



# LINGUISTIC DISCOVERY

DARTMOUTH COLLEGE

Volume 8  
Issue 1  
2010

## **Commentary on de Haan - Evidentiality in Epistemic Modality: Let's Get the Whole Picture**

Heiko Narrog

Tohoku University

doi: 10.1349/PS1.1537-0852.A.367

url: <http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/1/xmlpage/1/article/367>

*Linguistic Discovery*

Published by the Dartmouth College Library  
Copyright to this article is held by the authors.

ISSN 1537-0852

[linguistic-discovery.dartmouth.edu](http://linguistic-discovery.dartmouth.edu)

# Evidentiality in Epistemic Modality – Let's Get the Whole Picture

## Comment on ‘Building a Semantic Map: Top-Down versus Bottom-Up Approaches’ by Ferdinand de Haan (2010)

Heiko Narrog  
Tohoku University

Although the papers by de Haan (2010) and Boye (2010) approach their topic from different perspectives, they essentially address the same issue: the relationship between epistemic modality and evidentiality. They differ in a number of details, such as de Haan’s emphasis on “bottom-up” procedure, which Boye assumes only implicitly, and the fact that de Haan accepts the idea of overlap between epistemic modality and evidentiality proposed by van der Auwera and Plungian (1998), while Boye rejects it. Nevertheless, they both essentially come to the same conclusion, namely that there is no “absolutely” correct answer to the question of categorial relationship between the two categories. This relationship ultimately depends on the choice of definition by the scholar—a definition that should observe, however, the criterion of contiguity of meaning (or of subcategories) as represented on a semantic map. That is, representation on a semantic map serves to illustrate that the category as it is defined is a coherent one.

Both papers are very welcome contributions to a relatively hot topic. Judgments on the relationship between the two categories have often not been very well-founded. As Boye poignantly notes, the currently most authoritative book on evidentiality (Aikhenvald 2004), for example, strongly claims a strict separation between evidentiality and epistemic modality without actually providing arguments in favor of this claim. Thus, in my view, both of the papers constitute an advance in the overall discussion. At the heart of the evidentiality vs. epistemic modality discussion, however, whether or not explicitly stated, often lies the question of how to categorize modal markers such as English *must* in Indo-European languages. They have been labeled as ‘epistemic’ for a long time, whereas in the past couple of years some scholars have started to reconceptualize them as ‘evidentials’. In my comments on de Haan’s paper here, this question, which always looms in the background of the epistemic modality/evidentiality debate, will be a central concern. I also intend to address a number of issues in the paper which I feel call for clarification, perhaps in a future publication where the author has more space to explain his reasoning.

The first set of issues concerns category labels and terminology. This may appear to be nitpicking, but the papers referred to here discuss the problem of category labeling after all, and in order for their approach to succeed, analysis and labeling at the level of what Cysouw (2010) calls “analytical primitives” is indispensable. One striking point in comparing Boye’s and de Haan’s paper is that although they essentially deal with the same semantic space, the category labels that they employ within that space, especially for epistemic notions, is so different. This immediately casts doubt on the validity of the analyses presented in the papers, and may furthermore even cast doubt on the validity of the whole enterprise of building semantic maps. How can the same space be represented so differently? Can these different representations be reconciled with each other? Frankly, it seems to me that the terminology in de Haan’s paper is more reason for concern than that in Boye’s paper. In de Haan (2010), three modal category labels are used for the three English modals discussed, and none of them seems unproblematic to me.

- 1) English *must* is labeled as an “evaluation of evidence”, and the author states that *must* has no other functions. The first and foremost problem is, of course, whether this is justified content-wise, that is, what the criteria are for classifying the use of a modal marker as “evaluation of evidence”, and whether this really applies to every instance of use, as is claimed here. The author refers to a different publication for the actual data analysis, and the reader cannot but trust that the answer is given there. However, purely with respect to category labeling, the term “evaluation of evidence” strikes me as lacking something. What is the actual result of this evaluation expressed by *must*? As evaluations are basically either positive or negative, does *must* express a positive evaluation or a negative evaluation? If the author had claimed that the speaker evaluates evidence, and an evaluation as “strong” would prompt the use of *must*, while an evaluation as weak or absent would prompt the use of *may*, this would be far easier to comprehend. It would also reflect the common notion that *must* and *may* stand in a paradigmatic relationship. That *must* should express the process of evaluation itself is hard to grasp. Would it really be meaningful to paraphrase a sentence like *This must be John* as ‘I evaluate evidence that this is John’? Why is a more common term such as “inference based on evidence” not applicable here?
- 2) *Bound to* is labeled as “strong epistemic modality”. The question that this label immediately begs is what the contrasting “weak epistemic modality” would be? *May*, perhaps? If so, would we have to conclude that the highly marginal *be bound to* (instead of *must*) stands in a paradigmatic or contrastive relationship to *may* in Modern English grammar?
- 3) *Will* is labeled as “predictive” even when referring to present and past state of affairs. Given that it is the goal of this research to build a semantic map “bottom-up”, wouldn’t it be more appropriate to start out distinguishing future-oriented ‘prediction’ from the purely epistemic present- and past-oriented uses (cf. e.g. Bybee et al. 1994:244, who define future through prediction, and thus link ‘prediction’ to ‘future’)? *Must* and *will*, according to the author, are both based on evidence. Is it possible then, that *must* in *John must be there*, and *will* in *John will be there* are nevertheless unrelated categories? This is what is suggested by the labeling (“evaluative” vs. “predictive”) and the spatial arrangement on the maps. Palmer’s (1990:36f.) solution to view them in a paradigmatic relationship of degrees of modality is intuitively more convincing.

The difference in labeling between Boye and de Haan can be at least partially attributed to the difference in the data on which the labeling is based. Boye’s terminology is based on an analysis of modality in 52 languages (although details of this analysis are not provided), while de Haan’s is based on a small selection of modal markers from English, Swedish and Dutch which are conceived as covering a related or adjacent area in semantic space. It seems to me that it is this isolated treatment of a few elements out of paradigmatic sets of markers which constitutes a fundamental problem in de Haan’s analysis, and that this finds a reflex in the terminological problems broached above. Are we to deny a paradigmatic relationship between the English modals as posited by Palmer (1990) or Halliday (2004)? What is the relationship between *must*, *will* and *may*? And if the focus is solely on epistemic modality expressing relatively high likelihood (epistemic necessity), why, for example, is the marginal *bound to* discussed, but not the more common *should*, *have (got) to*, or epistemic adverbs? Shouldn’t the latter have priority, at least in terms of description? It is easily conceivable that de Haan’s labeling and his classifications would have led to a quite different outcome if a more comprehensive picture of epistemic modality had been taken into consideration.

The central issue of the paper is the relationship between epistemic modality and evidentiality. Important contributions concerning this issue have been made by Nuyts (2001a, 2001b, and

others). Nuyts conceives of evidentiality as a dimension that is potentially present in all epistemic expressions,<sup>1</sup> and he covers in his study—in contrast to the paper discussed here—the whole range of epistemic expressions including modal verbs, adverbs, modal adjectives and mental verbs. Shouldn't this well-published and highly relevant approach at least be cited? In terms of methodology, Nuyts' approach of covering epistemic modality as a whole (and analyzing its expression comprehensively with the help of corpus data) seems to be extremely sound, and, decisively, it leads to different conclusions than those in de Haan's paper.

Let me return to the question of *must* here, and suggest an avenue of thinking about it (i.e. a hypothesis) which is closer to that of Nuyts, and which I believe deserves consideration. If *must*, as traditionally assumed, expresses something like a 'high likelihood' (or 'epistemic necessity') semantically, and 'strong confidence of the speaker' pragmatically, the dimension of evidence may be more or less entailed by this meaning? That is, if a state-of-affairs with undetermined factuality is presented as having 'high likelihood' or with 'strong confidence', isn't it simply the default case that the speaker must have some evidence for his or her claim? It is usually pragmatically odd to express strong confidence without any evidence. The expression of strong confidence without any evidence should be limited to specific pragmatic contexts, such as strong personal beliefs—including religious beliefs—which transcend the need for evidence. Therefore, the presence of evidence is commonly, but not necessarily, implied by the use of a marker like *must*. De Haan's hypothesis, by contrast, hinges on the assumption that *must* is always evidential, indicating an 'evaluation of evidence' ("we find no other function for English *must*"). This still needs to be backed up with evidence. In order to reject my hypothesis, it would be necessary to find commonly used grammaticalized markers of high likelihood like *must* which basically express a speculation without evidence. I am doubtful that *be bound to* would be a good example for this. First, it is extremely marginal (many treatments of English modality don't even mention it), and its marginality would rather confirm that an assessment of high likelihood is commonly linked to the presence of evidence. A further problem is that, similar to *must*, no usage (corpus) data are presented to support the author's claim as to its "unevidential" nature. A quick look at the British National Corpus reveals that doubt is in fact in order. Numerous examples of *be bound to* can be found in which, pace de Haan, evidence is present in the context, even within the same sentence:

- (1) While this downswing will not be as sharp as the previous downturn (in 1979-81), not least given the very much lower level of inflation that we now have, a dull 1989 *is bound to* be followed by a difficult 1990.
- (2) Probably a majority would recommend treatment of the female sexual partner(s) with at least the first attack of NSU, but, as with gonorrhoea where up to one third of female gonorrhoea contacts can be shown not to have the disease, such a policy of treatment without diagnosis *is bound to* lead to a certain amount of overtreatment.
- (3) With increasing longevity more families will become four and five generation ones, often influenced emotionally by some very elderly relative, so that development of this area of therapy *is bound to* be fruitful.

In an alternative view, *must*, *will*, and *may* are paradigmatically related to each other, with a declining degree of likelihood (and confidence) being linked to a declining amount (or degree) of evidence available for the modal judgment. That is, *must* is most strongly associated with the

---

<sup>1</sup>Note that Nuyts views evidentiality as determining the "subjectivity" of epistemic modal expressions, but this is a line of thought which does not need to be followed here.

presence of evidence, and *may* most weakly. But even *may* can have evidence as a background, cf. the following example from Collins Cobuild:

- (4) Carl and Martin may inherit their grandmother's possessions when she dies.  
The thought agitates her.

Again, due to the nature of the data, we are dealing with an isolated sentence here, but it is most probable that the judgment expressed by *may* is associated with evidence, such as the existence of laws which would lead to the realization of the state of affairs in the embedded proposition.

For category labeling, the decisive factor is the criteria according to which the data, i.e. each example, are analyzed. De Haan claims the function of an "assertion of evidence" for Swedish as illustrated by the following example.

- (5) *Några mål på hörnor och frislag lär det inte bli i VM*  
[11] any goals on corner.PL and free.kicks LÄR it NEG become in world.cup  
'There won't be any goals on corners or penalties in the world cup.'

Again, context is lacking, but the content of the proposition appears to be purely speculative, even contradicting common sense and available evidence (such as knowledge about frequent goal-scoring situations in football). In my view, (5) would be an excellent example not to support but to question the labeling "assertion of evidence". What would be the criterion to analyze this example as an "assertion of evidence"?

To conclude my comments, the basic point that de Haan makes (and Boye as well) is highly appreciated. The main objection that I have raised here specifically with respect to de Haan's contribution is that the analysis (and concomitantly also the terminology, a crucial point in this paper) may become distorted if only a few markers in a specific area are analyzed in isolation. I suggested instead to always keep the area as a whole in sight. Specifically, a dimension of evidentiality may be interacting with epistemicity as a whole, and this point is lost if markers are viewed in isolation from the modal systems in their respective languages. If an interaction between the two categories is acknowledged, there are still different ways to conceptualize the interaction between epistemic modality and evidentiality (cf. e.g. von Stechow and Gillies 2007, who claim that epistemic modals are in principle evidential markers). However, such interaction is a possibility which merits serious consideration. In this vein, I suggested that the expression of high likelihood ('epistemic necessity', 'strong confidence') is pragmatically closely associated with the existence of evidence in contrast to the expression of low likelihood. This association is not absolute but a strong tendency whose force depends on context.

The point that I am making here can also be extended to constructing semantic maps in general. The question is whether distortions may arise if the focus is on a very small area within a larger semantic area, and if data are drawn from just a small number of linguistic expressions (five markers in three languages, in this case). Wälchli's (2010) paper in this issue convincingly demonstrates the possibility of distorting effects from small sample sizes. The choice of a very small number of linguistic expressions may be justified by a particularly fine-grained semantic analysis, but this does not apply to the analysis presented in this paper. On the contrary, various uses of modal markers like *must*, as analyzed in detail by Coates (1983) and other previous researchers, are simply subsumed under a single label.

I have also expressed my concern about the accuracy of the data analysis for individual markers, which apparently could not be presented in more detail in the paper under discussion for reasons of space. A more in-depth analysis of the data and particularly delicate attention to labeling functions and meanings should be of particular importance for a bottom-up approach that relies on the investigation of a very small number of linguistic forms. The author will

presumably have the opportunity to demonstrate the validity of his analysis by presenting his data more comprehensively in other (future) publications.

## References

- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Boye, Kasper. 2010. Semantic maps and the identification of cross-linguistic generic categories: Evidentiality and its relation to Epistemic Modality. *Linguistic Discovery*, this issue.
- British National Corpus. World edition, 2001.
- Bybee, Joan, Revere Perkins and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Coates, Jennifer. 1983. *The semantics of the modal auxiliaries*. London: Croom Helm.
- Collins Cobuild on CD-Rom. 3rd revised edition, 2003.
- Fintel, Kai von and Anthony Gillies. 2007. An opinionated guide to epistemic modality. *Oxford Studies in Epistemology*, ed. by Tamar S. Gendler and John Hawthorne, vol. 2, 32-62. Oxford: Oxford University Press.
- Halliday, M.A.K. 2004. *An introduction to Functional Grammar*, 3rd edition. Revised by Christian M.I.M. Matthiesen. London: Arnold.
- Nuyts, Jan. 2001a. Epistemic modality, language, and conceptualization. Amsterdam: Benjamins. (Human Cognitive Processing 5).
- , 2001b. Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics* 33.383-400.
- Palmer, F.R. 1990. *Modality and the English modals*. Second edition. London: Longman.
- van der Auwera, Johan and Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2/1.79-124.
- Wälchli, Bernhard. 2010. Similarity semantics and building probabilistic semantic maps from parallel texts. *Linguistic Discovery*, this issue.

Author's contact information:

Heiko Narrog  
Graduate School of Information Sciences  
Tohoku University  
Kawauchi 41  
Aoba-ku  
Sendai-shi, 980-8576  
Japan  
[narrog@gmail.com](mailto:narrog@gmail.com)