The role of the Czech words prý / prej / pré in expressing evidential (reportative) and epistemic meanings

This article undertakes to analyze the Czech particle prý

1) as a reportative "hearsay" marker, signal of quoting, reproducing, paraphrasing the speech of somebody else or one's own speech (sometimes taking part in textual transitions between direct and indirect speech)

2) as a marker of epistemic modality signalling the speakers' uncertainty, their doubts about the truthfulness or validity of the quoted utterance, their lack of belief in an unsubstantiated assertion.

Data from the Czech National Corpus enable us to study the combinations of prý with other modal particles (asi, možná, snad, údajně, zřejmě), with the modal verb mít, with some words and constructions with hearsay semantics (říká se, proslyčká se, pověst, fáma, legenda, tradice...), etc. The author tries to pinpoint syntactic positions that require prý absolutely (as a signal of quoting), and structures where prý seems to be redundant. From a stylistic point of view, the author especially pays attention to specific functions of prý in narratives and in press interviews. In addition to the corpus SYN2000, three Czech corpora of spoken language (Prague Spoken Corpus, Brno Spoken Corpus, Oral2006) have been used as sources of data for colloquial prej / pré (von prej, co bude k vobědu; schází se s Mirkem, prej). It appears that in spontaneous conversation prý particularly functions as an evidentiality marker while overlaps of evidentiality and epistemic modality are peculiar to written texts.
The major claim of this study is that the concord account for such sequences is in many cases more adequate than an analytic or a scopal one. I will argue that the concord readings are triggered by an overlap in the functions of the grammatical evidential marker and the epistemic or evidential word. The second important claim is that such concord readings provide a good analytical tool for diagnosing the functions of the grammatical evidentials of any language.

Section 2 introduces the major theoretical problems relevant to this study. Section 3 discusses the specific goals and the method used in the study. The types of interactions between the grammatical evidentials and the epistemic-or-evidential words are presented in Section 4. Section 5 looks for the triggers of what we call the 'concord (or holistic) reading', while Section 6 presents further evidence in support of the concord-hypothesis. Section 7 discusses in detail some general consequences instigated by the results of the previous sections. Section 8 summarizes the overall results of the study.

2. Key issues

Since the early 1980s there has been an upswing in interest towards delimiting the conceptual boundaries of the notion of 'evidentiality'. More specifically, the main puzzle has been (and still is) the relationship between the notions of 'evidentiality' and 'modality'. There is agreement among scholars that what is often referred to as 'evidentiality' is a somewhat Janus-faced category residing partly in the epistemic sector of the modal domain and partly outside it. One can distinguish between three major viewpoints as regards the relationship between these two notions: 1) (epistemic) modality and evidentiality are distinct functional categories with no overlap between them (see e.g. Nuyts 2001, 27-28; Aikhenvald 2003a; 2004, 7); 2) there is a submission relation between them, i.e. one of these notions includes the other one (see e.g. Chafe 1986; Kiefer 1994; Ifantidou 2001; Boye 2006: 21), and 3) there is an area of overlap, but no full subsumption (see e.g. Kozintseva 1994; Plungian 2001); the main candidate for such an area is provided by the notion of 'inferentiality' (see e.g. Palmer 2001, 8-9, 24; Dendale & Tasmowski 2001).2

For the purposes of this study, I will not presuppose any of the above possibilities. The only preliminaries applied henceforth are first, that these notions are conceptually graspable (if not necessarily linguistically discernable), and second, that for any form which expresses both meanings – epistemic judgement and reference to the source of information – one of these meanings is in a given context supposed to be pragmatically foregrounded or more prominent than the other.

2 In some studies the second and the third possibility are not differentiated.
The possibility of combining grammatical evidentials with epistemic or evidential lexemes within a single sentence is recognized and seen as a proof of the conceptual sovereignty of the notion of evidentiality (Aikhenvald 2004, 257). Unfortunately, there are virtually no studies exclusively devoted to such combinations. One may ask why such combinations should at all be of interest. In essence it boils down to the more general problem about the restrictions on combinability of two or more sentential modifiers, and lately there have been several attempts to formulate the principles underlying such restrictions (see e.g. Hengeveld 1989; Van Valin & LaPolla 1997, 40-52; Alexiadou 1997; Van Valin 2005, 19-21; Ernst 2006, 92-148). We will take a closer look at two studies which are of particular interest to us: Cinque (1999) and Nuyts (forthcoming). These studies are instructive not only because of their explanatory power, but also because they represent opposite theoretical traditions. Cinque (1999), who approaches the problem from the perspective of the Chomskyan generative paradigm, is interested in the relative order of adverbials and functional heads as a possible sign of universal phrase-structure constraint. Nuyts, on the other hand, working within the framework of Cognitive Linguistics, is interested in the combinability of such modifiers as an indication of the hierarchical nature of the qualifications of states of affairs. Both studies arrive at rather strict hierarchies.

To begin with, Cinque’s hierarchy is advanced as a universal hierarchy of clausal functional projections. The hierarchy derives from the observation that various classes of adverbs enter a strictly ordered sequence, and this ordered sequence coincides with the order of the dependent morphemes encoding various functional notions (such as mood, modality, tense, aspect and voice). It is then stipulated that these different classes of adverbs enter into a transparent Spec/Head relation with the different functional heads of the clause. Each specific class of adverbs (e.g. tense adverbs) is an overt manifestation of a distinct functional projection, which in certain languages may also be overtly expressed in the corresponding functional head position (e.g. as a tense affix). The restrictive sequential order of the adverbs or affixes is therefore nothing but a reflection of the hierarchical relations of the corresponding functional projections. Cinque’s list of projections is very comprehensive; Figure 1 presents only the first part of it (starting from left), which contains the projections ‘evidential’ and ‘epistemic’.

As can be seen from this figure, evidentials (lexical or grammatical) precede (lexical or grammatical) epistemic expressions. A direct indication for this relative order is the acceptability of (2a), where the evidential adverb evidently precedes the epistemic adverb probably, compared to the unacceptability of (2b) where the opposite order is present (see Cinque 1999, 135).

(2) (a) Evidently John has probably left.
(b) *Probably John has evidently left.

Cinque (1999, 141) claims that although many of the relative orders among such functional elements may eventually reduce to scope relations (as maintained in the functionalist literature; see e.g. Bybee 1985), not all orders are so explicable. If the relative order is determined by the fact that the notion of evidentiality is semantically superordinate with respect to the notion of epistemicity, then one should not expect (3) to be acceptable:

(3) It is probable that it is evident that he is the guilty one. Cinque (1999, 135)

He assumes therefore that the hierarchy above should be considered a property of the Universal Grammar rather than only a reflection of the semantic structure (see also Cinque 2006, 119-144 for discussion).

The hierarchy of Nuyts (forthcoming) (see Figure 2) is postulated in two-dimensional format, but due to space limitations we will present it here as one-dimensional, thus also simplifying the comparison with Cinque’s hierarchy. While in Cinque’s hierarchy the relation ‘A hierarchically higher than B’ was

---

3 Cinque’s notion of ‘functional head’ corresponds in this case to a dependent (mood, tense, aspect, or other) morpheme.

4 As the figure shows, Cinque draws a distinction between evaluative and epistemic modalities. Evaluative modalities do not refer to the degree of certainty in the truth of the proposition, “but rather express the speaker’s (positive, negative, or other) evaluation of the state of affairs described in it” Cinque (1999, 84). The following English adverbs could be considered evaluative: (un)fortunately, luckily, regrettaully, surprisingly, strangely/oddly (enough), (un)expectedly.

5 See Svenonius (2001, 211) for an explanation of examples like (3), which rescues the semantic scope account.
indicated with ‘A [B]’, in Nuyts’ hierarchy the relevant indexation is ‘A > B’. A basic primitive in this hierarchy is the cognitive-functionalist notion of ‘qualification’.

Figure 2. Hierarchy of qualifications of states of affairs [Nuyts, forthcoming]

_evidentiality > epistemic modality > deontic modality > time > quantificational aspect (frequency) > phrasal aspect (parts of the) STATE OF AFFAIRS_

Nuyts believes that although this hierarchy is part of syntactic and/or lexical semantic representation in grammar, it does not belong in grammar at all, but rather constitutes a primary dimension of human conceptualization (see also Nuyts 2001, 353-357). This means that although the hierarchy postulated is based on solely linguistic facts about scope relations between (and the grammatical behaviour of) qualificational expressions in language, it obviously reflects basic principles to a much greater extent, which are characterized by Nuyts as principles of human perception and conceptualization.

Both Cinque (1999) and Nuyts (forthcoming) arrive at their hierarchies using material which is strictly constrained with respect to the level of linguistic expression. As many other scholars, they explore the combinability of items, which are either lexical or grammatical, but not the combinability of lexical and grammatical markers. Therefore, by addressing combinations of grammatical and lexical markers of evidentiality and epistemicity respectively, we enter into an unexplored area in the research paradigm (see also Makarcz, this volume) for another contribution on this topic. The sequences of grammatical evidentials and lexical markers of epistemicity or evidentiality are significant because they involve more conventionalized meaning-to-form mapping than the sequences of two lexical markers. According to Mushin (2001, 170), we should expect ‘a much higher degree of conventional mapping between actual source of information and adoption of epistemological stance in languages with grammatical evidentiality than in languages which lack such systems.’ As a rule, the degree of conventionalization is mirrored in the frequency of the given pattern (see Hopper & Traugott 2003, 126-130; Brinton - Traugott 2005, 100). This means that we should expect in Bulgarian and Estonian, which have grammatical patterns of evidentiality, such combinations to be more frequent than for example in Russian, where no grammatical patterns of evidentiality exist.

Interestingly, Bulgarian and Estonian seem to allow all possible orders and combinations of epistemic words and grammatical evidentials, and therefore we are left with the possibility of drawing generalizations based on the semantic and pragmatic interpretations of such sequences, but not on their word order pat-
terns. Furthermore, looking at the degree to which such sequences correspond to the above hierarchies, one has to admit that in Bulgarian and Estonian we are dealing with tendencies rather than with rules. Provided that in example (1) we have a co-occurrence of the arguably epistemic lexical marker _май_ and the arguably evidential grammeme encoded as _l_-form (auxiliary-less past participle), we can immediately identify a certain discrepancy with the above hierarchies. It is embodied by reading (1b), in which the evidential is in the scope of the epistemic — a pattern which does not conform to the above hierarchies. Nevertheless, these hierarchies appear to be strong structural tendencies as readings like (1b) are extremely rare in Bulgarian and Estonian.

3. Description of the items under consideration

Before analyzing such sequences, we need to specify which particular Bulgarian and Estonian items appertain to the notions of ‘grammatical evidential’ and ‘lexical marker of epistemicity and/or evidentiality’. We can easily delimit the universe in which we operate by referring to notions used by the traditional descriptive grammars.

Let a sequence of a grammatical evidential and an epistemic or evidential lexical marker be a sequence in which a grammatical marker classified as ‘evidential’ by the descriptive grammars of Bulgarian and Estonian co-occurs with a lexical marker classified as ‘a word (adverb or particle) expressing epistemic assessment’. The latter includes both epistemic and evidential lexemes and this is due to the fact that descriptive grammars do not postulate a separate class of evidential adverbs or particles, but include such items in the class of modal (epistemic) words (see GBE II, 405-406, 494-495 for Bulgarian and EKG II, 187-190 for Estonian). With the help of this definition, we considerably restrict the array of sequential types factored in the study. First, we retrace those potential grammatical markers of evidentiality which are not descriptively promoted to a categorial status, and second, we retrace those lexical markers of epistemicity or evidentiality that belong to inflectional classes, such as verbs for example. The reason for this latter delimitation is that with non-inflectional classes it is easier and faster to look for natural examples on the Internet.

In Bulgarian, the grammaticalized evidential category _преки_нано наклоне_ ‘renarrative mood’ (or – for those who consider its mood status problematic – the class of _преки_нан жарм_ ‘forms of renarration’) is encoded by past active participles (ending in _-l_), which in third person singular and plural are not accompanied by the auxiliary verb _съм_ ‘to be’. Compare the minimal pair in (4);

(4a) conveys direct and (4b) indirect evidentiality:
From a typological perspective, the classificatory notions of ‘firsthand’ and ‘non-firsthand’ capture the difference between (4a) and (4b) and the Bulgarian evidentiality system corresponds to type A1 in Aikhenvald’s classification (see Aikhenvald 2004, 29-31). The ‘non-firsthand’ term comprises the meanings of reported evidentiality, inferentiality and mirativity, while the ‘firsthand’ term takes default reading in which the information is acquired through vision.

In contrast, the grammatical evidentiality system of Estonian represents the type A3 in Aikhenvald’s classification (see Aikhenvald 2004, 33). Type A3 encodes the distinction between ‘reported’ and ‘everything else’. The ‘reported’ term of the Estonian evidentiality system is labelled kaunud kõnevis ‘oblique mood’ and is manifested by the dedicated marker -vat suffixed to the first verbal form of the predicate. Compare the difference between the direct evidence in (5a) and the reported evidence in (5b):

    Jaan leave-PST.3SG Pärnu-ILL

    ‘Jaan left for Pärnu.’

    Jaan be-EV leave-PST.PTCP Pärnu-ILL

    ‘Reportedly, Jaan has left for Pärnu.’

As for the particular items characterized as ‘a word (adverb or particle) expressing epistemic assessment’ we will use the notion of epistemicity as a cover term even though some of these items have evidential meanings. It is well known that markers of evidentiality imply different degrees of certainty about the state of affairs under consideration. In other words, at this stage we will not distinguish between items with focal evidential meanings and items (such as evidential words) with only implicational epistemic meanings. Instead, we will catalogue lexical items according to the degree of certainty they express (or imply), i.e. we will assign each Bulgarian and Estonian item a rough position on an ‘epistemic scale’ (see Givón 1982; Akatsuka 1985; Nyuys 2001, 22 about this

We have 13 Bulgarian and 11 Estonian items whose co-occurrences with the relevant grammatical evidentials will be studied in the remaining part of this paper. It is worth noting that we will only be concerned with those cases where these words function as sentential modifiers, i.e. with cases where their scope is equivalent to the scope of the grammatical evidentials, which always operate at sentential level. This means that co-occurrences of grammatical evidentials and epistemic or evidential words where the latter have constituent scope are not factored in the present study.

An important caveat is that the borders between the four degrees of certainty are drawn intuitively, and it is possible that exclusive testing of the degrees of certainty would slightly modify the above classification, especially with regard to the middle area in the table. Another intuition is that the invariant meanings of the majority of the items in the table are epistemic. Only Bulgarian очевидно

<table>
<thead>
<tr>
<th>Degree of Certainty</th>
<th>Bulgarian</th>
<th>Estonian</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WEAK CERTAINTY</strong></td>
<td>едва ли ‘hardly, scarcely’, надо ли ‘hardly, scarcely’</td>
<td>vaevalt ‘hardly, scarcely’,</td>
</tr>
</tbody>
</table>
'apparently, evidently' and Estonian ilmselt 'obviously, apparently' and nõhvatavasti ‘apparently, evidently’ always convey evidential meanings. With regard to Bulgarian найверно ‘most likely, presumably’ and най ‘probably, it seems (that)’ as well as to Estonian õenäoliselt ‘probably’, they seem vague with respect to the distinction between epistemicity and evidentiality. As it is irrelevant for the purposes of this study whether these intuitions are fully sound or not, they will not be further tested.

The sentences where a grammatical evidential co-occurs with an epistemic or evidential word were collected from the Internet. To make the search as simple as possible, I looked only for co-occurrences where the epistemic or evidential lexeme immediately precedes or follows a verbal form grammatically marked as evidential.

4. Types of interactions

We already saw that there are four ways in which the sentence in (1) can be accessed. Accordingly, we will draw distinctions between four types of interaction between evidential grammemes and epistemic lexemes. In this section, I will discuss these types in greater detail, illustrating them with attested Bulgarian examples. Each type is introduced with a Roman number, cf.:

I. An epistemic lexeme is in the scope of an evidential grammeme. This type is exemplified in (6), an example from an online forum. The evidential form (the auxiliary-less past participle имах ‘have.PST.PTCP’ in bold) indicates that the speaker refers to someone else’s words. Although the referent is non-specific — what is referred to is rather the attitude of the ruling political class in the US —, the adverb несъмнено ‘undoubtedly’ is perceived as a part of the reported statement and is thus within the scope of the report. If EV stands for a grammatical marker of evidentiality, EP for an epistemic word, p for a proposition and square brackets indicate scope relations, this type can be formally represented as [EV [EP [p]]].

(6) Днес всички интересуващи се знаят, че това е била чиста лъжа. Чиста лъжа беше и поводът за obvious lie obvious lie be.IMPF.3SG also occasion-DEF for

неката в Ирак – Саддам несъмнено имах intervention-DEF in Iraq Saddam undoubtedly have.PST.PTCP

ОМП. Оказа се, че не е weapons_of_mass_destruction it_turn.AOR.3SG_out that NEG be.3SG

имах, ама какво от това.
have.PST.PTCP but what of this

‘Now all people who are interested know that this was an obvious lie. An obvious lie was also the excuse for the military intervention in Iraq – reportedly, there wasn’t any doubt that Saddam has weapons of mass destruction. Well, he didn’t have them as it turns out, but so what.’ (http://www.pro-anti.net/show.php?article=1&issue=735)

II. An evidential grammeme is in the scope of an epistemic lexeme. Consider the background of the story in (7). For a month or so Nasko has had some strange wounds on his legs. As he is not able to recover, he visits his Turkish friend whose grandmother says that Nasko will die if he does not follow her advice. She gives him the address of a certain imam and Nasko goes to see him. Despite the fact that the whole story is grammatically marked for indirect evidentiality (every predicate is in the form of auxiliary-less past participle), the evidential form да ‘give.PST.PTCP’ (in bold) is within the scope of the word май. As already noted, this word lends itself both to the domain of epistemicity and to the domain of evidentiality, and can be roughly translated as ‘probably’ or ‘it seems’. The speaker in (7) no longer remembers the details reported in the sentence containing the forms in bold. By embedding the reference to the source of information in the scope of май, he indicates that he does not trust his memory. This type can be formally stated as [EP [EV [p]]].

(7) Дала му координатите по някакъв ходка и той give.PST.PTCP him coordinates-DEF of some imam and he

отпишъл при него. Ходка-та му дал май go.PST.PTCP to him imam-DEF him give.PST.PTCP probably as it seems

никаква кърпа / не си спомням точно вие / да спи върху нея some piece_of_cloth NEG remember.1SG exactly anymore to sleep on it

или нищо такова и след това да му я занесе. or something like this and after this to him bring.3SG

‘(Reportedly), she had given him the address of some imam and he had gone to him. And then (I think the story went like this), the imam had given him a piece of cloth, – well, I don’t remember exactly anymore – to sleep on it, or something like this, and to bring it back to him after that.’ (http://forum.rozali.com/viewtopic.php?p=32156&sid=94472861ac1e962874a819e64d2620b9)
Examples like (6) and (7) show that in Bulgarian scope relations are not coded by word order. In both examples the direction of scope dependency is opposite to the relevant order of the grammatical evidential and the epistemic word with respect to the propositional core. The word order is also irrelevant for the description of the remaining two types of interaction. In these two types neither the grammatical evidential nor the epistemic word takes scope over the other one, which means that these items share exactly the same semantic scope. Although often disregarded in the study of evidentiality (and related categories), such cases are not exceptional cross-linguistically (see [Aikhenvald 2004, 87-95] and [Boye 2006, 191-194] for examples).

III. The two items modify independently the proposition; witness the example in (8). The sequence in bold indicates that although the speaker does not have direct evidence to show how educated Abraham was, he is confident (based on common knowledge) that Abraham was a very educated man. Here the sequence of the grammatical evidential and the epistemic word can be seen as a free collocation of two forms, which both maintain their functional identity – the first expressing reported evidentiality and the second certainty. This type can be formally stated as [EV [p]] & [EP [p]], i.e. as a conjunction of two qualifications of the status of p. As already noted in Section 1, I will conventionally speak in this case about ‘analytic reading’ of the sequence of evidential grammeme and epistemic word.

(8) Днес обаче знаем, че Авраам в никакъв случай не може да be.3SG place.PASS.PTCP at one level and compare.PASS.PTCP with PrimitiveDEF superstitious Bedouins on the contrary he be.PST.PTCP със сигурност високообразован човек, потънал на културно и with confidence highly educated man descendant of cultivated and высокоцивилизовано общество.

‘Yet, we know by now that there is no way in which Abraham can be put on the same level as the primitive and superstitious Bedouins. On the contrary, he is supposed to be, and we are confident about this, a well-educated man, a descendant of highly civilized society with great cultural achievements.’ (http://www.bgbible.sdabg.net/arheolog/a-6.htm)

IV. The two items are understood as a single entity, which as such includes both reference to the source of information and reference to the speaker’s epistemic judgement. Which one of these referential meanings prevails in this unified form, depends first on discourse-pragmatic factors and second on the genuine semantics of the items classified in the grammars as ‘evidential’ and ‘epistemic’. Using a hyphen as a mark of structural blending, this type can be stated as [EV-EP [p]]. Consider (9), in which the speaker addresses his forum mates, asking them for a favour. The string in bold cannot be disassembled into an expression of epistemicity and evidentiality. Rather, it is perceived as one unit, which as a whole conveys uncertainty based on indirect evidence. This effect is due to the fact that neither of the forms involved can be identified as only evidential or only epistemic; instead, both can express both meanings. We saw that the word май can be translated as ‘probably’ or ‘as it seems’. In addition, the auxiliary-less past participle бил (‘be-PST.PTCP’) conveys here inferentiality (an inference based on remote evidence) and has a scent of doubt in its presupposition. As a result, the two formal elements merge into one unit expressing hesitation and doubt over the state of affairs expressed in the proposition. As noted in Section 1, I will speak in such cases about ‘concord (or holistic) reading’ of the sequence of grammatical marker of evidentiality and epistemic word.

(9) Ако някой скоро ще проверя в архив-а в Търново [...] нека if somebody soon put check.3SG in archive-def in Tarnovo HOR да видя информация за име-то Цеко Иванов (или) Иванчов to see.3SG information about name-def Tseko Ivanov or Ivanov Draшански от гр. Бяла Статина бил се no Drašanski from town Byala Statina fight.PST.PTCP on фронтове-те на Добруджа през ПСВ, не знам в fronts-def of Dobrudja during First_World_War NEG know.1SG in коя част е бил или какъв вид е имал, which unit be.3SG be.PST.PTCP or what rank be.3SG have.PST.PTCP

Мисля че май бил убит от свой другар по think.1SG that may=as it seems be.PST.PTCP killed by his friend by посрещка докато оня нещо си изравя пушка-та, но това mistake while that something fix.PST.PTCP rifle-def but these са само догадки, be.3PL only guesses
'If anybody is going to the archives in Târnovo in the near future, please let him check for information about the name Tseko Ivanov (or) Ivanov Drašanski from the town of Byala Slatina, who fought on Dobrudja front during the First World War; I don't know in which unit or what rank he would have had. I think he may have been killed by mistake while his friend was fixing his rifle, but these are only guesses.'

(http://forum.boinaslava.net/archive/index.php/t-6691.html)

Note that when we mentioned ambiguity between different interpretations of (1), we did not refer to any actual ambiguity in the specific speech situation. Rather, we were concerned with the possibility of different interpretations of a particular sentence in different contexts. As can be seen from examples (6)-(9), the co-occurrences of grammatical evidential and epistemic word are usually given specific interpretation by the context, or in other words, they are disambiguated by the contexts.

In the next section we will focus on the properties of type IV, which is of main interest in this study, and we will try to identify the conditions which license such concord interpretations.

5. The triggers of the concord (or holistic) interpretation

The opposition 'analytic vs. holistic' access to a linguistic element was introduced into the study of grammaticalization and lexicalization by Lehmann (2002b). In philosophy of language, the doctrine of semantic holism defended by Quine (see Quine 1953) refers to the effect that a certain part of language can only be understood through its relations to a (already understood) larger segment of language. In our specific case, the concept of 'concord (or holistic) reading' refers to a very early phase of semantic coalescence of two formally distinguishable propositional modifiers. I will maintain that this effect is triggered by an overlap in the functions of what we referred to as 'grammatical evidential' and 'epistemic word'. Previous acknowledgments of this effect on co-occurring modals include Halliday (1970, 331), Lyons (1977, 807-808) and Coates (1983, 46, 138). In a recent study, Kasper Boye classified cases with two co-occurring epistemic markers', which share the same scope into four types (see Boye 2006, 189-196). His second type coincides with our notion of concord interpretation, cf.:

Two or more epistemic items or constructions which have overlapping meanings (...) may co-occur non-obligatorily in a unified expression of

---

6 In Boye's system evidentials belong to the epistemic meaning domain.

justificatory support for a proposition. [The co-occurring epistemic items or constructions cooperate in what may be thought of as the specification of one epistemic meaning,' (Boye 2006, 191)

5.1. Kinds of functional overlap

Let us first elaborate on the different ways in which the meanings of the linguistic elements may overlap. In both Bulgarian and Estonian the element characterized as 'grammatical evidential' is a constant and the element characterized as 'epistemic word' is a variable. This is obvious as in both languages there is only one grammaticalized category of evidentiality, but many lexical words expressing epistemic and/or evidential meanings. Therefore we may say that the variable 'epistemic word' takes different values in a single language. We already saw that one of the major distinctions that can be drawn between the items in Table 1 is that some of them have predominantly epistemic functions, whereas other predominantly evidential functions. Based on the distinction between words with primary (or foregrounded) epistemic meaning and such with primary (or foregrounded) evidential meaning we can distinguish between two subtypes of our type IV:


Once more, the index EP should not be misleading - it signifies a word which is classified in the descriptive grammars of Bulgarian and Estonian as expressing epistemic assessment. The indices EP and EV on the other hand stand for the actual functional value of the signatum of EP. The first subtype is exemplified in (10); the story is about the security surveillance cameras:

(10) Видеосъёмките на Лондон и Каир не са интегрирани video_surveillance_areas of London and Cairo NEG be.3PL integrated на определено ниво. После се оказа, че експлозив-бомба at certain level then turn_out(copy.3SG that explosive-DEF може би имал балкански произход – но дали maybe have.PST.PTCP Balkan origin but Q(yes/no) складове-не и въобще работа-та с врзаните материали у storehouses-DEF and in general work-DEF with explosive material at нас например се контролира по подходящ начин? us for_example monitor.IMP in adequate way
The areas with video surveillance in London and Cairo are not integrated at a certain level. It turned out then that the explosives were maybe of Balkan origin but who knows whether here (in Bulgaria, P.K.) for example the storage of the explosive materials and the work with them in general is monitored in an adequate way.

(http://newteck.orbitel.bg/computer/articles.php?issue_id=56&unit_id=495&article_id=1651)

We are dealing here with a semantically and pragmatically unified expression, although it can be formally deconstructed into an expression of inferential evidentiality inducing a certain amount of doubt, and into a corroborating adverb of middle certainty (може би ‘perhaps, maybe’).

Comparing the examples of concord reading (9) and (10), we can see that in both cases the function of the Bulgarian grammeme expressing indirect evidentiality is contextually specified as inferential. This observation is consonant with the view according to which the inferential meaning is located at the borderline of evidentiality and modality (see van der Auwera & Plungian (1998, 85); Palmer (2001, 8-9, 24-26); Dendale & Tasnowski 2001). It is crucial however, that the concord reading is licensed also in cases where the function of the evidential form is not specified as inferential. Witness (11), in which the writer sends a query to an online forum. The evidential form был ‘be.PST.PTCP’ conveys here a hearsay from unspecified source(s) and together with май ‘probably, it seems’, which is vague between epistemicity and evidentiality, expresses uncertainty in the truth of the proposition. Here also, I believe, a concord (or holistic) analysis is more adequate than an analytic one.

(11) Знае ли някой как стоят взгледът с цената на know.3SG Q(yes/no) anybody how stands question-DEF with price-DEF of хостинга за Java и PHP, за чертото искам да си правя hosting-DEF for Java and PHP since want-1SG to REF'L do.1SG нещата на Java, а не на PHP (не зная и не things-DEF on Java but NEG on PHP NEG it kwow-1SG and NEG искам да го уча ако може). Май бил want-1SG to it learn.1SG if possible probably=as it seems=be.PST.PTCP в пъти по-скъпи.

‘Does anybody know how the question stands with regard to the hosting price for Java and PHP since I want to do my things with Java, not with PHP (which I'm not familiar with and don't want to

learn if at all possible). It is supposed to be many times more expensive.’

(http://clubs.dir.bg/showflat.php?Board=java&Number=1937935356 &page=0&view=collapsed&sb=3&vc=1)

Further evidence for the claim that the concord reading is not triggered only in contexts marked for inferential evidentiality is provided by Estonian. As already noted in Section 3, the Estonian grammatical evidential only encodes the meaning of reported evidentiality. If the necessary condition for the concord interpretation were the inferential reading of the grammatical evidential, then one would not expect to find instantiations of this interpretation in Estonian. Nevertheless, such instantiations are attested. Consider (12) in a short commentary on black economy and prostitution, published in the Estonian newspaper ‘Sakala’ on 12 March 2004.

(12) Olen mitme tuttava kaest päris-nud, kas Viljandi-s be.1SG several acquaintance from ask-PST.PTCP Q(yes/no) Viljandi-INE on võimalik prostituaati tellida. Kõik nad on vasta-nud be.3SG possible prostitute order all they be.3SG respond-PST.PTCP õlakihituse-ga või õel-nud kõhkevalt, et mõne aasta eest shrug-COM or say-PST.PTCP hesitantly that some year ago vist ole-vat saa-nud kūll.

Perhaps be-EV get-PST.PTCP indeed

‘I have asked many of my acquaintances if it is possible to order a prostitute in Viljandi (a town in South Estonia; P.K.). All of them have answered by shrugging their shoulders, or hesitantly saying that some years ago it had (perhaps) been possible.’


Here the refusal to give a univocal answer to the question asked by the narrator, i.e. the avoidance of responsibility, is explicitly marked by the word kõhkevalt ‘hesitantly’. Therefore the sequence in bold is best accessed holistically. None of the other readings seems adequate. Reading this sentence one would hardly assume that the word vist ‘perhaps’ is in the scope of the report (cf. reading I). Even more unlikely is the reading with narrow scope evidentiality (cf. reading II), i.e. where the people answering the question are not sure whether they have heard the facts about prostitution or not. The analytic reading (cf. reading III) is discarded because the word vist ‘perhaps, maybe’ and the hearsay form ole-vat ‘be-EV’, which induces doubt, are felt too similar and somehow functionally
redundant, which means that we are not dealing with two independent qualifications of the status of the proposition.

The discussion so far brings us to the necessary conclusion that the functional overlap of the grammatical evidential and the epistemic words is larger than the area covered by the notion of inferentiality.

Consider now the second group of cases with concord reading – the subtype [EV-EP-EV [p]]. Here the grammatical evidential co-occurs with an evidential word and the resultant effect is different from the effect triggered by the first subtype. The difference is due to the fact that the functional domain covered by the grammatical evidential does not overlap with, but rather contains the functional domain covered by the evidential word. This is a natural consequence of the fact that lexical expressions are more specific in meaning than grammatical ones. Such sequences are therefore best accounted for in terms of specification and disambiguation (terms used by Boye 2006, 133, 191). Witness (13), where the grammatical evidential encodes a more general evidential term and the evidential adverb specifies/disambiguates the meaning of this term. The evidential form have.PST.PTCP expresses general indirect evidentiality and the adverb очевидно ‘apparently’ specifies it into inferential evidentiality and more specifically into an inference based on visual perception.

(13) Объедините впечатления та о състезание-то и от други discuss.AOR.1PL impressions-DEF from contest-DEF and from other неща случаещи се наоколо в една фургон, на студени things taking place around in one van at cold безалкохолни напитки. Charlie (който очевидно имал тежка non-alcoholic drinks Charlie who apparently have.PST.PTCP heavy вечер) прецени, че просто нещата нямаш така лесно да evening decide.AOR.3SG that simply things NEG.PUT so easily INF.COMP се избягат в главата му и бе принуден да си become_clear.3PL in head-DEF his and be.AOR.3SG forced to REFL. вземе бира.
take.3SG beer

‘We discussed our impressions of the contest and from other things going on while sitting in the van with our soft drinks, Charlie (who obviously had had a heavy night) decided that the things in his head wouldn’t become clear that simply and was forced to get himself a beer.’


A native speaker of Bulgarian may have doubts in the acceptability of the sentence containing the grammatical evidential form and the evidential adverb in (13). It may seem elliptic, in which case the omitted element would be the form бе ‘be.IMPF.3SG’ occurring between the two forms in bold. Adding this form, we would have a free collocation of the word очевидно ‘apparently’ and the regular form of pluperfect indicative бе имал ‘be.IMPF.3SG have.PST.PTCP’. One may speculate that the example in (13) is nothing but a case of auxiliary ellipse caused by carelessness. Nevertheless, similar occurrences are also attested in edited texts. An example is found in a historical survey of the uprising of the Bulgarian Paulicians in the eleventh century; cf. (14).

(14) Както разказва византийската историко Анна Комнина, през as narrate.3SG Byzantine-DEF historian Анна Комнина during 1084 г. (началото-то на въстание-то) Травъл бил вече 1084 year beginning-DEF of uprising-DEF Traval be.PST.PTCP already шест години лично и доверен ‘служа’ (т.е. един от близки-м sixth year personal and confidential servant i.e. one of close-DEF адютанти или оръженици) на нейния баща, прославени-м адъжантов или оръженици на нейния баща, прославени-м adjutants or armour-bearers of her father celebrated-DEF пътководец и император Алексий I Комнина (1081—1118). military_commander and emperor Алексios I Komnenos 1081-1118.

Като пресметнем време-то на тази служба, то поздравявам как calculate-1PL time-DEF of this position then PLOVD-ATTR павликянците очевидно бил сред най-верни-те хора на paulican apparently be.PST.PTCP among mostfaithful-DEF people of Алексий I Комнина още от 1078—1079 г. и е Алексios I Komnenos ever since 1078-1079 year and be.3SG преживявал заедно със своя ’патрон’ редица победи и experience.PST.PTCP together with his patron series of victories and triumphs, както и нямал тежки изпитания. Triumphant as well as and not a few hard probations

‘As is documented by the Byzantine historian Anna Komnene, in the year 1084 (in the beginning of the uprising) Traval had already for six years been a trusted personal servant (i.e. one of the close adjutants and armour-bearers) of her father, the celebrated military commander and emperor Alexios I Komnenos (1081—1118). If we add up the time of his service, then the Plovdiv-born Paulican had
apparently been one of the most trusted people of Alexios I Komnenos ever since the years 1078–1079 and had experienced together with his patron a number of victories as well as many hardships. (http://iteret.bg/publish13/p_pavelov/buntari/travy1.htm)

As already noted, in such occurrences the evidential word specifies/disambiguates the more general evidential term encoded by the grammatical evidential. This effect is possible only in Bulgarian, because the Bulgarian indirect evidential is such that it can be specified into narrower terms by the lexical items listed in Table 1. The Estonian evidential term is narrowly reported, and therefore co-occurrences with inferential adverbs like those in (13) and (14) cannot have concord readings, but only readings where the items in the sequence are in scopal dependency, cf. (15).

(15) Ta **ole-vat** nähtavasti tõö-l.
s/he be-EV apparently work-ADE

'It is said that s/he is apparently at work.'

This granted, we can now try to identify the licensing conditions of what we called ‘concord (or holistic) reading’. As we are dealing with overlapping domains, this can be done with the help of the tool from the set theory known as Venn diagrams. Before doing that, we will take for granted, that evidentiality and epistemic modality imply each other. Any explicit marking of the source of information correlates with the degree of the speaker’s commitment to the truth of the proposition, and vice versa, the degree of commitment correlates with the temporal or spatial distance between the state-of-affairs referred in the proposition and the participant which functions as a deictic center of the utterance. Consider now the types of intersection illustrated in Figure 3. We have only two sets, indicated respectively with A for the meaning domain covered by the grammatical evidential, and B for the meaning domain covered by the epistemic or evidential word. We will be interested here only in cases where the grammatical evidential and the epistemic (or evidential) lexeme co-occur and share the same semantic scope. Bearing this in mind we can identify the following types of intersections.

---

In the first diagram, there is no overlap in the meanings of the grammatical and the lexical item. In such cases we expect that their combination is accessed always analytically as a free collocation of two independent qualifications of the propositional content. This corresponds to our type III.

In the second diagram, the intersection is not an empty set, i.e. there is an overlap between the meanings of the two elements. In this case, their combination can be accessed holistically as subtype IV(a). The intersection comprises at least the notional domain of inferentiality, but does not seem to be restricted to it. As a result, the two elements are comprehended as somewhat redundant. In certain functionalist studies of modality (see [Lyons 1977, 807-808; Coates 1983, 45-46, 137-138; Palmer 2001, 35]), similar cases are labelled ‘harmonic combinations’ of two modals while in the formally orientated studies this phenomenon has been recently labelled ‘modal concord’ [Geurts & Huijtinck 2006] in parallel with the phenomenon called ‘negative concord’ where two overt negators yield a single operator. Halliday noted already in 1970 that in such cases the co-occurring modals reinforce each other [Halliday 1970, 331]. In our case, the ‘reinforced’ meanings are those meanings that are confined to the overlapping area in the diagram.

In the third diagram, one of the sets is confined within the other one. As already said, such cases can be accounted for in terms of specification or disambiguation. The item whose functional range is designated by the set A can be said to be underspecified, and the item whose functional range corresponds to the set B specifies it. This is our subtype IV (b), which as we saw, is attested only in Bulgarian. As the Estonian evidential grammeme encodes reported evidence, combining it with an inferential word would yield either scopal reading, i.e. reading in which either the inference is in scope of the report or the report is in the scope of the inference, or same-scope, but ‘analytic’ reading. If, on the other hand, the Estonian evidential is combined with a particle or an adverb expressing reported evidentiality, the meaning areas covered by these items would just coincide and bring about redundancy.

---

Footnote: This entails that evidentiality is a deictic notion, a position recently defended by de Haan (2005) that I agree with.
5.2. Medium certainty as a licensor of the concord reading

Let us now elaborate on the question of how the semantics of the particular epistemic and/or evidential words correlate with the different readings. In particular, we are interested to discover which items of those listed in Table 1 engage with the evidentials in combinations that can be accessed holistically. It seems that such readings are licensed only in combinations where the grammatical marker of evidentiality co-occurs with a lexeme expressing or implying a middle degree of certainty (or commitment) that the facts referred to by the proposition obtain. An apparent explanation for this tendency is that the evidential grammemes of Bulgarian and Estonian imply middle rather than full or low certainty. Using an expression of indirect evidentiality usually implies that the speaker does not want to commit himself with an opinion as to whether the reported state of affairs obtains or not. Therefore, in those combinations where the grammatical evidential and the lexical expression of middle certainty share the same scope, the latter are perceived as somehow superfluous. Due to the meaning intersection, the analytic reading is usually not selected and the only remaining way is that of convergence of the two items into a single unit, where the epistemic stance which can be paraphrased as 'I am not absolutely sure in p' or 'I do not fully commit myself to p' is reinforced.

With regard to the combinations with words expressing full or weak certainty, our evidence is that the first can receive either scopal (types I and II) or only in Bulgarian – analytic interpretation (type III), whereas the latter seem to be restricted to only scopal interpretation. The reason for this is that words expressing low certainty (or commitment) can be considered as implicit negators: qualifications like ‘hardly p’ normally imply ‘not p’. As a result such combinations yield clearly two distinct operators, one of which (evidentiality) is not truth-functional whereas the other (negation) is truth-functional, and hence the combination cannot be accessed holistically. These observations are illustrated in Figure 4 which follows the conventional distinction between four degrees of certainty used in Table 1.

![Figure 4. Correlation between the type of interaction and the degree of certainty adduced by the epistemic (or evidential) word in Bulgarian and Estonian.](image)

The distribution in the figure leads to a number of repercussions for the licensing conditions of the different readings. First, it is clear that the scopal interpretation is possible regardless of the degree of certainty adduced by the lexical word. In both Bulgarian and Estonian there were instantiations of scopal dependency between the grammatical evidential and the epistemic (or evidential) word all along the spectrum of certainty adduced by the latter.

With regard to the difference between the analytic and concord reading, it is not easy to decide between them without a thorough examination of the relevant context. The area on the scale where the analytic reading is available is larger or equivalent to the area permitting concord interpretation. This is a data-driven generalization, which can also be reached deductively. If we assume that any free collocation of two independent qualification expressions is prior to cases where these expressions are felt as parts of a single semantic unit, it follows that the analytic reading is licensed at least in the same environments where the holistic reading is licensed. In Bulgarian, cases where the grammatical evidential co-occurs with an epistemic or evidential word inducing full degree of certainty could only have scopal or analytic reading, whereas the concord reading is reserved for co-occurrences with words inducing medium-to-strong certainty. The instantiations of concord reading with words inducing strong certainty are rarer
and more specific than those with words expressing medium certainty. Strong
certainty items can only yield the type of concord reading that we characterized
as evidential specification (recall examples (13) and (14)).

Compared to Bulgarian, Estonian seems to be more restrictive with regard to
the availability of the analytic and the concord interpretations. Combinations
with words inducing strong certainty always receive scopal interpretation. For
example, the Estonian sentence in (16) where the grammatical evidential is com-
combined with an adverb expressing strong certainty can only receive scopal inter-
pretation:

(16) Ta ole-vat tõenäolselt tõõ-l.
     s/he be-EV probably work-ADE

*It is said that s/he is probably at work.

**It is said and I think it is probable that s/he is at work.'

The difference between Bulgarian and Estonian in the availability of the differ-
ent readings mirrors the functional range of the respective grammatical markers
of evidentiality. The Bulgarian evidential grammeme has broader semantics and
so the array of epistemic words with which it can form a ‘harmonic combi-
nation’ is larger than in Estonian. In Section 7, we will return to this regularity
and look for an operational device that can be used to detect the precise semantic
range of any evidential grammeme in any language.

5.3. Non-specific referent as a licensor of the concord (or holistic) reading

We noted in Section 5.1. that the concord interpretation is not restricted to cases
where the grammatical evidential receives an inferential specification, but also
to cases where it has hearsay interpretation. The analysis of the Bulgarian and
Estonian data pinpoints a specific condition that triggers the concord reading in
such cases of reported evidentiality. It relates to the identity of the referent of the
report. The grammatical evidential may be used in a context which specifies the
individual from whom the speaker has acquired the information concerning p or
in a context which does not specify the source of the report, but rather indicates
that the speaker has acquired the information about p from different sources at
different times or that determining the referent of the report is irrelevant in the
given speech situation. In the first case we can talk about a specific referent of
the evidential expression, in the second about non-specific referent of the
evidential expression.

In Bulgarian and Estonian, the concord reading is found in contexts with non-
specific referent. Consider the following example from Bulgarian. It comes from
an Internet discussion concerning an earthquake which took place the day

before. The earthquake was light; it was felt only by some people in certain
districts of Eastern Sofia, where the protagonist of the story lives.

(17) Аз пък си помислих, че съсед-а Тошо се е
     I but refl:DAT think.AOR.1SG that neighbour-DEF Tošo refl be.3SG
изтърсил по град. И нищо чудно, нами епицентър-а
     tumble_down.PST.PTCP on ass and nothing surprising isn’t epicentre-DEF
май бил в Младост ...
     probably-as it seems be.PST.PTCP in Mladost

‘And I thought that my neighbour Tošo fell on his backside. It
doesn’t surprise me, wasn’t the epicentre supposed to be in Mladost
(a residential area in Sofia; P.K.).
(http://muro.biz/old/?p=253)

It seems that the author of (17) has gathered the information about the epicentre
of the earthquake from one or several sources, none of which is fully reliable.
The referent of the evidential qualification is not contextually specified, which
in turn triggers an interpretation in which the hearer cannot identify any scopal
relation between the two elements in bold. Considering that these elements have
partly overlapping meanings, the speaker is left with the possibility for interpreting
them in terms of redundancy and reinforcement, in which case the concord
interpretation is activated. These stipulations are easy to check. If we insert into
the second sentence of (17) a clause referring to a specific source of information
about the epicentre then the scopal interpretation [EV [EP [p]]] arises leaving no
space for any other interpretation. In other words, if we added a clause like
‘according to my uncle’ to (17) the word mай automatically receives a narrow
scope interpretation with respect to the expression of evidentiality.

This observation has important consequences. It seems that utterances with a
non-specific referent of the report involve an inferential step by the speaker.
This means that the evidential qualification in sentences like (17) can be
paraphrased as ‘From what I have heard, I infer that p’ or ‘I guess from hearsay
that p’. In fact, it is this inferential step which binds the reported evidentiality
and the degree of certainty in a unified expression. It is interesting that Estonian
– a language with a grammaticalized term of reported evidentiality, also allows
such inferential interpretations in contexts with non-specific referent. Witness
(18), where the story is about the Eurovision song contest.
6. Further evidence for the ‘concord’ hypothesis

The reader may have noticed that I have so far not presented any empirical evidence showing that what was called the concord (or holistic) reading really exists. Speaking about the holistic reading presupposes at least some degree of formal bondeness, i.e. an increase of the intimacy with which the two collocating elements are connected to each other (see [Lehmann 2002b, 131-139]). In our case, however, there are no direct signs whatsoever of increased bondeness, and therefore we have to admit that the ‘holistic’-claim belongs to the realm of native linguistic intuitions that cannot be easily verified. The ‘concord’-claim, on the other hand, is a weaker version of the ‘holistic’-claim and therefore seems more suitable for the description of cases where a certain amount of semantic but not necessarily formal coalescence is at play.

One sign for the increase of the intimacy between the two elements is provided by their collocational frequencies. There is evidence indicating that the Bulgarian and Estonian grammatical evidentials show preferences with respect to particular lexical items. Table 2 presents some (preliminary) statistical data for the existence of such preferences. The first column of the table indicates the number of the most frequent collocations of the grammatical marker of evidentiality and epistemic or evidential word. In Bulgarian, the most frequent collocation is “evidential grammeme + mai”. This means that the number in the first cell indicates the frequency of the cases where the word mai ‘probably’ immediately follows or precedes the auxiliary-less past participle. In Estonian, the most frequent collocation is ‘evidential grammeme + vist’. Thus, the number in the second cell of the first column indicates cases where the word vist ‘perhaps, possibly’ immediately follows or precedes the evidential vat-form. The second column contains the respective numbers for the second most frequent collocations; in Bulgarian this is nepomno ‘most likely’ preceding or following the auxiliary-less past participle, and in Estonian ehk ‘maybe, perhaps’ preceding or following the evidential vat-form.

<table>
<thead>
<tr>
<th></th>
<th>the most frequent collocation</th>
<th>the second most frequent collocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulgarian</td>
<td>(ev. grammeme + mai) 183</td>
<td>(ev. grammeme + nepomno) 82</td>
</tr>
<tr>
<td>Estonian</td>
<td>(ev. grammeme + vist) 161</td>
<td>(ev. grammeme + ehk) 14</td>
</tr>
</tbody>
</table>
The distribution in the table is significant (chi-square=31.19425, p=0.000)\(^8\) showing that the lexical items май ‘probably, it seems’ in Bulgarian and виет ‘perhaps, possibly’ in Estonian are much more preferable in collocation with the relevant grammatical markers of evidentiality than any of the remaining lexical items. This, of course, applies only in case we assume that the overall frequency of these lexical items is similar. A Google search, which shows only approximate numbers, confirms that this assumption is more or less correct. The search for 'маи' yields approximately 1 680 000 Google hits and the search for 'вероятно' approximately 1 420 000 hits. In Estonian, 'вя' shows approximately 1 740 000 hits, and 'эх' approximately 1 500 000 hits. Thus, despite the similarity in the token frequencies of the Bulgarian май and вероятно, the first occurs over twice more often than the second in collocation with the grammatical evidential. In Estonian, this tendency is even more striking. Despite the similarity in token frequency of 'вя' and 'эх', the first occurs over eleven times more frequently in collocation with the grammatical gram than the second. The only conclusion that forces itself upon us is that the grammatical evidential is sensitive to a particular word, which expresses medium certainty and thus generates redundancy.

Is it interesting that native speakers of Bulgarian and Estonian often consider sentences containing combinations of the grammatical marker of evidentiality and a word expressing middle certainty somewhat overburdened and redundant if such sentences are out of the context, but if they are surrounded by the original context their acceptability to the speakers increases substantially, and the characterization of ‘redundancy’ is replaced with something like ‘motivated reinforcement’. This granted we can draw a parallel with the reinforcement occurring within the paradigms of grammatical evidentiality. The paradigm of the Bulgarian evidential is composed by aorist or imperfect past participles which can occur with or without the past participle of the auxiliary verb. The past participle of the auxiliary usually denotes criticism and distrust on behalf of the speaker (see [Demina 1959, 323; GBE II, 360]). Therefore some studies (such as [Nitsolova 2006]) postulate a separate paradigm of dubitative forms which is to be distinguished from the paradigm of the ‘renarrative’ forms. Compare the evidential renarrative form in (19a) with the evidential dubitative in (19b):

(19) a. май казал
he say,PST.PTCP
‘Reportedly, he said’...

b. май би бил казал
he be,PST.PTCP say,PST.PTCP
‘Reportedly (but I doubt it), he said’...

Due to its function to increase the distance between the speaker and the event referred to by the proposition, the participle of the auxiliary can be characterized here as a ‘distance particle’, a term originally used by Johanson (see Johanson 1998, 146). Such distance particles enhance the epistemic component in meaning of the compound.

Consider now an analogous case in Estonian. Besides the dedicated marker of evidentiality -vat, Estonian has the multifunctional modal verb pidama (with premodal meaning ‘to hold’) which in addition to dynamic, deontic and epistemic necessity is used as a marker of evidentiality. The dedicated marker -vat is often suffixed to this verb, which enhances the sense of doubt; witness (20) adopted from Kehayov (2002, 136). In (20a), where the suffix -vat is the only marker of evidentiality, we are dealing with a typical case of reported evidentiality, which may, but need not, imply a sense of doubt. In (20b) -vat is suffixed to the verb pidama and the cooperative meaning of the whole verb form is that of report accompanied by a stronger sense of doubt.

(20) a. Та ole-vat Tallinna-s.
s/he be-EV Tall/gin-INE
‘Reportedly, s/he is in Tallinn’

b. Та pida-vat Tallinna-s olema.
s/he must-EV Tall/gin-INE be-SUP
‘Reportedly (but I do not subscribe to this view), s/he is in Tallinn.’

Now, what is common between the cases in (19) and (20) and the combinations of grammatical and lexical items with concord interpretation is the semantic effect of reinforcement. The difference, on the other hand, between these cases is that in (19) and (20) we are dealing with a grammaticalized means of reinforcement whereas in the combinations of grammatical and lexical items with concord reading the relationship between the two elements is not grammaticalized.

---

\(^8\) The calculation tool uses Yates’ correction for continuity, which reduces the magnitude of the difference between expected and observed frequencies by 0.5.
It was stressed on several occasions that there is an implicational relation between the meanings of evidentiality and epistemicity as the cognitive remoteness of the source of information correlates with the degree of certainty. Despite this, only one of these meanings is considered a focal meaning (cf. Wiemer 2006, who draws a distinction between focal and associated meanings in his analysis of Polish lexical evidentials). Therefore, we can claim that in a complex sequence of two items with similar meanings, a certain meaning (evidential or epistemic) can be focalized both by grammatical or non-grammatical structural means. The only reason why the concord interpretation does not seem so obvious in cases where grammatical and lexical items are combined is that it is not overtly marked in the morphosyntax of the language.

7. Some consequences of general relevance

The effect of reinforcement can be compared with cases of reduplication where the property denoted by the repeated word is enhanced. If someone says good good dog it normally means that the dog is very good, i.e. the concept of ‘goodness’ is reinforced. We saw that certain combinations of ‘evidential’ and ‘epistemic’ yield an increase of doubt in the truth of the proposition, in which case we may say that the concept of ‘doubt’ is reinforced. Although this parallel might seem speculative, both examples involve reinforcement of a term. The very existence of such effect leads us again to the question whether we should look for an umbrella term for evidentiality and epistemicity. Such an umbrella term would be a narrowly defined functional category which encompasses only two subcategories: the one of evidentiality and the one of epistemicity. In such a way, the concord analysis provides another piece of evidence for the conceptual affinity of evidentiality and epistemic modality. Bulgarian and Estonian grammatical evidentiality systems are among the first discovered and best described evidentiality systems in the world. Moreover, due to the work of Jakobson (Jakobson 1971), the Bulgarian evidentiality system has played an important role in establishing the cross-linguistic category of evidentiality. It is thus beyond doubt that the cognitive basis that warranted an establishment of a new grammatical category distinct from modality in these languages is firm enough. Nevertheless, the surprising number of cases where the grammatical evidential and an epistemic word ‘reinforce’ a common meaning component urge us to reconsider whether even these languages do not warrant an umbrella term for the notions of evidentiality and epistemicity from which this meaning component can be abstracted.

The second important consequence emerges as we look at the size of the area of overlap of the linguistic elements which are said to be evidential and epistemic. We know from the previous research that Bulgarian encodes the broader term of ‘indirect’ evidentiality while Estonian encodes the narrower term of ‘reported’ evidentiality. This in turn leads us to the idea that the functional overlap between the grammatical evidential and the domain of epistemicity is larger in Bulgarian than in Estonian. This idea is articulated in Plungian’s notion of ‘modalized’ evidentiality system used for the description of the Balkan systems (Plungian 2001). Consider Figure 5, which illustrates the size of the overlapping area. In the figure we have two diagrams with overlapping rings. The left ring stands for the functional domain covered by the relevant evidential grammeme, whereas the right one stands for the domain of epistemicity.

Figure 5. The size of the epistemic component in the functions of the evidential grammemes

These diagrams portray the situation not only in Bulgarian and Estonian, but in any two languages with different semantic ranges of their grammatical evidentials.

Why is the size of the overlapping area important? First, in languages with minor overlap between the functions of the grammatical evidential and the domain of epistemicity we can expect co-occurrences with epistemic items to be more common than in languages with more significant overlap. The smaller is the overlapping area, the lesser is the chance for a functional ‘clash’ and redundancy. By virtue of the economy principle, the use of expressions with overlapping meanings is to be avoided.

On the other hand, in languages with greater overlap between the functions of the grammatical evidential and the domain of epistemicity concord interpretations are more likely to occur than in languages with minor overlap. If the area of overlap is larger, the relative frequency of the instantiations of concord reading – i.e. their percentage from the total of the attested co-occurrences with epistemic items – is expected to be higher. This is because a larger region of the semantic space is encoded either by more linguistic expressions than a smaller region or by the same number of linguistic expressions with more general
meanings. Both possibilities have frequent effect. In our case, the chance that the overlapping area is encoded linguistically is greater in Bulgarian than in Estonian.

Unfortunately, we cannot test these deductively achieved claims as we do not have comparable corpora for Estonian and Bulgarian. Nonetheless, these generalizations are significant as they seem to hold for any two languages with grammatical evidentials.

The evidence from Bulgarian and Estonian that we have looked at so far can be helpful if we want to elaborate a waterproof method with which we can identify the exact functional range of any grammatical evidential in any language. We will take for granted that any two markers with cognitively adjacent and/or partly overlapping meanings are subjected to specific restrictions as for their co-occurrence in the same sentence. Furthermore, we will adopt the assumption that if the co-occurrence of these two markers in a single sentence is accepted, there is still another set of restrictions which govern the semantic effects induced by such co-occurrence. Suppose we study the functions of a certain marker in a certain language. If we find out what the relevant restrictions are we would be able to determine precisely the functional boundaries of this marker and to better locate it in semantic space. Say we study the functions of a certain grammatical marker of evidentiality. Combining this marker with different items expressing epistemic modality helps to determine its functional range. In light of the above evidence, the following aspects should be taken into consideration:

1) Does the combination of the grammatical evidential with epistemic items generate concord readings?
2) If it does, with which particular epistemic items does this happen?
3) What is the position of these epistemic items on the scale of certainty?
4) Are there any epistemic items which particularly often enter into a concord relation with the grammatical evidential?

The fieldworker’s guides in Kozintseva (1994) and Aikhenvald (2004, 385-390) consider the compatibility of evidential and modal markers a relevant criterion for determining the type of evidential coding. The questions above could be considered as a supplement to these guides. By answering these questions we could test the functional boundaries of the grammatical evidential of any random language. The last question is crucial as for whether a particular combination of evidential and epistemic item is in process of becoming conventionalized. The concord readings of the collocations of Bulgarian and Estonian grammatical evidentials with the items maa ‘probably, it seems’ and viss ‘perhaps, possibly’, respectively, form the majority of cases where the collocation of grammatical evidential and epistemic or evidential word has concord reading in these languages. Moreover, as we saw in Table 2, these collocations make up a considerable share of the total amount of attested co-occurrences of grammatical evidential and epistemic or evidential word. It seems therefore that we are dealing with sensitivity between the grammatical evidential and a certain epistemic word, which might reflect an early stage of conventionalization of such complex expressions. An advanced stage of such conventionalization, on the other hand, would be a situation where their co-occurrence has become obligatory. Boye reports for a good number of cases from different languages where two epistemic or constructions with overlapping meaning co-occur obligatorily in a unified qualification expression (Boye 2006, 78-80, 189-191) and, following the postulates of the grammaticalization theory, we could assume that such expressions originate in non-obligatory syntagmatic patterns.

As a final point, it should be noted that the co-occurrences of grammatical and lexical expressions that we studied are surprisingly common compared to the co-occurrences of two grammatical or two lexical expressions of the relevant categories. We saw in Section 2 that there are a number of studies concerned with combinations of lexical or grammatical markers of epistemicity and/or evidentiality. We focused on the third possibility, namely on co-occurrences of grammatical and lexical marker. This choice turned out to be successful as we easily managed to gather a considerable body of examples. In contrast, both in Bulgarian and Estonian, combinations of grammatical evidentials with ‘epistemic’ moods, such as the conditional mood, are ungrammatical. The compatibility of two lexical means of evidentiality and epistemicity in Bulgarian and Estonian has not yet been studied, but the intuition is that many of the possible combinations are not acceptable. It seems therefore that if evidential and epistemic modifiers are combined at different levels of linguistic expression (lexicon and morphology, for example), they are more acceptable than if they are combined at the same level of linguistic expression. This pattern might be due to some general principle which blocks redundancies at the same level of grammar, but allows them at different levels.

---

9 I will refrain from presenting exact frequencies, because many of the examples of concord reading might be considered ambiguous, also allowing for other readings and which may therefore be considered controversial among speakers. It is clear, however, that they form the majority of all cases of concord reading.

10 Except from the short conditional forms in Bulgarian, which permit evidential marking (see GBE II, 370 for examples), but which are obsolete by now. These evidential conditional forms were in many cases homonymous with the relevant imperfective past participles and therefore do not qualify as adequate examples.
8. Conclusions

In this contribution, I outlined the types of interaction between grammatical markers of evidentiality and lexical markers of epistemicity and evidentiality. These types were stipulated based only on Bulgarian and Estonian data, but were assumed to be cross-linguistically pertinent. The following four types of interactions were distinguished:

1. The grammatical marker of evidentiality outscopes the epistemic (or evidential) word.
2. The epistemic (or evidential) word outscopes the grammatical marker of evidentiality.
3. The two items are not in scopal dependency, but represent two independent qualifications of the status of the proposition;
4. The two items are understood as parts of a single entity, which 'reinforces' a common meaning component.

This fourth type was referred to as 'concord (or holistic) reading' of the sequence of the grammatical marker of evidentiality and an epistemic (or evidential) word. The domain of medium certainty as well as the absence of specific referent of the report were shown to trigger the concord reading.

The central claim of this study was that the possibility for a concord reading of such sequences should be seriously considered in the description of the evidential and/or modal system of any language. It was argued that if one wants to determine the array of meanings expressed by a given grammatical evidential, the possibility for concord readings, the regularity of these readings and the range of the specific semantic effects should be explanatory as for the exact functions of this grammeme.

Abbreviations


References

Makarcev = Макарчев, М. М. "Связь лексических и грамматических показателей в индикативности в болгарском языке", this volume. 

