** mit der Mafia unter einer Decke stecken.*

---

[24] (20') could be interpreted at best in the sense that the people spoken about left some written documents about, say, their conditions and habits of labour. But how could those people have compared their way of life with conditions at the beginning of the 21st century?

'I **am said to** be cooperating with the mafia.'
(26b) ? *Ich **will** mit der Mafia unter einer Decke stecken.*
    i.   #'I **am said to** be cooperating with the mafia.'
    ii.   'I **want to** cooperate with the mafia.'

In sum, *wollen* and *sollen* distribute complementarily with respect to a property that might be called 'referential selectivity'.

Remberger (to appear), in a recent, thorough study on evidential and modal shifts of WANT-verbs, has compared Germ. *wollen* with It. *volere* and Fr. *vouloir* and their alleged hearsay meanings. She concluded that in the two Romance languages quasi-evidential WANT „is restricted to a specific class of non-animate nouns (...) or folklore" (p. 17f.). Such uses can be found not only in German, but also in English (*The tradition wants...*, *Die Tradition will (es), daß ...*). For this reason we cannot count these verbs as sufficiently conventionalized markers of reportive evidentiality.

3.4. Predicatives

The class of lexemes to be treated in this subsection is peculiar, inasmuch as they are neither modifiers of the verb phrase, nor do they modify clauses (or sentences), but constitute sentences themselves. In this view, predicatives are intermediate between modifiers of the nucleus and lexemes that are not integrated into clausal syntax, but take scope over whole sentences (as do particles, see 3.5).

Predicatives are a specific word class of heterogeneous morphological origin, whose members share three properties: (i) they are uninflected (including originally finite verb forms deprived of their paradigms), (ii) their only syntactic function is a predicative one (thence their name), i.e., they constitute the nuclei of (simple or complex) sentences, (iii) they imply an animate argument (sentient being, perceiver) whose syntactic realization is either blocked or relegated to an oblique case (mostly dative) or a PP. Instead, many predicatives require a second argument, very often of a propositional nature, which is syntactically realized by complementiser + clause; exceptionally this argument may be encoded with a PP and this results in a specific evidential function (see ex. 33 and the comment on it). In any case sentences with predicatives as nucleus do not have a nominatival subject, i.e., no NP to trigger agreement on the predicate (verb).

Predicatives are numerous in Slavic and Baltic languages, where we also observe certain tendencies toward morphological unification. As concerns lexical semantics, the predominant number of predicatives denote transient emotional or physical states (e.g., Pol. ***Przykro nam***.DAT*, że nie zdałeś egzaminu* 'We **are sorry** that you didn't pass the exam', ***Zimno mi***.DAT 'I'm **cold**'), attitudes (e.g., Lith. *Visiems*.DAT *buvo **gėda**, kad teko sumeluoti* 'Everybody **was ashamed** that they couldn't but lie') or modal notions like obligation, dynamic and deontic possibility or their negative counterparts (e.g., Russ. *možno, nel'zja, nado*, Pol. *trzeba*, Lith. *galima*).

Beside the lexical groups just mentioned there is a small subgroup of predicatives related to perception which have evolved into evidential markers, an even smaller number specializes in hearsay. These are:

(27) **Russian**   reportive: *slyšno* 'to be heard (to be told)'
                     inferential: *vidno* 'to be seen, inferred' (in the same meaning also dialectal *vidat'*; cf. Kuz'mina 1993: 129f.)

(28) **Polish**   reportive: *słychać* 'to be heard, told'
                     inferential: *widać* 'to be seen, inferred', *czuć* 'to be sensed', *znać* 'to be known; as is known'

(29) **Lithuanian**   reportive: *girdėti* 'to be heard, told'
                       inferential: *matyti* 'to be seen'

The tiny lists given here are most probably exhaustive; apparently, Latvian does not have a single such unit (though, like the three other languages, it has predicatives describing emotional and physiological states).

    By their morphological form, the Lithuanian and Polish units in (27-28) are infinitives. In Polish these forms have been petrified, because they have lost all other forms of their former paradigms. On the contrary, Lith. *girdėti* and *matyti* still function as the main, stylistically neutral verbs of seeing and hearing, respectively. It is the infinitive (which does not allow for a nominatival subject) and the way how it codes a sentential argument which set them apart as hearsay units from the „homonymous" verbs with intact inflectional paradigms and a diversified array of syntactic patterns (see below; on *matyti* cf. Usonienė 2001). Contrary, the Russian predicatives *slyšno* and *vidno* go back to adjectives (or adverbs) derived from the same common Slavic roots *vid-* 'see' and *slyš-/slych-* 'hear' (as do the Polish predicatives in 28). In modern standard Russian *slyšno* still occasionally shows up in adverbial usage (see 30a), whereas *vidno* can no longer be used as an adverbial at all; thus both units have split off semantically from the adjectives *vidnyj* '1. visible, 2. prominent, 3. stately (about good-looking men)' and *slyšnyj* 'audible', which are tightly associated to direct evidentiality.

    Depending on the type and coding of the argument, *slyšno* (as well as *vidno*) mark different evidential functions. With a nominal object NP *slyšno* refers to direct perception (30b); if the argument is propositional and adjoined by a complementiser, *slyšno* can still refer to direct perception (or inferences based thereof, see 30c), but it then can also acquire reportive function (30d):

(30a)   *Na berezax ele **slyšno** oživajut razmjakšie vetki.*   → adverb
         'On the birch trees the withered twigs come back to life hardly **audibly**.'

(30b)   *Marfen'ku vsegda **slyšno** i **vidno** v dome. Ona to smeetsja, to govorit gromko.*
         'Marfen'ka **can** always **be heard** and **seen** in the house. Either she is laughing or she is shouting around.'
         (I. Gončarov, „Obryv")   → predicative, direct perception

(30c)   *Gluxota moja byla nepolnoj: skvoz' nee bylo **slyšno, čto** golos Belokurogo vysok i nazojliv.*
         'My hearing loss was not profound; it was still **possible to hear that** Belokuryj had a high and piercing voice.'
         (I. Grekova, „Bez ulybok", 1975)   → predicative, direct perception (inference ?)

(30d)   *Poslednee vremja tol'ko i **slyšno, čto** oligarxam vredno zanimat'sja politikoj.*
         'Recently it **can** only **be heard that** it is detrimental for the oligarchs to engage in politics.'
         („Sovetskaja Rossija", Feb., 15th, 2003) → predicative, hearsay

Predicatives functioning as markers of indirect (reportive or inferential) evidentiality never allow for a realization of the implied animate argument. This salient syntactic feature distinguishes them from predicatives that belong to other semantic domains (in particular from those denoting emotional or physical states, but see also *oligarxam*.DAT.PL *vredno* 'for the oligarchs it is detrimental' in the second part of ex. 30d). Cf., for instance Lith. *girdėti* (31) and Pol. *słychać* (32):

(31)   *Gaila, kad Žemaitės mes negalim skaityti žemaitiškai... – Gal galėsim? **Girdėti** (*visiems), kad susipratę leidėjai ketina parūpinti autentiškai parengtus jos "Raštus".*

'Too bad that we cannot read Žemaitė in Samogitian … – Maybe it will be possible one day? (*To all) **there are rumors** that publishers who are skilled in it intend to get the authentically revised "Writings".'
("Šiaurės Atėnai", 1995/5-7)

(32) **Słychać** (*_nam_)_, że w Iraku znowu dokonano zamachu._
'**It is said** (*to us) that yet another assault has been plotted in Iraq.'

Like Russ. _słyšno_, these predicatives convey a reportive meaning only if they code a propositional argument; see (31-32). Token frequency of such items is however very low (cf. Wiemer 2006b: 29 for Polish) and outranked by a pattern in which these predicatives take a nominal argument and, consequently, mark direct auditory evidentiality; compare Pol. _Na ulicy było_.PAST.N **słychać** _krzyki_.ACC.PL 'On the street one could **hear** shouts', Lith. _Už kalno_ **girdėti** _griausmas_.NOM 'Behind the hill thundering can **be heard**'.

Somewhat more frequently _słychać_ is used with the valence frame _o_+LOC 'about'. In this case it refers not to the content of reported speech, but only to its topic; e.g.

(33) _Często było   słychać o   nim._
     often   COP.PAST.3.SG.N REP   about he.LOC
lit. ≈ 'Often (it) was heard about him.', i.e. 'Often he was spoken about.'

The same applies to Russ. _słyšno_, which is however more colloquial (noticeably with the preposition _pro_+ACC) and occurs more rarely than Pol. _słychać o_,[25] whereas Lith. _girdėti_ practically does not realize this syntactic pattern at all.[26]

The units listed in (27-29) have "homonymous" particles. The distinctive criterion between particles and predicatives is that particles lack argument structure and, consequently, cannot be combined with a complementiser.

3.5. **Particles and sentential adverbs**

I will now proceed with those considerably more numerous expressions that belong to classes which lack an argument structure of their own: particles and sentential adverbs. Since, on the basis of the existing literature, I have found myself unable to figure out any crucial difference (either functional or structural) between sentential adverbs and (modal) particles, I will treat them here in one rubric, leaving out for future investigations whether there is any tenable functional difference that can be generalized over languages. I will refer to the relevant units as 'particles', although in various descriptions some of them might have been named 'adverbs' or else (for a survey cf., for instance, Ramat 1996); terminology varies considerably also due to national linguistic traditions.

The medley bag of words called 'particles' which we are concerned with here are sentential modifiers (as opposed to focus particles and particles that function as signals of turn-taking or attention-regulating devices in dialogue). Whether in European languages hearsay particles can take scope not only on the propositional, but also on the illocutionary

---

[25] Compare an instance from Šoloxov ("Tixij Don"): – _Pro Aksin'ju ne_ **słyšno?** _– sprosil Grigorij s delannym ravnodušiem v golose._ '"Is there nothing **to be heard** about Aksin'ja?", Grigorij asked in a voice that pretended indifference.'

[26] Even a search by Google did not yield a single conclusive example. This pattern was encountered only in the fixed expression _Kas (naujo) girdėti (apie)?_ 'What's the news; How are you?', which is structurally identical to its Polish equivalent _Co słychać (o)?_ In all other hits which included the preposition _apie_ 'about' the form _girdėti_ was used as an infinitive proper, quite often in the scope of a modal (e.g., _teko girdėti apie_... '(one) had to hear [= learn] about'). This result from a Google search is corroborated from an informant's reply: _girdėti_ as a predicative with _apie_ referring to the topic of speech sounds natural only in questions (e.g., _Ar_ **girdėti** _kas nauja apie studijų reformą?_ ≈ 'Is there anything new **to be heard** <u>about</u> the university reforms?').

level (for the difference cf. Faller 2006), cannot be investigated here. As far as I can see, no water-tight and cross-linguistically applicable criteria have been formulated of what should count as particle; often particles are "defined" in negative terms: they are uninflected; neither are they part of the clause's constituent structure, nor do they have to occupy fixed positions; they are highly heterogeneous in terms both of syllable structure, cliticizability and morphological provenance, etc. Ultimately the only valid criteria appear to be located on the level of pragmatic functions (for overviews and/or attempts toward defining particles as well as parentheticals cf., e.g., Burkhardt 1999; Grochowski 1986; 1997: 22-24; 2007; Mosegaard Hansen 1998: ch. 3).

More or less identical criteria are applied to determine parentheticals (see 3.6). The reason why particles and parentheticals are separated here, is that linguists treat as parentheticals not only words (word forms), but units consisting of more than one word (form), often of whole syntagms (e.g., small clauses). Thus, before "uniting" particles and parentheticals on functional grounds, it seems advisable to account for particles as one-word units first. Admittedly, this dividing line appears somewhat arbitrary and in some cases it will even be violated, at least insofar as we would need a universal definition of 'word' (in morphological and prosodic terms). We should however notice that one-word units, being short, often underlie rules of cliticization specific for the respective language. Beside that, particles can take parenthetical clauses into their scope (see ex. 47); they must therefore be distinguished from each other. Thus, in order to supply a transparent description I prefer, for the time being, to provide arguments and empirical data for parameters according to which reportive particles can be subdivided. Whether and which of these parameters can be applied to subclassify parentheticals, too – and how we may unite both classes – has to remain the task of a (hopefully) nearer future.

Since structurally particles are not integrated into the syntax of the clause which they modify, their scopal behaviour can vary, and since among all reportive markers they display probably the widest array of semantic-pragmatic diversification, they are best suited for the task of catalogizing structural and functional parameters along which reportive markers can be classified. Particles, as it were, can be considered as the best accessible playground of multifarious factors that influence the usage of reportive markers and, therefore, supply the empirical basis on which I will try to work out an important part of the complex network of these factors.

For the present purposes it should suffice to conceive of 'particles' as units which fulfil the usual criteria mentioned above. I will first give some representative examples of reportive particles from different European languages and then elaborate on parameters of their subdivision.

English *apparently* (cf. Wierzbicka 2006: 278ff., from where this example is quoted):
(34)   ***Apparently*** *there's eight times more sheep than there are people in New Zealand.*

German *angeblich*
(35) *Die Bundesregierung arbeitet **angeblich** an einem zweiten Konjunkturprogramm im Umfang von mindestens 30 Milliarden Euro.*
'The Federal Government **is reported to** be working [is **allegedly** working] on a second economic stimulus package totalling at least 30 billion euros.'
(ZEIT online, 13.12.2008)

Romanian (by courtesy of A. Merlan, Konstanz)
(36) ***Cică*** *se   căsător-eşte   cu   o   franţuzvaică.*
     REP   RM   marry.3.SG.PRS   with   IND.ART   French.F
     'He **is said to** be marrying a French woman.'

Spanish *dizque*

(37) *Esto* **dizque** *va a ser pantano*.
　　 this REP(=SayC) go.PRES.3.SG to be swamp
　　 'This is going to be swamp, **they say**.'

For this and other examples from different varieties of Romance cf. the systematic account of *dizque*-units in Cruschina/Remberger (to appear).

Polish *jakoby* (cited from Wiemer 2006b: 41)

(38) *List stryja wiele wyjaśniał, ale jeszcze więcej gmatwał; czego tam nie było! Potopy, które co tysiąc pięćdziesiąt lat nawiedzają Polskę; przodkowie rodem z Egiptu; historia Sieciecha, palatyna Władysława Hermana, który* **jakoby** *pragnął przekazać przyjacielowi brata koronę; przyszłe córki Alicji w roli zbawczyń ojczyzny...*
'The uncle's letter explained much, but then it caused even more confusion: there was all sorts of stuff in there! Floods that hit Poland every one thousand and fifty years; ancestors of Egyptian origin; the story of Sieciech, Wladyslaw Herman's Palatine, who **allegedly** wished to pass on the crown to his brother's friend; the future daughters of Alicja as saviours of the nation…'
(T. Mirkowicz: „Pielgrzymka do Ziemi Świętej Egiptu", 1999)

Croatian / Serbian *navodno*

(39) *Sve bi nas valjalo ponovo vratiti iza Karpata, odakle smo,* **navodno**, *došli*.
　　 'We all should be sent back beyond the Carpathian Mountains, from where we **are said** to come.'
　　 (Internet)

Lithuanian *esą*

(40) *Daugeliui teko ilgai laukti savo mokslinių laipsnių ir pedagoginių mokslo vardų patvirtinimo, nes,* **esą**, *jie „neturį" sovietinio diplomo*.
　　 'Many had to wait long for an approval of their academic titles or teaching diplomas, because, **so it was said**, they did "not have" a Soviet degree.'
　　 (Z. Zinkevičius "Prie lituanistikos židinio"; 1999, p. 76)

Latvian *it kā* 'as if'

(41) *Šobrid* **it kā** *viss ir, lai izglītība atbilstu darba tirgus prasībām*.
　　 '**Allegedly**, there is everything available now to render education fit for the job market.'
　　 (www.politika.lv, December 2006; by courtesy of I. Kļēvere-Wälchli)

Further cases: It. *dice* (Giacalone Ramat/Topadze 2007: 27; Pietrandrea 2007: 55-57), Gr. *léi* (Stathi, this volume), Russ. *mol, deskat', -de* (Arutjunova 2000; Plungjan 2008), Russ. *kažetsja, budto by, vrode* (Letučij 2008; Wiemer 2008), Pol. *podobno, ponoć, rzekomo* (Rytel 1982: 47f.; Wiemer 2006b), Cz. *prý, údajně* (Hoffmannová 2008; Karlík et al. 1995: 626f., 664; Rytel 1982: 46, 48f.), Slovak *vraj*, Macedonian *vele* (Mushin 1997: 294-296), Lith. *tarytum, girdì, neva* (cf. Wiemer 2007b), Georg. *turme, titkos, vitom* (Topadze 2008: 51f., 56-60). We should add Arm. *ɛʏɛɾ*, which is an etymological equivalent of Lith. *esą* (see ex. 40), as it is an isolated participle form of the copula. In semantic terms however *ɛʏɛɾ* differs from Lith. *esą* insofar as it is not restricted to hearsay (cf. Donabédian 1996: 95-97).

Parameters pertinent for a subclassification of reportive particles can be divided and ordered into (at least) five global groups:
(i) diachronic background, with a further subdivision into:

     (a) etymology and semantic evolution,
     (b) changes of morphosyntactic properties (along the "cline" in Fig. 1),
(ii) criteria of formal (morphosyntactic) properties and behaviour,
(iii) criteria of semantic-pragmatic properties,
(iv) distributional criteria, i.e., how do such particles combine
     (a) with other evidential markers (regardless of their "morphosyntactic format") and
     (b) with other units that function as sentential modifiers (inter alia, epistemic ones)?

In this paper I will focus only on criteria of groups (ii-iii). A fifth group of criteria concerns (v) the variability of grammatical status (syntactic distribution), i.e., the question whether a unit with a particular phonological shape can occur in different types of constituency (according to the classification substantiated in section 2) without noticeable changes of lexical meaning (sc. evidential function). This question is not to be neglected as for the proportion of units concerned, nor ought we to dismiss it in our theorising. It is a kind of synchronic correlate to the diachronically oriented issue in (i-b). The manner we handle issue (v) will be crucial for any solid lexicographic account of evidential markers in a language as well as for approaches towards the lexicon—syntax-interface. I will pick up this issue in 3.9.

Probably the most important criterion of group (ii) concerns the scopal behaviour of a unit, first of all whether its scope is only sentential or variable, in the latter case it can narrow its scope down to even attributes of a noun (see 42b, 43, 44). Although one needs to have checked a sufficient amount of relevant data in order to be able to state firmly that the predominant number of reportive particles displays variable scope, I have by now got the general impression that reportive particles are capable of variable scope by default. Let us assume this for the time being. Another default rule to hold for most hearsay particles appears to be that they need not specify the author of the original utterance. Probably there is an inherent semantic or pragmatic link between these two co-occuring properties.

All units illustrated so far can have wide (propositional) scope; most, but not all of them can occur with narrow scope. If they do they normally take the narrowest possible scope, i.e., scope only over a part of a NP (an attribute). At least this is the most frequent case with text tokens known to me. In this respect, Germ. *angeblich* 'allegedly' is representative; cf.:

- wide scope

(42a) *Gestern* [*kam er **angeblich** erst um fünf mit seiner bereits sechsjährigen Tochter*].[27]
    '**Allegedly**, yesterday [he arrived only at five with his already six-year-old daughter].'

- narrow scope

(42b) *Er kam gestern erst um fünf mit seiner **angeblich*** [*bereits sechsjährigen*] *Tochter*.
    'Yesterday he arrived only at five with his **allegedly** [already six-year-old] daughter.'

Here and in the following examples narrow scope is indicated by square brackets. See some further typical illustrations:

Polish *jakoby*
(43) *Wiele osób (...) w dyskusjach nad problemem przywrócenia do życia owego czołowego pomnika pamięci narodowej podnosi* [*zbyt długi*] ***jakoby*** *okres odbudowy oraz ogromne jej koszty*.

---

[27] Actually, scope may be ambiguous and embrace only the adjacent constituents to the right (the adverbial *erst um fünf* 'only at five' together or without the comitative PP). Prosody (sentence stress) plays an important role in disambiguating scope. We must however leave this general problem, as it does not contribute essentially to the present points.

'In debates on the issue of the revitalization of this outstanding site of the national memory many have addressed the **allegedly** [too long] period of reconstruction and the enormous costs.'
 („Młody Technik", 1971-3)

Russian *jàkoby* (cited from Rakhilina 1996: 299)
(44) *Vy, razumeetsja, slyšali o moem jakoby* [*romane*] *s vami?*
        'For sure, you have heard about my **alleged** [affair] with you.'

Russian *budto by* 'as if, apparently'
(45) *Muxin do six por budto by ničego ne znaet ob ėtom.*
        '**Apparently**, until now Muxin knows nothing about it.'

Utterances like (45) allow for inferential readings, too.[28] Pol. *jakoby* and Russ. *budto by* are also encountered as complementisers (see 3.7, 3.9).
   Also attested, though less often, are cases in which a particle takes an adverbial PP into its scope (46) or a parenthetical expression (small clause, see 47):

Lithuanian *esą* (cited from Wiemer 2007b: 177)
(46) *Pirmą kartą nepažįstamas vyriškis Juozui paskambino šių metų balandžio pabaigoje. Užuominomis jam pasakė, kad turėsiąs mokėti pinigus esą* [*už blogą tarpininkavimą*]. *Buvo leista suprasti, kad reikia duoklės.*
'The first time an unknown man phoned up Juozas at the end of April this year. Using allusions he told him that he would have to pay money **supposedly** [for bad mediation]. They let him know that a tribute was required.'
("Lietuvos Aidas" 1995-8)

Croatian / Serbian *navodno* (by courtesy of A. Drobnjaković)
(47) *Povodom teksta u estradnom časopisu "Mega hit", u kojem je osvanula informacija da se razvodi, navodno* [*zbog jedne plave žene*], *upitali smo Ljubu Aličića da li je to istina. "Naravno da je u pitanju trač", kaže Ljuba.*
'On the occasion of the article in the magazine 'Mega hit', where the information appeared that he is going to divorce his wife, **allegedly** [because of a blond woman], we asked Ljuba Aličić if that is true. 'Of course it's a gossip,' says Ljuba.'

There is at least one peculiar case of a reportive particle which can also function as a sort of word-formation device. This is Pol. *niby* 'as if': like *jakoby* it can be used both as particle and conjunction (see 3.7, 3.9); it then not only refers to a foreign speech act, but also implies that the actual speaker does not subscribe to the implicit consequences which the original speaker might have had in mind (cf. Wiemer 2007b: 45-48). If used like a prefix, *niby-* adds an ironical touch:

(48) *Widocznie te twoje niby-badania w tej szopie to wcale nie taki cymes.*
        'Apparently, your **so-called** [= pseudo-]research in that shed there is not such a big deal.'
        (R. Antoszewski: "Kariera na trzy karpie morskie", 2000)

This is reminiscent of what Brendel et al. (2007: 6) have called 'modalized citation' ("modalisierendes Zitat") and similar to Travis' (2006: 1278) 'labelling'. Thus, *niby-* raises

---

[28] For an analysis of *budto by* and other etymologically closely related units cf. Letučij (2008: 222-232) and Wiemer (2008: 347-353).

the general problem of whether it is justified to include adnominal modifiers of the SO CALLED-type into an inventory of reportive markers. There are reasons in favour of such a decision. First, labelling is not restricted to word-formation devices like Pol. *niby-* or attributive adjectives like Germ. *so genannter*, It. *cosidetto*, Russ. *tak nazyvaemyj*, but is commonplace among particles with variable scope, i.e., to the assumed absolute majority of them. Second, adpositions – which are adnominal modifiers par excellence – can also be considered reportive units (see 3.8), although they do not participate in word-formation. Third, notice that SO CALLED-modifiers have a specific quotative function *sensu stricto*. Meibauer (2003) argued that quotative meaning components[29] can be integral parts of words. If we accept these arguments we should however be aware that we are likely to be running into problems with delimitation both in semantic terms (how deeply can the element referring to a previous utterance be anchored in the semantic explication of a linguistic item?) and in morphosyntactic terms. As for the latter, remember Anderson (1986: 275), who excluded compounds and derivational forms from the inventory of evidential markers (see his fourth defining criterion).

A further problem arises: if *niby* becomes a bound morpheme and thereby loses its syntactic freedom even more than clitics, how should its categorial status be handled and, consequently, how many lexical entries of *niby* are there? (On this general issue see section 3.9.) Apart from this, notice that *niby* used as a free morpheme cannot be regarded as a clitic (just like most markers in Polish, Russian and Lithuanian do not behave like clitics); we cannot therefore assume a cline: free > cliticized > agglutinated (> fused) morpheme, as might be natural in other languages with a more salient propensity towards clitics (e.g., Romance or South Slavic); instead, we would have to assume that the "clitic stage" is circumvented.

Although most particles with hearsay meanings supposedly have variable scope, there are some of them which demonstrate only wide, sentential scope. On preliminary terms, two rather systematic types figure saliently. Remarkably, both can be characterized as quotatives in the narrow sense (cf. Aikhenvald 2004: 64, 177f.). First, reportive markers which only recently have split off from complementisers; in such cases it may even be arguable whether they have at all to be treated as units distinct from complementisers. A nice case to illustrate this is supplied by Cz. *že*. See the following examples from spontaneous discourse: in (49) *že* occurs after a subject pronoun (*voni* 'they'), we can assume an ellipsis of the predicate. In (50) however *že* occurs so frequently without a speech act verb, both after intonational breaks and without any breaks, that we get the impression that it is becoming a syntactically independent unit marking quoted speech:

(49) *měla sem pěkný seno (...) a tak sem šla Kostelce / a <u>voni</u> že už maji všechno plný a že už to seno nebudou brát jo / (...)*
'I had some nice hay (…) and so I went to Kostelec / but <u>they</u> *že* already had enough of it, and *že* they won't take any more hay, will they'

 (50) *vona mě prosila (...) že zkrátka bych mohla že čeká rodinu že bych mohla u nich bejt / že se bude mít dobře // ale že byli sami chudáci / že sem se nic moc neměla / že sem se tam hrozně zhubla (...)*
'she asked me (…) *že* I could, at short notice *že* she is expecting a child *že* I could stay with them / *že* s/he will be well off // but *že* they were all poor / *že* I wasn't well off / *že* I terribly lost weight there'.

---

[29] Such a component need not be marked morphologically, it can also manifest itself as a prosodic (or graphic) feature in combination with the conversion of whole phrases. But with SO CALLED-units they are part of their lexical meaning.

Hoffmannová (2002: 374) treats the repeated use of *že* as hypertrophy which leads to a mixture of quoted, indirect and free-indirect speech. Now, on account of the status of *že* as oscillating between a complementiser "floating" without its predicate and a particle, it is unlikely that *že* gains scope over constituents on levels lower than the clause.

Admittedly, some reportive particles which arguably began (and can still be used) as complementisers with an 'as if'-meaning, are now used with variable scope; compare Russ. *jakoby* and *kak budto*, *budto by*, Pol. *jakoby* (see ex. 43), probably also Latv. *it kā*. These units however are likely to have gained the status of particles a more considerable time ago than Cz. *že*. Most plausibly all of them started as clause connecting devices in constructions of comparison with an 'as if'-meaning.[30] To this group we can add Lith. *esą*, although its history is different: it derives from a participle used in complementation (see 3.2.1). It is reasonable to assume that all these items acquired variable scope only after they had separated from "their" subordinating verbs, although this assumption has not yet conclusively been tested on diachronic data.

Second, another group of quotative units lacking variable scope originates from syntactic positions other than clause connectives. Colloquial quotatives like Engl. *like*, Russ. *tipa* (= *tip*.GEN 'type, sort') should be named here in the first place. Again, in this group expressions used for comparison are particularly commonplace, but demonstratives and quantifiers (in a broad sense) are also mentioned quite often, as, for instance, Germ. *und sie so*, lit. 'and she so' (+ quote). Such syntactically indeterminate units as a rule precede the reproduction not only of a literal quote, but also of paralinguistic features, mimics and gestures of the original speech event; cf. the title of Golato's (2000) article dealing with reference to "embodied actions". This distinguishes them from units like Cz. *že*, which started their lives as connectives. Since quotatives of the LIKE-type are so tightly associated to the complex reproduction of speech, it is hard to imagine that they will ever occur with narrow scope. Moreover, since they refer to the entire "behavioural accompaniment" of utterances (the manner of pronunciation, characteristics of voice, etc.), the notion of sentential (or propositonal) scope sounds somewhat misplaced. Utterances need not take on the shape of sentences, and in spoken language (i.e., where quotatives of this type occur) they notoriously decline from models of standard syntax. Thus, if we want to say that these quotatives have scope over propositions, they additionally, and on equal rights, mark illocutionary and emotional aspects of speech.

Some reportive particles can behave like clitics, e.g. Cz. *že*, *prý*, Russ. *mol*, probably also It. *dice*. Notice that the degree of coalescence (in terms of Lehmann 1995) does not automatically correlate with a reduction of scope. For instance, the Georgian hearsay morpheme *-o* is a clitic, but it nonetheless operates on phrase level. According to Boeder (2000: 279), "in colloquial speech [it] can occur on each constituent boundary". The following example is quoted from Giacalone Ramat/Topadze (2007: 18):

(51) *gaero-m,   v   -i   -ziareb   -t   tkven-s   mc'uxareba-s-**o***
     UN.ERG   1.PL   VERS   share:PRS   1.PL   your.DAT   grief.DAT.**REP**
     'The United Nations said, we share your grief!'

Furthermore, certain reportive clitics show tendencies towards agglutination. As cases in point we can count Pol. *niby* mentioned above and Georg. *metki* (see ex. 55). A further item which might belong into this subclass is Russ. *-de* (cf. Plungjan 2008). Arutjunova (2000: 438) cites examples like the following from the Dal'-dictionary (composed during the 19th c.), in which -

---

[30] A preliminary investigation of Russ. *jakoby* and its Polish cognate *jakoby* corroborates this assumption (for details cf. Wiemer 2005: 115, 120-123; 2008: 362-369). As for the development of evidential marking from comparison constructions, in general, and the Russian series of expressions centering around *budto* 'as if', in particular, cf. Letučij (2008).

de occurs repeatedly on successive syntagms (and which can barely be translated into English other than with the aid of a colloquial quotative marker):

(52)   *On govorit, ja-**de** ne pojdu-**de**, chot'-**de**, čto choš' delaj.*
      'He **goes** "I won't go, you can take it or leave it".'

Such occurrences have however clearly diminished, (52) sounds rather hypertrophical in contemporary Russian.

Let us now turn to criteria of group (iii). Among them three are salient:
(a) the relation to speaker's assessment of the trustworthiness or veracity of the utterance(s) referred to (absence vs. presence of "epistemic overtones");
(b) known—anonymous speaker—folklore;
(c) the question which level, or aspect, of the reported utterance(s) is highlighted by the respective marker: propositional content—topic of speech—manner of speaking; propositional content—literal quote; can the speaker refer to an utterance of him/herself?; must the speaker have heard the reported utterance(s) him/herself?, etc.

As concerns the relation of hearsay to epistemic assessment, I refrain here from any discussion. Factor (b) has casually been commented on above and in sections 3.2.4, 3.3.2 and 3.3.3. In the following I will concentrate on criteria of group (c).

We can distinguish a bunch of aspects within the reference to speech acts which prove to be relevant at least for some units. The criteria named immediately below do not pretend to be exhaustive, but they are outstanding provided we regard as the most trivial (widespread) kind of reportive markers those ones that (i) focus on the propositional content of speech (ii) uttered previously (iii) by another person (other people) (iv) in the absence of the actual speaker. On the background of these four "default conditions" the following more specific aspects can be singled out:

1. Literal reproduction of speech. This is what quotatives proper do. We have already mentioned a few. Notice that, as a rule, quotatives usually imply that the actual speaker him/herself has witnessed the utterance s/he is referring to.

We may further ask whether particles that normally fulfil the four default conditions just formlated can also be used with direct quotes. Theoretically nothings precludes such a usage; however in practice this seems to happen rarely, labelling appears to be an exclusion (see ex. 48). This shows that reference to utterances with "embodied action" (Golato's 2000 term) differs in principle from reference to denominations (etiquettes). Again, these impressions open up further issues remaining for empirical investigation. In this connection it seems sensible to test also whether the following unidirectional implication holds: if a reportive particle (or, generally, a reportive marker) is normally used under the four default conditions, it can also be used with quotes, the reverse does not apply (i.e., quotative markers cannot replace less specific reportive markers if these refer only to the propositional content).

2. There are markers that allow referring not only to previous speech acts, but also to their semiotic substitutes. In fact, quotatives, especially from the second subgroup mentioned above, often do refer to semiotic substitutes or thoughts (assumed by the speaker), too. But this is not an exclusive characteristic; other, non-quotative subgroups of hearsay markers are able to highlight semiotic substitutes of speech as well. I have found the following two. First, the Russian "xenomarkers" *mol* and *deskat'*, which do not primarily refer to the propositional content of the reported message, but to the manner in which it was formulated (Plungjan 2008: 306f.). In the following example (from Arutjunova 2000: 444) *mol* does not refer to any speech act, but to another person's way of behaviour (*delaja vid, čto* 'pretending / showing that') interpreted by the actual speaker:

(53) *Na lice rabočego prenebrežženie, stal vozle tolpy bokom, <u>delaja vid, čto</u> on priostanovilsja na minutu, dlja zabavy: **mol**, zaranee znaju, čto vse govorjat čepuxu.*
'The worker's face shows contempt. He positions himself at the edge of the crowd, <u>giving the impression as if</u> he had stopped only for a minute, just for fun: {mol} I know from the start that everybody talks rubbish.'
(Bunin)

Second, reference to semiotic substitutes of speech occasionally occurs with general hearsay markers that focus on the content of speech. See an example of Pol. *jakoby* 'as if' (quoted from Wiemer 2006b: 40):

(54) *Innym znów razem, gdy upuszczony ołówek niefortunnie potoczył się za daleko, podniosła go niezwłocznie, nie przerywając wykładu i, jakby nigdy nic, włożyła do szuflady. Na rozpaczliwe gesty poszkodowanego (że nie ma **jakoby** czym pisać, żeby mu oddać ołówek) nie zwracała najmniejszej uwagi, kontynuując lekcję.*
'Some other time, when a dropped pencil rolled too far away, she picked it up immediately, without interrupting her lecture, and, as if nothing had happened at all, put it in her desk drawer. She paid no attention whatsoever to the injured party's desperate gestures (meaning that he **allegedly** had nothing to write with, so that he would need the pencil back) and she carried on with the lesson.'
(A. Libera „Madame", 1998)

In both cases given an air of irony arises. It remains an open question whether this is an accidental effect or the consequence of a more stable implication of the "semiotic transfer" of the respective hearsay markers.

3. As shown for the Polish predicative *słychać* (see 3.4), reportive markers happen to focus only on the topic of speech, provided their argument is coded with a PP. I am unaware of similar cases from among particles, or any other group of markers (in terms of their "format").

4. Usually reportive markers are used by speakers who have not participated in the speech events from which they report, they rather exclude the actual speaker's personal witness. An exception to this rule are the Russian "xenomarkers" already mentioned. Not only must the speaker have witnessed him/herself the reported utterance, s/he can even refer to his/her own speech uttered previously. In addition, *mol*, *deskat'* and *-de* require that the original speaker be not anonymous (or at least easily retrievable from the linguistic context), that reference to a reported speech act have been prepared in the preceding context and the conditions under which it occurred be described; cf. Arutjunova (2000: 442-447), Letučij (2008: 223f.), Plungjan (2008: 306f.).

5. Reportive markers happen to underlie restrictions with respect to grammatical person and/or to participants of the reported or the reporting speech act. A certain kind of 'referential selectivity' was already demonstrated for the reportive use of the German auxiliaries *sollen* and *wollen* (see 3.3.3). There are however other dimensions of referential peculiarities, one of them is characteristic of the Georgian particle *-metki* (< *me vtkvi* 'I said'; cf. Topadze 2008: 44f.). It marks "exclusively quotations in the first person singular when the speaker reports an utterance s/he had already made or reflected on in the past" (Giacalone Ramat/Topadze 2007: 16). The following example given there can be characterized as an indirect directive speech act:

(55) *ramdendžer   g -i -txar-**i**   k'ar-i   ar   dak'et'-o – **metki**!*
how many times 2.SG-VERS.tell.AOR.**1SG**   door.NOM NEG close.OPT.2.SG—**QUOT**
'How many times have I told you not to close the door (**I said**)!'

Georgian has another marker, *-tko* (< *-tkva* 's/he said'), which "is used when the addressee is a mediator between the speaker and a third person. The utterance is addressed to the third person, the information source is the speaker" (Giacalone Ramat/Topadze 2007: 16):

(56) *utxar-i   male   mod-i-**tko**.*
tell.2.SG-IMP   soon   come-IMP.2.SG-**QUOT**
'Tell him/her, to come soon.' (literally: 'tell him/her, come soon, **s/he said**')

In this case, the reporting utterance has not been realized yet and is, as it were, made in advance; by the same turn, the addressee of the speech event uttered in (56) is "circumvented". The suffix *-tko* functions as a quotative proper, i.e., with literal quotes, whose reference to discourse participants does not agree with the indication of grammatical person: *-tko* is a quotative marker reserved for the "prospective", intended addressee, who is absent in the speech event of (56). This usage resembles Free Indirect Speech, insofar as there is a deictic switch of person reference which in the second part of (56) re-aligns the deictic system from the point of view of the prospective addressee.

In a sense, this reminds us of some (not all!) uses of the Russian xenomarkers, namely those which seem to be typical of an older, 18th-19th century pattern; cf. an example from Arutjunova (2000: 441):

(57) *Poslušaj, podi skaži Petru, čtoby on skoree sjuda prišel: barynja, **deskat'**, sprašivaet.*
'Listen, go and tell Peter that he should come here as quickly as possible: the landlady {**deskat'**} asks (for him).'
(Krylov)

Here the speaker is the landlady (the barin's wife, *barynja*) herself, but in the last part of the utterance she refers to herself by using the third person (*sprašivaet* 'asks') and marking by *deskat'* the phrase she wants the servant to convey.[31]

Another item showing restrictions as for grammatical person appears to be It. *dice*. Having been isolated from its paradigm, this originally verbal form can be used in clauses with predicates (finite verbs) in the 3rd person, both singular and plural (see ex. 58a-b), but neither with 1st or 2nd person predicates (see 58c-d). This restriction holds at least if *dice* is used after or within direct speech (C. Guardamagna, p.c.), so that it might turn out to be itself restricted to the quotative use of *dice*:

(58a) *Mi ha chiesto / detto – **dice** – [continuation of quote].*
     'S/he has asked / said to me – **dice** …'
(58b) *Mi hanno chiesto / detto – **dice** – [continuation of quote].*
     'They have asked / said to me – **dice** …'
(58c) *Ho / Abbiamo chiesto – ***dice** – [continuation of quote].*
     'I / we have asked – ***dice** …'
(58d) *Hai / Avete chiesto – ***dice** – [continuation of quote].*
     'You (SG/PL) have asked – ***dice** …'

---

[31] During the last centuries Russian xenomarkers must have experienced a functional change. In former times (obviously until the 19th c.) they were quite common as markers of direct quotes, whereas in recent times they have been moving out of this domain and been specializing as markers of speech that, on the contrary, has to be r e f o r m u l a t e d by the actual speaker (cf. Plungjan 2008).

3.6. Parentheticals

From a functional point of view, particles and parentheticals[32] are hardly distinguishable: they both (a) take scope over whole utterances, (b) their position is not fixed (neither is their scope), and (c) their relation to the sentences (or constituents) they modify cannot be captured in terms of dependency.[33]

Here are some examples of markers where the decision as for whether the given unit counts as parenthetical or particle proves to be really arbitrary:

(59)   **French** *paraît-il* [34]
   *(...) le charmant roi mage (...) avec lequel on lui avait trouvé autrefois – **paraît-il** – une grande ressemblance.*
   'The charming king of magicians (…), whom people once had found him to resemble closely – **so it seemed** [**as they said**].'
   (M. Proust: *À la recherche du temps perdu*; cited from "Le grand Robert…" 1985/VII: 74)

(60)   **Alemannic** (Swiss etc.) **German** *schints* (*schynts* etc.)
   *Vermutlich, weil das Senioren-Konzert in Dübendorf zu Ende war, Herr Woody. Dort waren ja **schins** nicht alle gleich zufrieden mit dem Gebotenen.*
   'Presumably because the concert for the elderly people in Dübendorf was finished, Mister Woody. **Apparently** not everybody was equally content with what they had presented.'
   (Campionatischer, 2006.08.07)

(61)   **Russian** *kažetsja* (cf. Bulygina/Šmelev 1993: 80)
   *Ivan, **kažetsja**, uexal v Kiev.*
   'Ivan **apparently** left for Kiev.'

It is no accident that all examples given contain units deriving from verbs with the meaning 'seem, appear'. Other examples are Sp. *al parecer* (Cornillie 2007: 34-36), It. *a quanto pare* (Squartini 2008: 932f.), Pol. *zdaje się* (Wiemer 2006b: 53-59; 2008: 356-358), Engl. *it seems* (Chafe 1986: 268). Characteristically, all these units seem to have acquired hearsay function via inferential meanings, and most of them must now be qualified as undifferentiated markers of indirect evidentiality[35] (see section 4). Interestingly, the 3rd person singular form of the present indicative of SEEM-verbs[36] highlights a reportive function only when it loses its dependency relations with a host sentence, i.e., when it becomes petrified as a particle viz. parenthetical. As long as it has not lost syntactic links with their immediate surrounding – in

---

[32] The German equivalent is 'Schaltwörter, Schaltsyntagmen' (literally 'switching words/syntagms'), the Russian one 'vvodnye slova' (lit. 'introductory words').

[33] Cf. Kaltenböck (2007: 31) on parenthetical clauses: „parentheticals cannot be defined by themselves. (…) They derive their existence, as it were, from their interaction with a host clause. This interaction, however, takes place purely on the linear plane, not on a relational (i.e., dependency) level." (Similarly in Dehé/Kovalova 2007: 9 and Grochowski 2007: 71.) The same two characteristics – a functional one: their interpretation depends on a host to whose truth-conditional semantics they bear some relevance, and a negative syntactic one: no dependency relations – are typical of modal, or epistemic, particles (under which evidential particles have usually been subsumed).

[34] The same intermediate status can be ascribed to *dit-on*.

[35] This is probably true also of some units which are usually classified as sentential adverbs, but of identical etymology, like Engl. *apparently*, Dutch *schijnbaar* and their German counterpart *anscheinend*, which lately seems to be increasingly used for the purposes of hearsay.

[36] To this extent, Sp. *al parecer* (= infinitive) is an exception.

particular, if it is still used with a complementiser (basically as the head of a bi-clausal sentence) or in constructions known as raising – it functions only as markers of inferences or direct perception (often with epistemic "overtones"). Diewald (2001) and de Haan (2007), in their presentation of data from Germanic languages, do not concentrate on reportive meaning, nor do they point to this tendency, but it can be read between the lines. It becomes evident also in Cornillie's (2007) discussion of constructions and items based on Sp. *parecer*, and I have got the same impression from my own preliminary investigation of Russian and Polish forms based on SEEM-verbs. An analogous tendency can be observed with 'as if'-connectives (see 3.7). Of course, at present these observations allow only for the careful formulation of a hypothesis to be verified by diachronically oriented in-depth analyses of every single item.

Other source expressions frequently encountered in reportive parentheticals (particles ?) are the basic illocutive verbs of a language. Compare, again, It. *dice*, which in colloquial speech is still often used as a unit petrified jointly with the complementiser *che* (Giacalone Ramat/Topadze 2007: 27; Pietrandrea 2007: 55-57). In Southern Italy however it occurs also without a complementiser (C. Guardamagna, p.c.); see above the comment on (58a-d).

In fact, at least some parentheticals differ from particles in at least one important respect. Parenthetical clauses can contain anaphorical expressions which link the whole clause to their 'host' without exerting an influence on the syntactic structure of the latter; they can thus exhibit one-sided syntactic dependencíes on their hosts (cf. Dehé/Kovalova 2007: 6, basing on Ackema/Neeleman 2004: 96-99). This property brings them close to non-restrictive relative clauses, it is clearly associated to their function as metatextual comments (see f. 33). However, I am unaware of any such cases from among parentheticals with a reportive function. Of course, we know of numerous cases in which a grammatical form of a SAY/TELL-verb combines with connectives like Engl. *as*, Germ. *wie*, Russ. *kak*, Pol. *jak*, It. *come* 'as (people say)'. The question however is whether (and to whch degree) these syntagms can count as conventionalized holistic units and not just as productively used combinations of words to be analysed compositionally.

3.7. Conjunctions (complementisers)

In some of the languages of the Eastern part of Europe we encounter hearsay units that have to be qualified as complementisers or conjunctions. These units center around the meaning 'as if' and introduce sentential arguments of speech act verbs or link sentential modifications to NPs. Clear cases in point are Pol. *jakoby* (62) and Lith. *esą* (63):

(62) *Wspomniano mi, **jakoby** miał taki zamiar.*
      'I was told **that apparently** [lit. *as if*] he had such an intention.'

(63) *Seniai girdėti ir noriai cituojami argumentai, **esą** Baltijos valstybės yra Maskvos išlaikytinės.* ("Lietuvos Aidas" 1995-8)
      'For a long time we have been hearing arguments, **as if** the Baltic states were nurtured by Moscow.'

Lith. *esą* is the petrified obsolete neuter form of the present active participle of the copula *būti*, its etymology thus differs from units meaning 'as if' (Wiemer 2007b: 177-179). All these units however have developed further into "homonymous" particles (see 3.9); as for Lith. *esą* see (40), as for Pol. *jakoby* see the following example, in which *jakoby* combines with the complementiser *że* (cf. Stępień 2008: 327-330; Wiemer 2006b: 40-45; 2008: 365-368):

(64) *Paweł twierdzi, *że* Michał **jakoby** znał Marka.*
      'Paweł claims that Michał **allegedly/apparently** knew Marek.'

As for the genetically closest languages (Russian, Czech, Slovak, Ukrainian, Belarusian and Latvian, respectively), I have by now been unable to find text instances with cognate or functionally similar units that are used both as particles and conjunctions (complementisers) and in either case would mark h e a r s a y . In other words: hearsay conjunctions seem to belong to a rare species of lexical units, whereas etymologically related particles are more numerous, since they prove to be more apt to develop a reportive function (see 3.9). For instance, Letučij (2008) does not give a single example of Russ. *kak budto* as a conjunction in speech act contexts; Letučij (p.c.) considers such a use to be ungrammatical, contrary to *budto by* (e.g., *Govorili, *kak budto / budto by on uexal* 'They said **that** [lit. *as if*] he had left'; cf. also Wiemer 2008: 347-353). In the same vein Latv. *it kā* 'as if' does function as a complementiser, but, again, according to examples known to me (thanks to I. Kļēvere-Wälchli and N. Nau) and confirmed by J. Chojnicka (p.c.), it then introduces sentential complements in exclusively inferential contexts (perception verbs etc.), whereas as a particle *it kā* can mark hearsay (see ex. 41). It seems thus that, contrary to Polish and Lithuanian, the syntactic distribution of the Russian and Latvian 'as if'-connectives (particle vs. complementiser) tightly correlates with the evidential function (perceptive-inferential vs. reportive).

3.8. Adpositions

The last word class to be considered in connection with hearsay is adpositions. In European languages relevant units are encountered throughout, with prepositions being in the majority. Semantically adpositions modify NPs, whereas syntactically they are their heads, rendering PPs. Consequently, we have to distinguish their constituent-internal scope, which is purely adnominal, from the scope taken by the PP a s a w h o l e . PPs with reportive meaning always function as adverbials, their scope is therefore propositional. Adpositions themselves show various restrictions as for the admissible lexical noun class. Here I must restrict myself to a general remark on a coarse level of classification, which will be illustrated from Polish.

Among adpositions with a reportive function we should distinguish at least two types. In the first type the hearsay meaning is the only possible one, i.e., the adposition collocates only with names of persons (or groups of people), see (65a); this collocation restriction is loosened only in favour of names of institutions or bodies of people with some official function (65b). Adpositions of this sort cannot be used with names of products of speech, see (65c). Compare Pol. *zdaniem*:

(65a) ***Zdaniem*** *rzecznika rządu (przewodniczącego / naszych specjalistów / profesora ...) kryzys gospodarczy nie powinien wprowadzać nas w panikę.*
'**Due to the opinion of** the government spokesman (the chairman / our specialists / the professor …) the economical crisis should not cause panic.'

(65b) ***Zdaniem*** *zarządu (redakcji / komisji / rady ...) kryzys gospodarczy nie powinien ...*
'**In the opinion of** the board (the editorial office / the commission / the council …) the economical crisis …'

(65c) *****Zdaniem*** *ogłoszenia rzecznika rządu (artykułu w gazecie / raportu ...) kryzys gospodarczy nie powinien ...*
*'**In the opinion of** the governmental spokesman's release (the newspaper article / the report …) the economical crisis …'

Units belonging to the second type mark hearsay only by virtue of a specialized instantiation of a more global meaning consisting in a reference to the basis of judgment or measurement. For this reason markers of this type can collocate also with names of products of speech, e.g.,

written texts, announcements, dictionary entries (see 66c). They can occur even with names of other products of intellectual activity such as research results, public surveys, etc. (see 66d). This is where, for instance, Engl. *according to* belongs to (see the translations). Compare examples with its Polish equivalent *według*:[37]

(66a) **Według** *wychowawczyni (przyjacół / babci ...) dziewczynka ma wielki talent do rysowania.*
    '**According to** the house-mistress (the friends / grandma ...) the girl is very gifted in drawing.'

(66b) **Według** *komisji (rady nauczycielskiej ...) dziewczynka ma ...*
    '**According to** the commission (the teachers' council ...) the girl …'

(66c) **Według** *przepisu (podręcznika kucharskiego / wskazań sąsiadki ...) nie można dawać tak dużo drożdży do ciasta.*
    '**According to** the recipe (the cooking guide / the neighbour's advices …) one must not put that much yeast into the dough.'

(66d) **Według** *wyników badań (ostatnich sondaży / wniosków ...) posunięcia rządu spotykają się z coraz mniejszym poparciem obywateli.*
    '**According to** the research results (the last public surveys / the conclusions …) treffen die Maßnahmen der Regierung in der Bevölkerung auf immer weniger Unterstützung.'

As far as I can see, beside Pol. *zdaniem* only Lithuanian happens to have units fitting to the first type, namely: *pasak*, *anot* (cf. Wiemer 2007b: 183f.).

On the contrary, examples of units of the second type are commonplace. Here belong Germ. *zufolge*, *gemäß*, *laut*, Engl. *according to*, It. *secondo*, Sp. *según*, Fr. *selon*, Russ. *soglasno*, Bulg. *spored*, Croat. *prema*, Georg. *mixedvit*, *tanaxmad*, *cnobit* (Topadze 2008: 49).

3.9. Intermediate cases

In the preceding subsections I have rather casually pointed at cases in which the morphosyntactic classification of the units varies or causes problems. To a large part these cases can be characterized as "heterosemy". One of the first to use this term was Lichtenberk (1991: 476); for him heterosemy occurs "where two or more meanings or functions that are historically related, in the sense of deriving from the same ultimate source, are borne by reflexes of the common source element that belong in different morphosyntactic categories". Heine (1990) used the term 'heterosemy' to capture the result of a split of a unit's status on a lexicon—grammar cline if this split was not accompanied by erosion. Correspondingly, I apply 'heterosemy' to label a unit which, in an unaltered phonological shape, occurs in variable syntactic environments and thereby belongs to different syntactic morpheme or word classes without however changing its evidential value. As a rule, this syntactic variation reflects different layers of diachronic development (a phenomenon which in work on grammaticalization has been called 'divergence' or 'levelling'). For a similar approach cf. Autenrieth (2002).

We will first look at units that are on the verge of focal points of a morphological cline. In 3.5 we mentioned the Russian xenomarker *-de*. Normally it is called a particle, but we could equally well treat it as an agglutinative suffix. In fact, very few units in Russian behave this

---

[37] For further details cf. Wiemer (2006b: 32-38). In the last resort the semantic difference between these two types of hearsay adpositions can be captured by differences in metonymic shifts and conceptual distance: units of the first type allow only for shifts from 'person' to 'institution embodied by that person (or several persons)', i.e., to an immediately contiguous subdomain. Units of the second type allow for more drastic shifts (or chains of metonymic transfers).

way oscillating, as it were, between bound morpheme and clitic status, and this is consonant with the fact that Russian generally has very few true clitics.

We face an analogous phenomenon with the Turkish 'copula particle' *ImIş* (a term used among others by Csató 2000). *ImIş* is cognate to the *mIş*-suffix treated in 3.1. In contrast to this suffix, the evidential meaning of the copula particle is temporally indifferent, i.e., not restricted to anterior events and their results (Csató 2000: 37f.; Johanson 2000: 80). This functionally stable feature correlates with its distributional properties. Look at the following pair of examples given by Csató (2000: 37; her translations):

(67) *Ali* *hasta-ymış*.   (< *hasta* + *ımış*)
     Ali   ill:ImIş
     'Ali is / was / has (apparently) been ill.'

(68) *Ali* *hasta-la-n-mış*.
     Ali   become_ill:DNV.RM.MIŞ(= ANT)
     'Ali has (apparently) become / became ill.'

The *ImIş* particle can be attached to nominal stems (including extended non-finite verb forms) and thereby mark nominal predicates (67); this property brings it into complementary distribution with *mIş*, which is attached to primary and finite verb stems (the latter may be extended and derived from nominal stems); see (68). Furthermore, *ImIş* cannot be accentuated, whereas *-mIş* is capable of carrying pitch accent. From this we see that *ImIş* is less integrated into the word form than *-mIş*. Actually, this property seems to be the only argument in favour of counting *ImIş* as a clitic, not as an agglutinated affix.

Now let us continue with items difficult to classify in terms of syntactic distribution, inter alia because changes of word or morpheme class do not hinge solely on morphologization (coalescence). Such a case was noticed for Cz. *že*. Originally being (and still used as) the main complementiser void of any additional semantic elements, in colloquial speech it can be disconnected from speech act verbs after which it introduces sentential arguments (see 3.5). This can be considered as a change into particle status, which however is still bound to clause initial position. We can say that the categorial change of *že* has resulted from the loss of the syntactic link with a subordinating verb.

The mirror image to this process, i.e., the loss of an otherwise necessary complementiser following a speech act verb, has repeatedly been attested in Romance. Here we come across reportive modifiers on clause level, which originate from the fusion of a petrified form of a general SAY-verb (as a rule, the form of 3.SG.PRS.IND) with a complementiser. These are typical cases of lexicalization as understood by Lehmann (2002) and Himmelmann (2004), referred to in section 2. If these modifiers lose all their verbal features they should be classified as particles (or sentential adverbs; see 3.5). Interestingly, some of the relevant items have not stripped off all verbal features: although they have ceased to be analyzable into their original morphemes (root/stem+COMP), in some Romance varieties such uninflected function words can still take a sentential complement. This becomes obvious if they occur with another complementiser. According to Cruschina/Remberger (to appear, p. 11f.), such cases are attested in Sardinian and Romanian. Compare an example they give for Sardinian:

(69) *E* **nachi** <u>*chi*</u>   *issa* *no* *b'*   *andaiada*   *nudda*.
     and   SAYC   COMP   she   NEG there.CL go.IMPF.3.SG   nothing
     'And it **is said** <u>that</u> she didn't go there at all.'

Analogous behaviour can be observed with Rom. *cică* and colloquial Italian *dice (che)*; for the latter see 3.5. On the basis of the criteria formulated in 3.4 such units should best be classified as predicatives; complement-selecting clausal modifiers of this kind do not allow to name the

implicit human referent. This is also characteristic of particles (see 3.5), but particles do not combine with complementisers.

There is another class of predicates coming close to the definition of predicatives, which is enormously widespread in European languages: the 3.SG.PRES.IND form of SEEM/APPEAR-verbs, e.g. Eng. *it seems that P*, *es scheint, daß P*, Fr. *il paraît que P*,[38] It. *pare/sembra che P*, Pol. *zdaje się/wydaje się, że P*, Russ. *kažetsja, čto P*, Lith. *atrodo, kad P*, as well as Gr. *fenome* (Stathi, this volume), etc. By their etymology they denote non-agentive, non-specific perception, their „impersonal" forms regularly evolve into inferential matrix predicates with sentential complements, in many cases further into particles/parentheticals (cf. de Haan 2007 and Diewald 2001 on Germanic). Not every item among these crosslinguistic equivalents develops into a marker of hearsay, and there are borderline cases for which it is difficult to decide whether a reportive interpretation is to be considered as yet a context-conditioned implicature or already a conventionalized meaning (as for Russian and Polish cf. Wiemer 2006b: 55f.; 2008). But remarkably, the transition from inferential to hearsay meaning seems to correlate with a change from predicative to particle status.

A similar observation was made in 3.7 with respect to 'as if'-complementisers from which hearsay particles evolve. From the syntactic point of view they "mirror" units lexicalized from SEEM/APPEAR-verbs, since 'as if' often links complements to such verbs. We have here another prominent type of heterosemy, the most complex case probably being Lith. *esą*; although its etymology does not derive from 'as if', it functions like an 'as if'-unit and deserves special attention. Compare (40), where *esą* was used as a particle, with example (63) and the following one, where it links a clause to the verb *priekaištauti* 'to reproach' (cf. Wiemer 2007b: 177):

(70) *R. Ozolas sakė, kad jam bandoma priekaištauti, **esą** pasitraukus A. Šleževičiui, į jo vietą ateis kitas pareigūnas, kuris darytų lygiai tą patį.*
'R. Ozolas said that there had been attempts to reproach him, (for) **as people say** when A. Šleževičius will have retired, another official will occupy his place, who will do exactly the same.'
(Lietuvos Aidas 1995-9)

Moreover, *esą* can be encountered in contexts where it can be regarded as existential or copular predicate within a hearsay-context, although this happens much more rarely; see, for instance (cf. also Wiemer 2007b: 178, ex. 4):

(71) *Per patį šaudymosi ir sprogimų įkarštį (…) pro mus praėjo du vokiečių kariai ir pasakė, kad už jų jau slenka rusai, kurių **esą** labai daug.*
'In the heat of the gunfire and between explosions two German soldiers passed by saying that behind them the Russians are advancing, and **allegedly** there were many of them.'
(Z. Zinkevičius "Prie lituanistikos židinio", Vilnius 1999; p. 55)

This usage reflects the original use from which the conjunctional and particle use must have developed.

Furthermore, a challenge for an account in terms of word and morpheme classes is offered by some quotative markers. Engl. *like* is a good case in point. See the following example from colloquial English (Google, by courtesy of M. Lampert):

---

[38] The appearance of an expletive subject pronoun in English-type languages is due to their non-PRODROP-character and irrelevant for the argument, because expletive subjects do not fill argument positions.

(72) *How did you come up with the concept for "I Don't Need a Hook"? The concept came from one of my friends. My friend recorded it in his studio in Patterson, he's an upcoming producer, Dustin Hill, and he was **like**, "Oh, yeah, I have this beat for you!" And I'm **like**, "OK, let me hear it." And I'm **like**, "Wow, that's really a hot beat. I have to really get on that beat." He **like**, "Yeah, I think this song should be called 'I Don't Need a Hook.'" And I'm **like**, "OK." So, from there I just wrote the song and it's all history from there.*

Should we say that *like* is an adjective in predicative use and that the direct quote follows on it asyndetically? Or should we, instead, assume there to be an elliptical predicate, just as this could be argued for in the case of the Czech complementiser *že* when it started losing its connection with the speech act verb (see ex. 49-50)? On the latter analysis *like* could be considered as a particle. In contrast to Cz. *že*, it does not originate from a complementiser, but from an expression used in comparisons. We may say that its semantic evolution into a quotative marker has been accompanied by a syntactic change into an adverbial modifier of a zero predicate with very general meaning ('to say' or 'to do').[39]

    A similar point can be made for colloquial Russ. *tipa* 'like, as if'; morphologicaly this is the genitive of *tip* 'type, sort'. See the following internet example (Google, by courtesy of M. Makarcev; idiomatic translation approximative):

(73) *Blin, devki, mne segodnja mama takoe zarjadila. Posmotrela v telefon, a tam Tomik i govorit takaja **tipa** na tebja pohož!!! Ja čut' ne vypala!!!*
    'Hi guys, fuck you, today my mum just knocked me flat out. She looked at the phone, and there was Tomik and she was **like** he's like you!!! It almost threw me for a loop!!!'
    (http://forum.tokiohotel.ru/showthread.php?t=1316&page=28)

From the syntactic point of view we can hardly qualify *tipa* as the predicate, even if we do not assume a syntactic break after the preceding speech act verb *govorit* '(she) says'. We can regard *tipa* being coded along with a zero predicate. Zero predicates are a typical phenomenon of colloquial Russian, and speech acts are one of three semantic domains in which they occur particularly often.[40] If the zero interpretation proves an adequate option, Russ. *tipa* should be treated as a particle, too.

4. Recollecting the threads and data (in place of a summary)
As mentioned at the beginning, from among an enormously complex picture only some factors conditioning the meaning and the behaviour of hearsay marking devices could be discussed and shown to be relevant for a comprehensive typology of such markers, which would include not only grammatical means (bound morphology, auxiliaries, periphrastic TMA grams), but also lexical units. Since until very recently the latter ones have largely been neglected, the remarks made above could only to a lesser extent been based on firmly achieved linguistic knowledge. It is they where at many places I had but to "scratch" on observations whose empirical validity must be ascertained not only by thorough comparative in-depth studies, but also on the basis of a transparent methodology, which seeks to distinguish properly evidential from epistemic meaning components, pragmatic implicatures from non-detachable semantic elements as well as diverse subfunctions within evidentiality, and which is anxious to unite markers of different morphosyntactic format and status under a

---

[39] The same applies, e.g., to It. *dice* ('say'), which presents us with the syntactic "mirror image" of what happens to complementisers or adverbials.
[40] The other two domains are movements (as a rule, one-directional) and intense physical actions like beating and hitting.

common functional roof. The concluding resumption to follow now is meant only to highlight some important, recurrent phenomena and problems mentioned in section 3.

First, we have to stress that many markers with hearsay interpretations are not reportive units *per se*, but should rather be characterized as general markers of indirect evidentiality, since they do not differentiate between reportive and inferential functions. This indiscriminate semantics is a recurrent pattern of the following classes of markers:

(i)   evidential extensions of the (present) perfect; the Baltic languages and Estonian are an exception to this "rule".

(ii)   evidential extensions of the conditional: modern Italian is an exception. Since however the restriction of the Italian conditional to hearsay is a recent development, it rather confirms the indiscriminate nature of this mood within evidentiality. In this respect the Romance conditional strikingly differs from the German subjunctive, which is restricted to hearsay, as is its "analytic substitute" *würde* + infinitive.

(iii)   evidential extensions of modal auxiliaries of deontic necessity; compare Romance descendents of Latin *debēre* and Dutch *moeten*. For the latter a hearsay function is debatable, even more so for Germ. *müssen* and Sp. *deber*. It remains an open question whether evidential extensions of MUST-auxiliaries, if they occur at all, are acquired via epistemic necessity or directly from deontic necessity. Notice further that Italian and French modals of strong obligation acquire evidential functions rather in their conditional forms. Again, their pattern remarkably differs from Germ. *sollen* and West Slavic HAVE-verbs, which are representatives of weakened obligation and have extended only into hearsay, not into inferentiality.

(iv)   Throughout Europe the form of the 3rd person singular present indicative of SEEM/APPEAR-verbs extends into evidentiality, but a hearsay function becomes more salient only when this form, by losing its argument structure, starts being used as a particle (or parenthetical), i.e. without any complementiser. This allows for a careful hypothesis saying that reportive meanings correlate with the least possible integration of paradigmatically isolated verb forms into clausal syntax, whereas inferential functions of the same units are available if the respective unit still functions as a predicate with a sentential argument in a syntactically definable dependency relation.

(v)   The same applies to sentential adverbs like Engl. *apparently*, units in which a form of a SEEM/APPEAR-verb has fused with a complementiser, e.g. Sic. *parica* 'apparently' (Cruschina/Remberger, to appear: p. 21), or stable collocations like It. *a quanto pare* Squartini (2008: 927-939), which should be regarded as parentheticals.

(vi)   An analogous remark holds for 'as if'-connectives. They have to be seen as the syntactic mirror image of petrified forms of SEEM/APPEAR-verbs, insofar as they can serve as their complementisers. Again, hearsay function is typical rather for particles (not for complementisers) with an 'as if'-etymology. If 'as if'-units function as evidential c o m p l e m e n t i s e r s, they are either indifferent, since an inferential vs. reportive reading has to be figured out from the context (Pol. *jakoby*), or they are in almost complementary distribution with cognate units; e.g. Russ. *kak budto* (inferential, never reportive as complementiser and seldomly so as particle) vs. *budto by* (rather reportive if used as a particle, but also inferential if used as a complementiser). Georg. *turme, titkos, vitom* as well as Latv. *it kā* mark hearsay only as particles; Russ. *jàkoby* (a cognate of the Pol. *jakoby*) has become obsolete as a complementiser with inferential function and developed into a particle marking only hearsay. Notice that Lith. *esq* shows the same heterosemy as do 'as if'-units, but it is restricted to hearsay, probably due to its different etymology (petrified

copula going back to logophoric constructions). Curiously, I have not come across instances of evidential 'as if'-units outside East Slavic, Polish, Baltic and the Caucasus (Georgian, Armenian); apparently they are typical for (or even restricted to ?) the eastern half of Europe.

On the basis of these observations we can hypothesize that lack of discrimination between inferential and reportive functions is the result of a slow transition from interential to reportive meaning, in other words: that inferential meaning is acquired before shifts into reportive interpretations take place; in some cases this shift eventually ended with a stable, non-detachable hearsay component of the respective unit's meaning. A plausible pragmatic mechanism favouring such a transition is based on the need to ground one's inference on other people's assertions. For diachronic and discourse pragmatic arguments in favour of this explanation cf. Wiemer (2005; 2008) based on Russian and Polish.

As concerns grammatical marking in the strict sense, the overview in 3.1-3.2 showed that both the distribution of evidential extensions of TMA-grams (however marked) and the distribution of inflectional vs. agglutinative morphology with evidential functions do not scatter randomly. Instead, within Europe clear areal biases can be figured out: the eastern half of Europe „prefers" an extension of the perfect, whereas in the western „hemisphere" (continental Germanic and Romance) marked moods (subjunctive and its analytical substitutes in German and Dutch, conditional in Romance) are chosen as the basis of evidential extensions. The fact that Romanian in this regard patterns basically like the other Romance languages, although it otherwise has acquired many Balkanisms and Slavicisms due to language contact, demonstrates that contact-conditioned features (among others, functional extensions of TMA-paradigms) can be overridden, or resisted, by remarkable genetical stability (for a survey on Romance cf. Squartini 2001; 2005). On the other hand, Retoromance uses the subjunctive, not the conditional, as the rest of Romance, a fact which could be viewed as a result of German influence.

Agglutinated morphemes indicating hearsay, either as a central function or as a functional extension, are rare in Europe; they appear only in three languages of its south-eastern and eastern periphery: in Estonian, Albanian and Turkic.[41] Only Est. -*vat* specialises in hearsay, whereas the reportive function of Turk. -*mIş* and Alb. -*kam* is only a contextually conditioned reading of a broad evidential (including mirative) meaning potential. Georgian is the only language using inflectional morphology, as its perfect forms part of a system based on alternations of stems.

Finally, notice that there is no tight correlation between type of morphology and gram type: evidential extensions from the perfect can remain analytic constructions (Balkan Slavic, Baltic languages, Armenian), can betray signs of coalescence (Albanian), or be as agglutinative (Turkish) or inflectional (Georgian) as the respective perfect proper. Est. -*vat*, in turn, does not derive from the perfect, but is nonetheless agglutinated. Finally, we must remain aware that the extensions of mood paradigms into evidentiality attested in Germanic and Romance can count as inflectional markers of hearsay (or evidentiality in general) only insofar as these inflections developed long before evidential extensions started, and the latter ones have by no means led to a change in the endings of these moods.

List of abbreviations

---

[41] From a broader areal perspective the Turkic and Estonian (at the least) should be considered as the "western-most" representatives of a vast Eurasian territory (cf. Johanson 2000: 83f.; Wälchli 2000 and the contributions to Xrakovskij (ed.) 2007).

Acknowledgments

References

Achema, P., Neeleman, A. (2004): *Beyond Morphology. Interface Conditions on Word Order*. Oxford etc.: Oxford U.P.

Aikhenvald, A.Y. (2003): Evidentiality in typological perspective. In: Aikhenvald/Dixon (eds.): *Studies in Evidentiality*, …, 1-31.

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

Aikhenvald, A.Y. (2007): Information source and evidentiality: what can we conclude? In: Squartini, M. (ed.): *Evidentiality between lexicon and grammar*. (*Rivista di Linguistica* 19-1.), 209-227.

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/Nichols (eds.): *Evidentiality* …, 273-312.

Arutjunova, N.D. (2000): Pokazateli čužoj reči *de, deskat', mol*. K probleme interpretacii rečepovedenčeskix aktov. In: Arutjunova, N.D. (ed.): *Jazyk o jazyke*. Moskva: Jazyki russkoj kul'tury, 437-449.

Autenrieth, T. (2002): *Heterosemie und Grammatikalisierung bei Modalpartikeln (Eine synchrone und diachrone Studie anhand von »eben«, »halt«, »e(cher)t«, »einfach«, »schlicht« und »glatt«)*. Tübingen: Niemeyer. (LA 450.)

Boeder, W. (2000): Evidentiality in Georgian. In: Johanson/Utas (eds.): *Evidentials* …, 275-328.

Brendel, E., Meibauer, J., Steinbach, M. (2007): Aspekte einer Theorie des Zitierens. In: Brendel, E., Meibauer, J., Steinbach, M. (eds.): *Zitat und Bedeutung*. Hamburg: Buske, 5-25. (= *Linguistische Berichte*, Sonderheft 15.)

Breu, W. (1987): Albanisches und makedonisches Verbalsystem in kontrastiver Gegenüberstellung. *Balkan-Archiv, Neue Folge* 12, 279-305.

Breu, W. (2009): Mood in Albanian. In: Rothstein, B., Thieroff, R. (eds.): *Mood in the Languages of Europe*. Amsterdam, Philadelphia: Benjamins. (forthcoming)

Bulygina, T.V., Šmelev, A.D. (1993): Gipoteza kak myslitel'nyj i rečevoj akt. In: Arutjunova, N.D. (ed.): *Logičeskij analiz jazyka: Mental'nye dejstvija*. Moskva: Nauka, 78-82.

Burkhardt, D. (1999): Govorjaščee lico und časti(cy) reči: Überlegungen zu den vvodnye slova. In: Böttger, K., Giger, M., Wiemer, B. (eds.): *Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV)* 2. München: Sagner, 83-92.

Chafe, W. (1986): Evidentiality in English Conversation and Academic Writing. In: Chafe/Nichols (eds.): *Evidentiality* …, 261-272.

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

Cornillie, B. 2007. *Evidentiality and Epistemic Modality in Spanish (Semi-)Auxiliaries. A Cognitive-Functional Approach*, Berlin, New York: Mouton de Gruyter. (Applications of Cognitive Linguistics 5.)

Cruschina, S., Remberger, E.-M. (to appear): Hearsay and reported speech: Evidentiality in Romance. In: Damonte, F. et al. (eds.): *Proceedings of the Incontro di Grammatica Generativa 2008* (= *Rivista di Grammatica Generativa*). Padova.

Csató, É.Á. (2000): Turkish MIŞ- and IMIŞ-items. Dimensions of a functional analysis. In: Johanson, Utas (eds.): *Evidentials …*, 29-43.

de Haan, F. (1997): *The Interaction of Modality and Negation (A Typological Study)*. New York, London: Garland Publ.

de Haan, F. (1999a): *Southwest Journal of Linguistics* 18, 83-101.

de Haan, F. (1999b): Evidentiality in Dutch. *Berkeley Linguistics Society* 25, 74-85.

de Haan, F. (2007): Raising as grammaticalization: the case of Germanic SEEM-verbs. In: Squartini (ed.): *Evidentiality …*, 129-150.

Dehé, N., Kovalova, Y. (2007): Parentheticals (An introduction). In: Dehé/Kovalova (eds.): *Parentheticals …*, 1-22.

Dehé, N., Kovalova, Y. (eds.): *Parentheticals*. Amsterdam, Philadelphia: Benjamins.

Diewald, G. (1999): *Die Modalverben im Deutschen (Grammatikalisierung und Polyfunktionalität)*. Tübingen: Niemeyer. (RGL 208.)

Diewald, G. (2001): *Scheinen*-Probleme: Analogie, Konstruktionsmischung und die Sogwirkung aktiver Grammatikalisierungskanäle. In: Müller/Reis (eds.): *Modalität und Modalverben im Deutschen*, 87-110.

Donabédian, A. (1996): Pour une interprétation des différentes valeurs du médiatif en arménien occidental. In: Guentchéva (ed.): *L'énonciation mediatisée …*, 87-108.

Dooley, Sh., de Haan, F. (2006): Epistemic Modality and Evidentiality: Swedish *lär*. Univ. of Arizona, http://u.arizona.edu/~fdehaan/papers/laer.html/. (ms.)

Ehrich, V. (2001): Was *nicht müssen* und *nicht können* (nicht) bedeuten können: Zum Skopus der Negation bei den Modalverben des Deutschen. In: Müller/Reis (eds.): *Modalität und Modalverben im Deutschen*, 149-176.

Erelt, M. (2001): Some notes on the grammaticalization of the verb *pidama* in Estonian. *Estonian Typological Studies* 5, 7-25.

Faller, M. (2006): *Evidentiality below and above speech acts*. http://personalpages.manchester.ac.uk/staff/Martina.T.Faller/documents/Evidentiality.Above.Below.pdf (ms.)

Faßke, H. (1981): *Grammatik der obersorbischen Schriftsprache der Gegenwart (Morphologie)*. Bautzen: VEB Domowina.

Friedman, V.A. (2000): Confirmative/nonconfirmative in Balkan Slavic, Balkan Romance, and Albanian with additional observations on Turkish, Romani, Georgian, and Lak. In: Johanson/Utas (eds.): *Evidentials …*, 329-366.

Friedman, V.A. (2003): Evidentiality in the Balkans with special attention to Macedonian and Albanian. In: Aikhenvald/Dixon (eds.): *Studies in Evidentiality*, …, 189-218.

Giacalone Ramat, A., Topadze, M. (2007): The coding of evidentiality: a comparative look at Georgian and Italian. In: Squartini (ed.): *Evidentiality …*, 7-38.

Golato, A. (2002): An innovative German quotative for reporting on embodied actions: *Und ich so/und er so* 'and I'm like/and he's like'. *Journal of Pragmatics* 32, 29-54.

Grochowski, M. (1986): *Polskie partykuły (Składnia, semantyka, leksykografia)*. Wrocław etc.: Ossolineum.

Grochowski, M. (1997): *Wyrażenia funkcyjne (Studium leksykograficzne)*. Kraków: IJP PAN.

Grochowski, M. (2007): Partykuły, parenteza a wyrażenia metatekstowe. *Z polskich studiów slawistycznych, seria XI: Językoznawstwo*. Warszawa, 69-74.

Grünert, M. (2003): *Modussyntax im Surselvischen (Ein Beitrag zur Erforschung der Morphosyntax des Verbs im Bündnerromanischen)*. Tübingen, Basel: Francke.

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

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

Heine, B. (1990): *Grammaticalization Chains as Linguistic Categories*. Duisburg. (L.A.U.D., Series A, Paper No. 291.)

Himmelmann, N.P. (2004): Lexicalization and grammaticization: Opposite or orthogonal? In: Bisang, W., Himmelmann, N.P., Wiemer, B. (eds.): *What makes Grammaticalization? A Look from its Fringes and its Components*. Berlin, New York: Mouton de Gruyter, 21-42.

Hoffmannová, J. (2002): Intertextualita v mluvených projevech (Úvahy o metodologii výzkumu a konceptuální soustavě). *Stylistyka* XI, 371-381.

Hoffmannová, J. (2008): Role českého *prý/prej/pré* při vyjadřování reproduktivních a modálních významů. In: Wiemer/Plungjan (eds.): *Lexikalische Evidenzialitätsmarker …*, 149-165.

Holvoet, A. (2007): *Mood and modality in Baltic*. Kraków: Wyd-wo UJ. (Baltica Varsoviensia 6.)

Johanson, L. (2000): Turkic indirectives. In: Johanson/Utas (eds.): *Evidentials …*, 61-87.

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

Kaltenböck, G. (2007): Spoken parenthetical clauses in English (A taxonomy). In: Dehé/Kovalova (eds.): *Parentheticals …*, 25-52.

Karlík, P., Nekula, M., Rusínová, Zd. (eds.) (1995): *Příruční mluvnice češtiny*, Praha: Nak-ví Lidové noviny.

Kehayov, P. (2002): Typology of grammaticalized evidentiality in Bulgarian and Estonian. *Linguistica Uralica* 38-2, 126-144.

Kozintseva [= Kozinceva], N. (2000): Perfect forms as a means of expressing evidentiality in Modern Eastern Armenian. In: Johanson/Utas (eds.): *Evidentials …*, 401-417.

Kozinceva, N.A. (2007): Perfect forms as a means of expressing evidentiality in Modern Eastern Armenian. In: Xrakovskij (ed.): *Ėvidencial'nost' …*, 67-84.

Kronning, H. (2007): Les auxiliaires «mediatifs» en suédois (à la lumière du conditionnel épistémique et du futur aléthique en français). In: Begioni, L., Muller, Cl. (eds.): *Problèmes de sémantique et de syntaxe. Hommage à André Rousseau*. Lille : Éd. du Conseil Scientifique de l'Université Charles-de-Gaulle – Lille 3, 287-309. (ULC 3 Travaux et recherches.)

Kuz'mina, I.B. (1993): *Sintaksis russkix govorov v lingvogeografičeskom aspekte*. Moskva: Nauka.

Lehmann, Chr. (1995): *Thoughts on Grammaticalization*. München, Newcastle: LINCOM Europa.

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

Letučij, A. (2008): Sravnitel'nye konstrukcii, irrealis i ėvidencial'nost'. In: Wiemer/Plungjan (eds.): *Lexikalische Evidenzialitätsmarker …*, 217-238.

Lichtenberk, F. (1991): Semantic Change and Heterosemy in Grammaticalization. *Language* 67, 475-509.

Meibauer, J. (2003): Phrasenkomposita zwischen Wortsyntax und Lexikon. *Zeitschrift für Sprachwissenschaft* 22, 153-188.

Moore, C. (2007): The spread of grammaticalized forms: The case of *be+supposed to*. *Journal of English Linguistics* 35-2, 117-131.

Mortelmans, T. (2000): On the 'Evidential' Nature of the 'Epistemic' Use of the German Modals *müssen* and *sollen*. In: van der Auwera, J., Dendale, P. (eds.): *Modal verbs in Germanic and Romance languages* (= *Belgian Journal of Linguistics 14*), 131-148.

Mosegaard Hansen, M.-B. (1998): *The function of discourse particles (A study with special reference to spoken standard French)*. Amsterdam, Philadelphia: Benjamins.

Müller, R., Reis, M. (eds.) (2001): *Modalität und Modalverben im Deutschen*. Hamburg: Buske, 149-176. (*Linguistische Berichte*, Sonderheft 9.)

Mushin, I. (1997): Maintaining Epistemological Stance: Direct Speech and Evidentiality in Macedonian. In: Singer, K. et al. (eds.): *Chicago Linguistic Society* 33: *The Main Session*. Chicago, 287-300.

Palmer, F.R. 22001. *Mood and Modality*, Cambridge etc.: Cambridge U.P.

Pietrandrea, P. (2007): The grammatical nature of some epistemic-evidential adverbs in spoken Italian. In: Squartini (ed.): *Evidentiality…*, 39-63.

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

Plungjan [= Plungjan], V.A. (2008): Ob èvidencial'nyx pokazeljax čužoj reči i nedostovernosti v russkom jazyke: *mol*, *jakoby* i drugie. In: Wiemer/Plungjan (eds.): *Lexikalische Evidenzialitätsmarker …*, 285-311.

Rakhilina, E.V. (1996): *Jakoby* comme procédé de médiatisation en russe. In: Guentchéva (ed.): *L'énonciation médiatisée …*, 299-304.

Ramat, P. (1996): «Allegedly, John is ill again»: strategies pour le médiatif. In: Guentchéva (ed.): *L'énonciation médiatisée …*, 287-298.

Remberger, E.-M. (to appear): The Evidential Shift of WANT. In: Dechain, R.-M., Peterson, T., Schenner, M., Sauerland, U. (eds.): *Proceedings of the GLOW Workshop on evidentiality. UBC Working Papers in Linguistics*. Utrecht.

Rytel, D. (1982): *Leksykalne środki wyrażania modalności w języku czeskim i polskim*. Wrocław etc.: Ossolineum.

Roels, L., Mortelmans, T., van der Auwera, J. (2007): Dutch equivalents of the German past conjunctive: *zou* + infinitive and the modal preterit. *Cahiers Chronos* 17, 177-196.

Smirnova, E. (2006): *Die Entwicklung der Konstruktion* würde + *Infinitiv im Deutschen (Eine funktional-semantische Analyse unter besonderer Berücksichtigung sprachhistorischer Aspekte)*. Berlin, New York: de Gruyter. (Studia Linguistica Germanica 82.)

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

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

Squartini, M. (2005): L'evidenzialità in rumeno e nelle altre lingue romanze. *Zeitschrift für romanische Philologie* 121-2, 246-268.

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

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

Stathi, K. (2009): ** (this volume)

Stępień, M. (2008): O wzajemnym przenikaniu się ewidencjalności i modalności (na przykładzie wybranych polskich czasowników i wyrażeń funkcyjnych). In: Wiemer/Plungjan (eds.): *Lexikalische Evidenzialitätsmarker …*, 313-333.

Thieroff, R. (1992): *Das finite Verbi m Deutschen (Tempus – Modus – Distanz)*. Narr: Tübingen. (Studien zur deutschen Grammatik 40.)

Topadze, M. (2008): *Mezzi di espressione dell'evidenzialità in Georgiano*. Pavia: Università degli Studi di Pavia, Facoltà di Lettere e Filosofia. (unpubl. ms. of doctoral dissertation)

Traugott, E.C. (1989): On the Rise of Epistemic Meaning in English: An Example of Subjectification in Semantic Change. *Language* 65-1, 31-55.

Travis, C. (2006): *Dizque*: a Colombian evidentiality strategy. *Linguistics* 44-6, 1269-1297.

Usonienė, A. (2001): Veiksmažodžio *matyti* komplementų tipai: formos ir reikšmės sąveika. *Baltistica* 36-1, 115-124.

Wälchli, B. (2000): Infinite predication as a marker of evidentiality and modality in the languages of the Baltic region. *STUF* 53-2, 186-210.

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

Wiemer, B. (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, 107-131.

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

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

Wiemer, B. (2007a): Kosvennaja zasvidetel'stvovannost' v litovskom jazyke. In: Xrakovskij, V.S. (ed.): *Èvidencial'nost' v jazykax Evropy i Azii. (Sbornik statej pamjati Natalii Andreevny Kozincevoj)*. Sankt-Peterburg: Nauka, 197-240.

Wiemer, B. (2007b): Lexical markers of evidentiality in Lithuanian. In: Squartini (ed.): *Evidentiality …*, 173-208.

Wiemer, B. (2008): *Pokazateli s citativnoj i inferentivnoj funkcijami v russkom i pol'skom jazykax – kommunikativnye mexanizmy semantičeskogo sdviga*. In: Wiemer, B., Plungjan, V.A. (eds.): *Lexikalische Evidenzialitätsmarker ...*, 337-378.

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

Wierzbicka, A. (2006): *English. Meaning and Culture*. Oxford etc.: Oxford U.P.

Xrakovskij, V.S. (ed.) (2007): *Èvidencial'nost' v jazykax Evropy i Azii (Sbornik statej pamjati Natalii Andreevny Kozincevoj)*. Sankt-Peterburg: Nauka.