Semantic Maps and the Identification of Cross-Linguistic Generic Categories: Evidentiality and its Relation to Epistemic Modality

Kasper Boye

University of Copenhagen

Semantic Maps and the Identification of Cross-Linguistic Generic Categories:
Evidentiality and its Relation to Epistemic Modality

Kasper Boye
University of Copenhagen

Cross-linguistic generic categories like evidentiality, tense, aspect, number, and person are entrenched in linguistic theory. However, it is not clear whether there is much empirical substance to them. There is a remarkable lack of criteria for what counts as a category. This paper tries to show that semantic maps can be used to give empirical substance to claims about cross-linguistic generic categories. It is argued that, as falsifiable cross-linguistic generalizations, semantic maps provide us with a criterion for categorial status and category membership and also provide us with a basis for identifying relations between different categories. However, it is also argued that there are limits to the use of semantic maps in evaluating claims about cross-linguistic generic categories, and that the criterion for categorial status and category membership provided by semantic maps ultimately needs to be supplemented by other criteria. In its argumentation this paper focuses on the category of evidentiality and on the relation between evidentiality and epistemic modality.

1. Introduction

In the most comprehensive study of evidentiality so far, Aikhenvald (2004) as one of her central points claims that "evidentiality is a category in its own right […]" (Aikhenvald 2004:7). Aikhenvald discusses a large amount of data that may be taken intuitively to support her claim. She also refers to studies that argue specifically in favor of setting boundaries between evidentiality and epistemic modality (de Haan 1999) or between evidentiality and mirativity (Lazard 1999, 2001, DeLancey 2001). But in fact, she does not present a single explicit argument in support of her claim. She does not give a single explicit criterion for what it means to be "a category in its own right" and is therefore unable to demonstrate that any criterion is met.

This is by no means meant as a critique of Aikhenvald alone. In general, discussions of cross-linguistic categories like evidentiality, tense, aspect, aktionsart, modality, number, and person are devoid of any interest in making precise what it is that qualifies them to be considered categories in the first place (cf. Boye and Harder 2009:15). The categories are normally taken to be notionally coherent. That is, each category is taken to cover a range of meanings that are similar in that they can be defined in terms of the same abstract notion— what Bache (1997:108) calls a "category concept". For instance, evidentiality is typically taken to cover meanings that can be defined in terms of the notion "source of information" (e.g. Willett 1988, Bybee et al. 1994, Aikhenvald 2004). But if similarity in meaning were all there was to categorial status, claims such as that made by Aikhenvald would be trivial. Aikhenvald's claim itself would be nothing but a statement that there is a range of meanings that can be defined in terms of the notion of source of information.[1] When we talk about cross-linguistic categories like evidentiality, tense,

---

[1] Aikhenvald explicitly considers evidentiality a grammatical category and thus takes it to cover only meanings of grammatical expressions (for instance, affixes or auxiliaries) (e.g. Aikhenvald 2004:11). But this has no impact on her claim that evidentiality is a category in its own right. It is interesting, of course, that the notion of "source of

aspect, etc., we imply that there is some adequate generalization over language-specific facts which goes beyond similarity in meaning. We imply that the notional coherence of a range of meanings is linguistically significant in some way.

Following Whorf (1956:113), I shall refer to cross-linguistic categories like evidentiality, tense, and aspect as cross-linguistic "generic" categories. In what follows, I argue that semantic maps can be used to give substance to claims about such categories. More precisely, I argue that, as falsifiable cross-linguistic generalizations, they may provide us with an empirically based criterion for categorial status and category membership as well as for identifying relations between distinct categories. In my argumentation, I focus on Aikhenvald's aforementioned claim about evidentiality. In particular, I try to show how a semantic map of epistemic expressions can provide substantial input into the ongoing debate on the relation between evidentiality and epistemic modality (van der Auwera and Plungian 1998, de Haan 1999, Dendale and Tasmowski 2001b; Guentchéva and Landaburu 2007b).

In Section 2, I specify what I mean by the notion of "cross-linguistic generic category" and argue that the notion, or any equivalent notion, cannot be dispensed with in linguistic theory. In Section 3, I propose what I shall refer to as "semantic-map continuity" as one criterion for identifying cross-linguistic generic categories. In Section 4, subsequently, I present a semantic map of epistemic expressions. And in Section 5, I show how the criterion and the map can be used to evaluate claims about evidentiality and the relation between evidentiality and epistemic modality. In Section 6, I go on to discuss the limits to semantic-map continuity as a criterion for identifying cross-linguistic generic categories, and eventually I propose an additional criterion. Finally, Section 7 is a brief conclusion.

2. Cross-Linguistic Generic Categories

The term "category" serves a number of different purposes in linguistic theory. For instance, the term is used both to designate specific linguistic "values" like past, present, and future—Whorf's (1956:113) "specific" categories—and to designate "groups of related values" like tense or time—Whorf's (1956:113) "generic" categories;[2] it is used both in descriptions of language-specific structures and in cross-linguistic generalizations; it is used both for linguistic phenomena and for the cognitive and communicative phenomena that are assumed to motivate them; it is used both to talk about purely theoretical constructs and to talk about entities that are assumed to have an ontological reality. And it is far from always clear in which way the term is used.

In the present paper, the term "cross-linguistic generic category" is restricted to entities like evidentiality, tense, and aspect. It is taken to refer to cross-linguistic generalizations over distinct

---

information" seems to provide a generalization over grammatical expressions in a large number of languages (cf. e.g. Chafe and Nichols 1986, Willett 1988, Guentchéva 1996, Dendale and Tasmowski 2001a, Aikhenvald and Dixon 2003, Guentchéva and Landaburu 2007a). More generally, it is an interesting hypothesis that cross-linguistically the meanings of grammatical elements can be described in terms of a limited range of notions (Slobin's (1997) "grammaticizable notions"; cf. e.g. Heine et al. 1991:32-39, Bybee, et al. 1994:10, Croft 2003:225). Surely, however, these observations and hypotheses are independent of any of the normal implications of claims about categorial status.

[2]Whorf characterizes his distinction between specific and generic categories as follows: "SPECIFIC CATEGORY: an individual class of [...] types, e.g. passive voice, durative aspect, vs. GENERIC CATEGORY: a higher hierarchy formed by grouping classes of similar or complementary types, e.g. voice, aspect" (Whorf 1956:113). While the distinction is a cornerstone in linguistic theory, it is rarely made explicit, and Whorf's terminology is used even more rarely. In a recent paper, for instance, Haspelmath (2007:123) refers to generic categories as "category-systems".

language-specific expressions that are related within as well as across the languages in which they are found. In so far as cross-linguistic generic categories are based on relations between expressions across languages, they must take the form of meaning generalizations—abstract meaning categories; arguably, linguistic expressions can be related across languages only in terms of their meanings—as opposed to their morphosyntactic properties.[3] But in so far as they are also based on relations between expressions within specific languages, cross-linguistic generic categories are endowed with a claim of being significant for the description of language-specific phenomena that do not necessarily pertain to meaning only. The assumption that cross-linguistic generalizations reveal something about—and may thus be accounted for in terms of—cognitive and communicative structures is well-founded, in so far as human language is understood as a cognitive and social phenomenon. However, this does not entail that cross-linguistic generalizations can be straightforwardly interpreted as descriptions of cognitive and social structures, as is often done. In the present paper, cross-linguistic generic categories are therefore conceived of as pertaining to linguistic phenomena only, as opposed to cognitive or social phenomena. And as generalizations over language-specific facts, they are conceived of entirely as the linguist's theoretical construct rather than as corresponding to any entity which is assumed to have an ontological reality.

Not all linguists would subscribe to the view that there is a need for cross-linguistic generic categories in linguistic theory. Among others, Bybee and her collaborators have argued that categories like evidentiality, tense, and aspect can easily be dispensed with in language-specific and therefore also cross-linguistic description. The argument turns on the traditional association of those categories with morphosyntactically delimited systems of linguistic expressions. While such systems are often found, the argument goes, they are usually not notionally coherent (Bybee 1985:17), and, whether or not they are notionally coherent, they must be considered "epiphenomenal": "the clues to understanding the logic of grammar are to be found in the rich particulars of form and meaning and the dynamics of their coevolution" (Bybee et al. 1994:1-2). For Bybee and her collaborators, linguistic universals are found only at the level of values like past, present, and future, as opposed to the level of groups of related values like tense (Bybee et al. 1994:3). They do not reject that values fall into distinct, notionally coherent groups. In fact, they readily group the universal values they study, "grams", into notionally coherent groups: "aspects and tenses", "agent-oriented modalities", "moods" and "evidentials" (Bybee et al. 1994:316-324). What they reject is the usefulness of any generalization over distinct linguistic values which, like Aikhenvald's category of evidentiality, is intended to go beyond registration of notional coherence.

However, notional generalizations are in themselves essentially intuition-based and subjective. Bybee et al.'s generalization over meaning paraphrases such as "direct evidence" and "indirect evidence" in terms of the notion of source of information may intuitively seem perfectly sound. But if the notional generalization itself were all we were left with, there would be no rational way to judge whether the generalization is adequate or not. If it were impossible to come up with empirical constraints on notional generalizations, we would have to abandon as

---

[3]Morphosyntactic properties are purely structural phenomena whereas meanings are not. As argued by the old structuralists and much more recently by functional typologists (cf. e.g. Haspelmath 2007 for a recent statement), structure is language-specific whereas substance is potentially universal.

Just like linguistic meanings, phonetic features can also be compared across languages in terms of a common substance. Thus, linguistic expressions can be related across languages also in terms of their phonetic properties. However, phonetic relations are irrelevant for the identification of cross-linguistic generic categories.

unscientific the use of terms such as "evidentiality", "tense", and "aspect" from linguistics. Bybee et al. would probably agree that if these terms were deleted from their book, it would have fatal consequences for its readability. Luckily, it is possible to come up with empirical constraints. Below I give one reason why cross-linguistic generic categories, in the sense specified above, cannot be dispensed with in linguistic theory.

The reason has to do with the fact that, as acknowledged by Bybee and her collaborators, some language-specific systems are actually notionally coherent. Whether it is an epiphenomenon or not, the fact that notionally related linguistic expressions occasionally, or perhaps even often, cluster in morphosyntactically delimited systems does not follow trivially from their notional relation. Consider, for instance, the often-cited Ngiyambaa (Australian) clitics *-gara* 'sensory evidence' and *-dhan* 'spoken, or by extension written evidence', illustrated in (1) and (2) respectively.

Ngiyambaa (Australian; Donaldson 1980:275-276)

(1)   *ŋindu-**gara**   girambiyi.*
      you.NOM-SENSEVID   sick.PST
      'One can see you were sick'.

(2)   *ŋindu-**dhan**   girambiyi.*
      you.NOM-REPEVID   sick.PST
      'You are said to have been sick'.

These two clitics make up a morphosyntactic system which is delimited by two properties: 1) The clitics mutually exclude each other (Donaldson 1980:244); 2) They always follow so-called "knowledge clitics" (Donaldson 1980:241) when they co-occur with them.[4] This system is notionally coherent in so far as both clitics have meanings that can be described in terms of the notion of source of information. What is important here is that systems like this, that are both morphosyntactically and notionally coherent, need not exist. Theoretically, all expressions of source of information in the languages of the world might either be completely absent in morphosyntactic systems or be found in systems that are not notionally coherent. For instance, one might perfectly well imagine a situation where there were no morphosyntactic properties that delimit the two clitics mentioned above from the Ngiyambaa "knowledge clitics". In support of this, it is far from always being the case that the morphosyntactic properties which delimit a notionally coherent system can be straightforwardly accounted for as a consequence of the notional feature in terms of which the members of the system are related. The second of the two properties that delimit Ngiyambaa *-gara* and *-dhan* from other Ngiyambaa clitics cannot. The fact that the two clitics are mutually exclusive can hardly be seen as a consequence of the fact that both clitics express source of information. In many languages, two or more expressions of source of information may co-occur. This is the case in Eastern Pomo (Hokan), where "the hearsay or reportative evidential" suffix *-·le* can follow one of two other evidential suffixes, including "the inferential evidential" suffix *-(i)ne*, as in (3) (McLendon 2003:109-113).

---

[4]According to Donaldson (1980), the Ngiyambaa "knowledge clitics" comprise *-wa* 'exclamative, what' and *-ga* 'ignorative, I don't know'. To judge from Donaldon's glossings, these clitics are notionally clearly distinct from the clitics illustrated in (1) and (2), *-gara* 'sensory evidence' and *-dhan* 'spoken, or by extension written evidence'.

Eastern Pomo (Hokan; McLendon 2003:112)

(3) *Ka·lél-xa   -kʰi   ma·ʔór-al   q'á·-ne-le.*
     simply-REPEVID-3.AG   daughter.in.law-PAT   leave-INFEVID-REPEVID
     'He must have simply left his daughter-in-law there, they say'.

In Qiang (Sino-Tibetan), likewise, the marker of "inferential evidence" and the marker of "visual evidence" can be used together, as in (4) (LaPolla 2003:69; cf. Aikhenvald 2004:87-96, and Boye 2006:180-196 for discussion and additional examples).

Qiang (Sino-Tibetan; LaPolla 2003:70)

(4) *Oh,   the: ʐbə   ʐete-**k-u!***
     Oh   3SG   drum   beat-INFEVID-VISEVID
     'Oh, he WAS playing a drum!'

To repeat, the existence of language-specific systems like the one found in Ngiyambaa that are both morphosyntactically and notionally coherent is a nontrivial fact. As a nontrivial fact it ought not to be neglected. Rather, systems like that found in Ngiyambaa suggest that a notional generalization over linguistic expressions of source of information is not only intuitively sound, but also linguistically significant.

It is the implication of linguistic significance which makes claims about cross-linguistic generic categories different from claims about notional coherence, and which thus makes cross-linguistic generic categories indispensable in linguistic theory.

3. Semantic-Map Continuity as a Criterion for Categorial Status and Category Membership

In the following, I outline how semantic maps can be used to give empirical substance to the notion of cross-linguistic generic categories—how they can be employed in an empirically based criterion for categorial status and category membership.

The criterion is this:

> *Criterion of semantic-map continuity*
> *for status as and membership in a cross-linguistic category*
> The meanings comprised by a cross-linguistic category together make up a
> continuous region of a semantic map.

That is to say, any claim about cross-linguistic categorial status would be rejected if in a semantic map at least one of the hypothesized members were not immediately adjacent – i.e. linked by what Haspelmath (2003) calls "connecting lines"—to any other member. Consider, for instance, Haspelmath's semantic map of indefinite pronouns in Figure 1 (Haspelmath 1997:119).
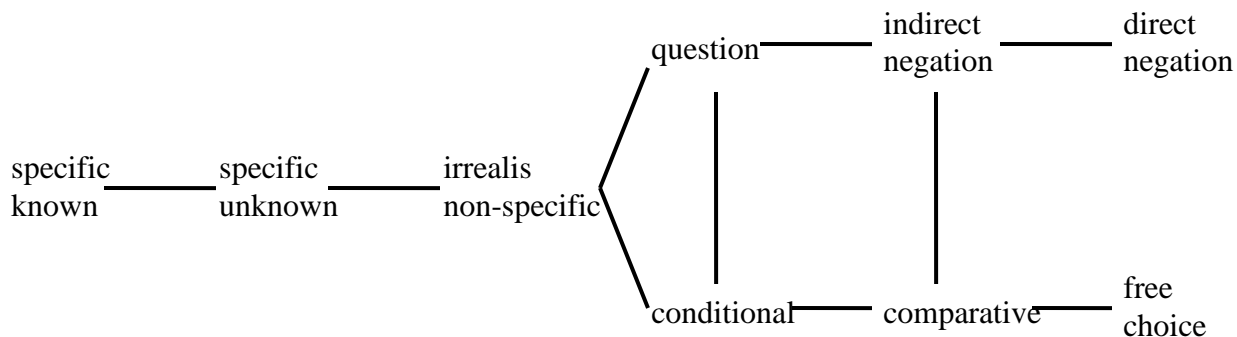
Figure 1**:** Haspelmath's (1997:64) semantic map of indefinite pronouns

If I were to claim the significance of identifying a cross-linguistic generic (sub)category of indefinite pronouns which comprises the meaning "specific known" and the meaning "irrealis non-specific", but not the meaning "specific unknown", my claim could be shot down immediately by checking it with the criterion above relative to Haspelmath's map: "specific known" and "irrealis non-specific" do not together make up a continuous region of the map. If, on the other hand, I took the category to comprise also the meaning "specific unknown", my claim would meet the criterion in so far as the three meanings "specific known", "specific unknown", and "irrealis non-specific" together make up a continuous region of Haspelmath's map.

The reason why semantic-map continuity may serve as a criterion for categorial status and category membership is that, just like the notionally coherent morphosyntactic systems discussed in Section 2, it is a linguistic empirical finding that does not follow, in any trivial way at least, from notional coherence. For instance, the fact that in Haspelmath's semantic map of indefinite pronouns the meanings "specific known", "specific unknown", and "irrealis non-specific" together make up a continuous region does not follow from the (apparent) appropriateness of generalizing over the three meanings in terms of the notion of (non-)specificity. The structure of a semantic map is an empirical result obtained by first identifying and generalizing over a number of comparable meanings across languages and then studying which meanings are directly related to each other in terms of synchronic polyfunctionality or diachronic change. A cross-linguistic notional generalization is the result of a different and far more subjective operation. Like the operation of constructing a semantic map, this operation involves an initial round of identifying and generalizing over a number of comparable meanings across languages. Subsequently, however, it submits the output of the first, already quite subjective, round to a second subjective round of generalizing over the distinct meaning generalizations. There is no reason to think that the result of constructing a notional generalization can be used to predict the result of constructing a semantic map.

The criterion of semantic-map continuity thus provides us with a guarantee of empirical linguistic significance which purely notional generalizations over linguistic expressions cannot provide themselves. It provides a basis for reinforcing claims about notional coherence into claims about *linguistically significant* notional coherence—that is, into claims about cross-linguistic generic categories.

To my knowledge, semantic maps have been employed in such reinforcing of claims only once and only implicitly. In their paper on "Modality's semantic map", van der Auwera and Plungian (1998) define modality in terms of the notions of possibility and necessity (van der Auwera and Plungian 1998:80). They argue that meanings that can be described in terms of the former notion and meanings that can be described in terms of the latter notion are adjacent in a semantic map: diachronically, they argue, expressions of necessity meaning may change into expressions of possibility meaning, and vice versa, and synchronically, expressions are found that are polyfunctional with respect to the two meanings (van der Auwera and Plungian 1998:97-104). In effect, then, they argue that modality defined in terms of the notions of possibility and necessity meets the criterion of semantic-map continuity for being identified as a cross-linguistic generic category.

4. A Semantic Map of Epistemic Expressions

I shall now return to Aikhenvald's claim about evidentiality as a category in its own right. In Section 5 I show how the criterion of semantic-map continuity and a relevant semantic map can be used to evaluate this claim in general and claims about the relation between evidentiality and epistemic modality in particular. To this end, below I present a relevant semantic map: a map of epistemic expressions.

Epistemic expressions are taken to comprise linguistic items and constructions (grammatical as well as lexical) that express either source of information or degree of (un)certainty, or both. Thus, they comprise the English expressions emphasized in (5)-(10).

English expressions of source of information

(5)   *The former president of France **seems** to have been bald.*

(6)   *__I hear__ that the former president of France was bald.*

(7)   *The former president of France was **evidently** bald.*

English expressions of degree of certainty

(8)   *The former president of France **may** have been bald.*

(9)   *It is **probable** that the former president of France was bald.*

(10)   *The former president of France was **certainly** bald.*

The semantic map of epistemic expressions below is based on data from 52 languages representing 35 phyla (according to the classification in Gordon 2005).[5] However, the map is

---

[5]The languages are (with the classifications in Gordon 2005 added in parentheses): Acoma (Keres), Basque (Basque), Bella Coola (Salishan), Berbice Dutch Creole (Creole), Carib (Carib), Cairene Egyptian Colloquial Arabic (Afro-Asiatic), Danish (Indo-European), English (Indo-European), Finnish (Uralic), French (Indo-European), German (Indo-European), Gooniyandi (Australian), Hanis (Penutian), Hdi (Afro-Asiatic), Hidatsa (Siouan), Hixkaryana (Carib), Hopi (Uto-Aztecan), Imbabura Quechua (Quechuan), Jacaltec (Mayan), Jarawara (Arauan), Japanese (Japanese), Kannada (Dravidian), Kashaya (Hokan), Khasi (Austro-Asiatic), Kobon (Trans-New Guinea), Kolyma Yukaghir (Yukaghir), Korean (Isolate), Koyra Chiini (Nilo-Saharan), Ladakhi (Sino-Tibetan), Lango (Nilo-

also compatible with data from a great number of languages discussed in Givón (1982), Akatsuka (1985), Bybee et al. (1994), and Aikhenvald (2004). The map may be represented as in Figure 2.
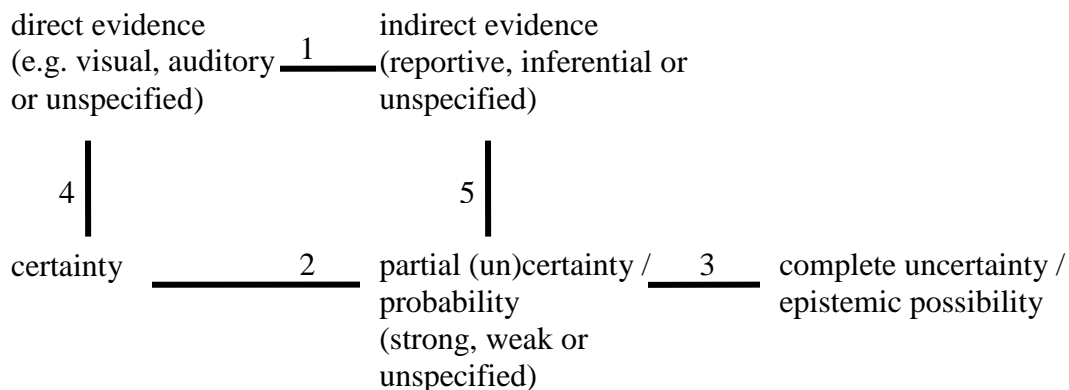

Figure 2: A semantic map of epistemic expressions

In Figure 2, as in Figure 1, the numbered lines represent what Haspelmath (2003) refers to as "connecting lines". The labels between the lines characterize epistemic meanings, or meaning regions. The map is a large-scale map. As indicated in the parentheses added to some of the labels, more fine-graded distinctions can be drawn for at least three of the meaning regions distinguished (cf. Anderson 1986 for a detailed semantic map of evidential expressions). However, this, as well as the ultimate empirical tenability of the map, is irrelevant for the present illustrative purpose.

In the rest of this section, focusing on the evidential meaning regions (direct evidence and indirect evidence), I will exemplify the linguistic basis for the distinctions (Section 4.1) and the connecting lines (Section 4.2) found in the map.

4.1 Distinctions made in the semantic map of epistemic expressions

All distinctions made in the semantic map of epistemic expressions are linguistically significant. Below, I will exemplify the linguistic basis for drawing a distinction between direct evidence and indirect evidence.

The distinction is taken to correspond to Willett's (1988:57) distinction between "direct evidence" and "indirect evidence", and roughly to Aikhenvald's (2004:65) distinction between "direct" or "firsthand" evidence on the one hand and "non-firsthand" evidence on the other. It is based on the finding that cross-linguistically, linguistic expressions are often found that indicate

Saharan), Lega (Niger-Congo), Limbu (Sino-Tibetan), Mangarayi (Australian), Mapuche (Araucanian), Montagnais (Algic), Ngiyambaa (Australian), Nunggubuyu (Australian), Qiang (Sino-Tibetan), Rapanui (Austronesian), Shipibo-Konibo (Panoan), Sherpa (Sino-Tibetan), Slave (Na-Dene), Southern Nambiquara (Nambiquaran), St'át'imcets (Salishan), Supyire (Niger-Congo), Tamil (Dravidian), Tidore (West Papuan), Tukang Besi (Austronesian), Turkish (Altaic), Wari' (Chapacura-Wanham), Western Apache (Na-Dene), West Greenlandic (Eskimo-Aleut).

either 1) (possibly a subtype of) direct evidence, but not (any subtype of) indirect evidence, or 2) (possibly a subtype of) indirect evidence, but not (any subtype of) direct evidence.[6]

An example of the first type of expression is the Ngiyambaa clitic *-gara* 'sensory evidence', illustrated in (1) above. Other examples are found in Imbabura Quechua (Quechuan), where both of the clitics *-má* 'emphatic first-hand information' and *-mi* 'first-hand information' indicate direct evidence (Cole 1982:163-166), and in Qiang (Sino-Tibetan), where direct evidence is indicated by the verbal suffix *-u/-wu* 'visual' (LaPolla 2003). In still other languages direct evidence is marked by the absence of other epistemic markers. In Hixkaryana (Carib), for instance, "zero marking", outside constructions involving the "nonpast uncertainty suffix", "marks 'eyewitness' in contrast to 'hearsay' " (Derbyshire 1979:143). In Turkish (Altaic), direct evidence is marked by "Ø" outside past-tense contexts (e.g. Kornfilt 1997:377-378). And "any declarative utterance in Bella Coola [Salishan] implies that the speaker has witnessed what he reports" (Davis and Saunders 1975:15; cf. Aikhenvald 2004:70-79 for a general discussion of "zero realization" in evidential systems).

An example of the second type of expression is the Ngiyambaa clitic *-Dhan* 'spoken, or by extension written evidence', illustrated in (2) above. Other examples are the Hixkaryana particles *ti* 'hearsay' and *mi* 'deduction' (Derbyshire 1979:143-145), the Basque particles *omen* 'citationnel' and *bide* 'apparentiel' (Jendraschek 2003:29, Oyharçabal 2003:316), and the West Greenlandic (Eskimo-Aleut) suffix *-sima* 'apparently' (Fortescue 2003:292-294). These clitics, particles, and suffixes all indicate a subtype of indirect evidence—inferential or reportive evidence. Expressions such as the Mangarayi (Australian) "past irrealis" construction (Merlan 1982:136-151) and the Montagnais (Algic) suffix *-shapan* seem to indicate indirect evidence in general, that is, to cover both the meanings of inferential and reportive evidence (cf. James et al. 2001:238 on the use of *-shapan* in the Montaignais of Betsiamites and Lower North Shore).

4.2 Connecting lines found in the semantic map of epistemic expressions

Every connecting line found in the semantic map of epistemic expressions has been established on the grounds that a number of language-specific expressions are polyfunctional with respect to, or move diachronically between, the meanings connected by the line.[7] Below, I exemplify the linguistic basis for identifying the connecting line between direct evidence and indirect evidence (connecting line 1 in Figure 2).

For this connecting line I have found no unequivocal diachronic evidence. That is, I have found no expressions for which diachronic movement from the meaning of direct evidence to the meaning of indirect, or vice versa, can be claimed unequivocally. However, I have found plenty

---

[6]This practice entails that if all attested expressions with the meaning of direct evidence covered also the meaning of indirect evidence, and all attested expressions with the meaning of indirect evidence covered also the meaning of direct evidence, the two meanings could not be distinguished in the semantic map.

[7]In order to reduce the risk that a connecting line between two distinguishable meanings is established on the basis of misinterpreted or erroneous data, it should in practice be based on several attested cases of polyfunctionality or diachronic movement between the meanings. In principle, however, a connecting line could be established between two meanings even if only one case of polyfunctionality or diachronic movement between the meanings were attested. Connecting lines are normally taken to represent functional or conceptual relations between the meanings they connect, and the existence of functional or conceptual relations must be independent of the cross-linguistic frequency with which they are actually exploited in the development of polyfunctional linguistic expressions. (Even the strength or closeness of functional or conceptual relations may be independent of the cross-linguistic frequency with which the relations are exploited–at least, any natural link has as yet to be demonstrated).

of expressions that are polyfunctional with respect to the two meanings, and it may be speculated that at least some of these expressions are the result of diachronic extension of an originally more narrow evidential meaning.

Examples of polyfunctional expressions include the Sherpa (Sino-Tibetan) verbal inflection *+nok*, which is used to indicate direct evidence in the present tense but indirect evidence in the past (Woodbury 1986:190-196), and the Kashaya (Hokan) suffix *-qă*, which may be used to mark "an inference based on circumstances or evidence found apart, in space and time, from the actual state or event", but may also be used as "a default category for [direct] evidence through senses other than those that have a specific suffix (Visual and Auditory)" (Oswalt 1986:38). In Lega (Niger-Congo), likewise, the particle *ampó* among other meanings covers both "direct sensory evidence, particularly from sight or hearing", as in (11), and "strong inferential [and thus indirect] evidence" (Botne 1997:517), as in (12).

Lega (Niger-Congo; Botne 1997:518)

(11) **ampó**   *ɛ́kurúrá*   *mompongɛ.*
 EVID   pound.3SG.PRS   rice.3
 'She's assuredly pounding rice (I can hear it)'.

(12) **ampó**   *Mɔkɛ́*   *ɛ́bilindɛ*   *bónɔ*   *Amísi*
 EVID   Moke   forget.3SG.RECPST   that   Amisi

 ɛ́ndilɛ   *ko*   *Pángé.*
 go.3SG.RECPST   to   Pangi
 '(It's evident, as his odd behavior indicates, (that)) Moke forgot that Amisi went to Pangi'.

A number of other examples of expressions that cover both the meanings of direct and indirect evidence are found in what Aikhenvald (2004:29-31) calls evidential "A2" systems. These systems of grammatical expressions of source of information consist of two members which differ in terms of a contrast between direct-visual evidence on the one hand and indirect evidence and direct-non-visual evidence on the other hand (cf. Table 2.1 in Aikhenvald 2004:65). Thus, the "non-visual" member cuts across the distinction between direct and indirect evidence.

5. Evidentiality and the Semantic Map of Epistemic Expressions

Armed with the, as yet unfalsified, semantic map of epistemic expressions and the criterion of semantic-map continuity, we are now ready to confront Aikhenvald's claim about evidentiality as a category in its own right in general, and claims about the relation between evidentiality and epistemic modality in particular.

5.1 Mapping evidentiality onto the semantic map of epistemic expressions

As for Aikhenvald's claim, it meets the criterion of semantic-map continuity when it is evaluated against the proposed semantic map of epistemic expressions.

When we apply the criterion of semantic-map continuity to a proposed cross-linguistic generic category, we ask whether the meaning regions covered by the proposed category together

constitute a continuum in a relevant semantic map. That is, we ask whether each of the meaning regions distinguished within the category is connected in a relevant semantic map to at least one of the other meaning regions distinguished. In the case of Aikhenvald's category of evidentiality, the answer to this question is yes. Aikhenvald defines evidentiality in terms of two features: 1) the notion of source of information and 2) grammatical status. Only the former feature is relevant here since, as conceived of by Aikhenvald (and others), it includes the meaning regions of direct evidence and indirect evidence as evidential, and excludes all other meaning regions as non-evidential. In the semantic map of epistemic expressions represented in Figure 2, these two meaning regions together make up a continuum. They are connected to each other by connecting line 1. Thus, Aikhenvald's cross-linguistic category of evidentiality may be mapped onto the semantic map of epistemic expressions, as illustrated in Figure 3.
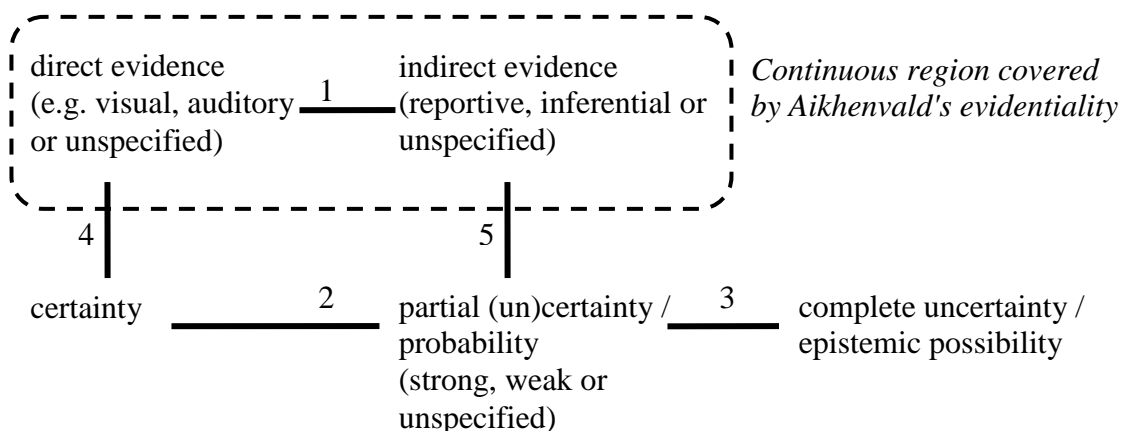

Figure 3: Aikhenvald's cross-linguistic category of evidentiality mapped onto the semantic map of epistemic expressions

Aikhenvald's evidentiality meets the criterion of semantic-map continuity then. However, this does not entail that her claim that evidentiality is a category in its own right is hereby confirmed beyond dispute. First, it might turn out that the semantic map of epistemic expressions can be falsified when confronted with more data or with a reanalysis of the data on which it is based. On a revised map the meaning of direct evidence and the meaning of indirect evidence might not be connected by a connecting line. Second, as will be demonstrated in Section 6, semantic-map continuity is not a sufficient criterion for categorial status and category membership. Meeting the criterion is not sufficient for being identified as a cross-linguistic generic category.

   Still, semantic maps and the criterion of semantic map continuity can be used to effectively reject some claims about cross-linguistic generic categories.

5.2 Evaluating claims about evidentiality by means of semantic maps

The categorial status of evidentiality has been intensively discussed in connection with attempts to describe the relation between evidentiality and epistemic modality. Several claims have been made about this relation (cf. the overviews given in Dendale and Tasmowski 2001b:241-242, and Nuyts 2006:10-12). Some works, like Aikhenvald's, take evidentiality and epistemic modality to be separate categories (e.g. de Haan 1999 and Nuyts 2001:27, 35). Others take the former to include the latter (e.g. Papafragou 2000:121, and Ifantidou 2001:5-8), or the latter to include the former (e.g. Palmer 1986:51, and Willett 1988:52). Still others take them to be separate but overlapping categories (e.g. van der Auwera and Plungian 1998:85-86, and Palmer

2001:8-9). And a couple of works take evidentiality and epistemic modality to be separate categories that are, however, both subcategories of the same superordinate category (Hengeveld 1989, and Boye 2006). The variety of claims is partly a result of the lack of empirically based criteria for identifying cross-linguistic generic categories. Categories are identified and distinguished mainly on the basis of notional coherence, and the choice between the claims is thus to some degree a terminological choice between different opinions about which notions—which ranges of meanings—the terms 'evidentiality' and 'epistemic modality' should be tied to.

Confronted with the criterion of semantic-map continuity and the semantic map of epistemic expressions, at least one of the claims must be rejected: van der Auwera and Plungian's (1998) claim that evidentiality and epistemic modality are separate but overlapping categories. Like Aikhenvald's category of evidentiality, van der Auwera and Plungian's categories of evidentiality and epistemic modality actually turn out to meet the criterion of semantic-map continuity when evaluated against the semantic map of epistemic expressions. As for the category of evidentiality, van der Auwera and Plungian share Aikhenvald's conception of it and explicitly take it to comprise the meaning regions direct and indirect evidence (as well as other meanings that are arguably covered by the two regions) (van der Auwera and Plungian 1998:85). Thus, van der Auwera and Plungian's category of evidentiality covers the same continuous region of the semantic map of epistemic expressions as does Aikhenvald's category (compare Figure 3 above with Figure 4 below). As for the category of epistemic modality, van der Auwera and Plungian define it as comprising two meaning regions: "epistemic possibility" (which they take to amount to "uncertainty") and "epistemic necessity" (which they take to amount to "[c]ertainty and a relatively high degree of probability") (van der Auwera and Plungian 1998:81). In so far as it comprises certainty, probability, and epistemic possibility, then, this category also covers a continuous region of the semantic map of epistemic expressions, as illustrated in Figure 4.
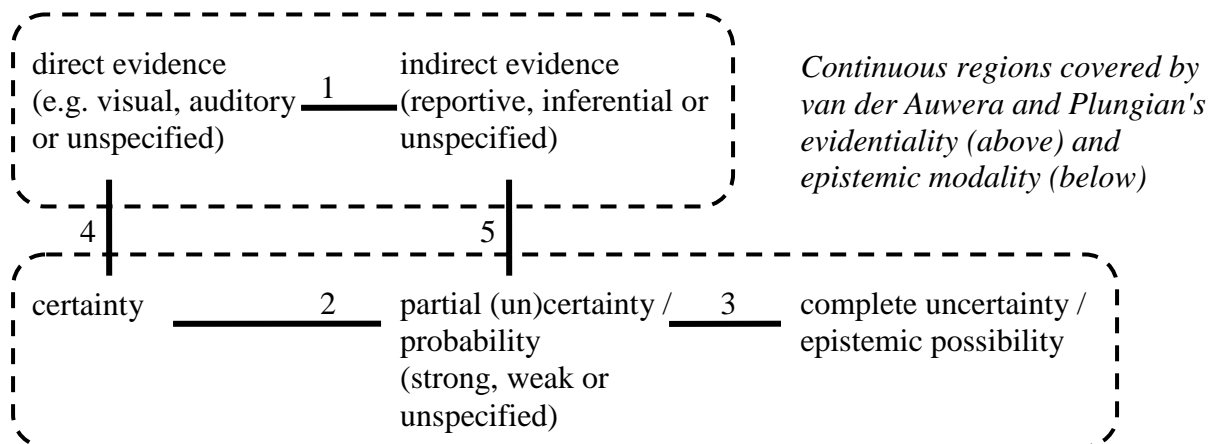

Figure 4: Van der Auwera's and Plungian's (1998) cross-linguistic categories of evidentiality and epistemic modality mapped onto the semantic map of epistemic expressions

The reason for rejecting van der Auwera and Plungian's claim is that their idea of an overlap between evidentiality and epistemic modality is incompatible with the semantic map of epistemic expressions. According to van der Auwera and Plungian, the overlap is located in the epistemic modal region of "epistemic necessity"—or "[c]ertainty and a relatively high degree of probability"—and the evidential region of inferential evidence, a subregion of indirect evidence.

These two regions, they claim, are identical—one amounts to the other (van der Auwera and Plungian 1998:85-86; cf. Palmer 2001:8-9 on "deductive" as an overlap region). However, in the semantic map of epistemic expressions the two regions are clearly distinct, and the distinction made between them is based on plenty of linguistic evidence. Expressions are found that indicate degree of certainty but not source of information, and expressions are found that indicate source of information but not degree of certainty. As an example of the latter, the Kashaya inferential-evidence suffix *-qă*, discussed in Section 4.2 above, has no implications of degree of certainty. What it implies is "merely lack of higher ranking evidence" (Oswalt 1986:38; cf. de Haan 1999 and Aikhenvald 2004:153-193 for additional evidence against the overlap view).

6. The Limits to Semantic Maps in Identifying Cross-Linguistic Generic Categories

Claims about cross-linguistic generic categories are claims about linguistically significant meaning generalizations. Semantic-map continuity is a nontrivial linguistic fact, and in the preceding section I have tried to show how semantic maps and the criterion of semantic-map continuity can be used to evaluate whether or not a given meaning generalization is in fact linguistically significant.

However, semantic maps are not an unequivocal means for identifying categories and relations between them. As already mentioned, the criterion of semantic-map continuity is not a sufficient criterion for categorial status and category membership.

To see this, consider again Aikhenvald's claim that evidentiality is a category in its own right. As demonstrated in Section 5.1, this claim meets the criterion of semantic-map continuity when evaluated against the proposed semantic map of epistemic expressions. As noted in the same section, however, this is no proof that her claim is correct in any sense. Claims that are incompatible with Aikhenvald's also meet the criterion. For instance, the following claim does:

> Aikhenvald's category of evidentiality is not a category in its own right. Rather, the meaning regions direct evidence and indirect evidence are comprised by a cross-linguistic category of epistemicity which in addition comprises the meaning regions certainty, probability, and epistemic possibility.

Arguably, the range of meanings covered by the category of epistemicity proposed in this claim is notionally coherent—the meaning regions direct evidence, indirect evidence, certainty, probability, and epistemic possibility can all be defined in terms of a notion like the philosophers' "justificatory support", for instance (cf. Boye 2006 for details). As illustrated in Figure 5, the category of epistemicity maps onto a continuous region of the semantic map of epistemic expressions.
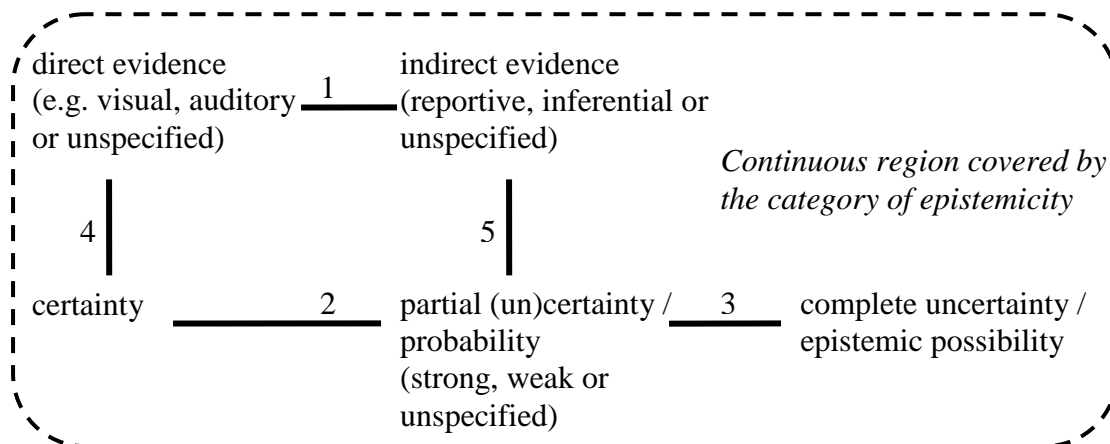
Figure 5: The cross-linguistic category of epistemicity mapped onto the semantic map of epistemic expressions

The following claim (made in Boye 2006) likewise meets the criterion of semantic-map continuity:

> Aikhenvald's category of evidentiality is not a category in its own right—at least not in the sense that it is not comprised by a superordinate category. Rather, a category of epistemicity comprises Aikhenvald's evidentiality as a subcategory side by side with a subcategory of epistemic modality (defined as to comprise the meanings of certainty, probability and epistemic possibility).

All of the three proposed categories arguably cover notionally coherent ranges of meanings, and each of the categories maps onto a continuous region of the semantic map of epistemic expressions, as illustrated in Figure 6.
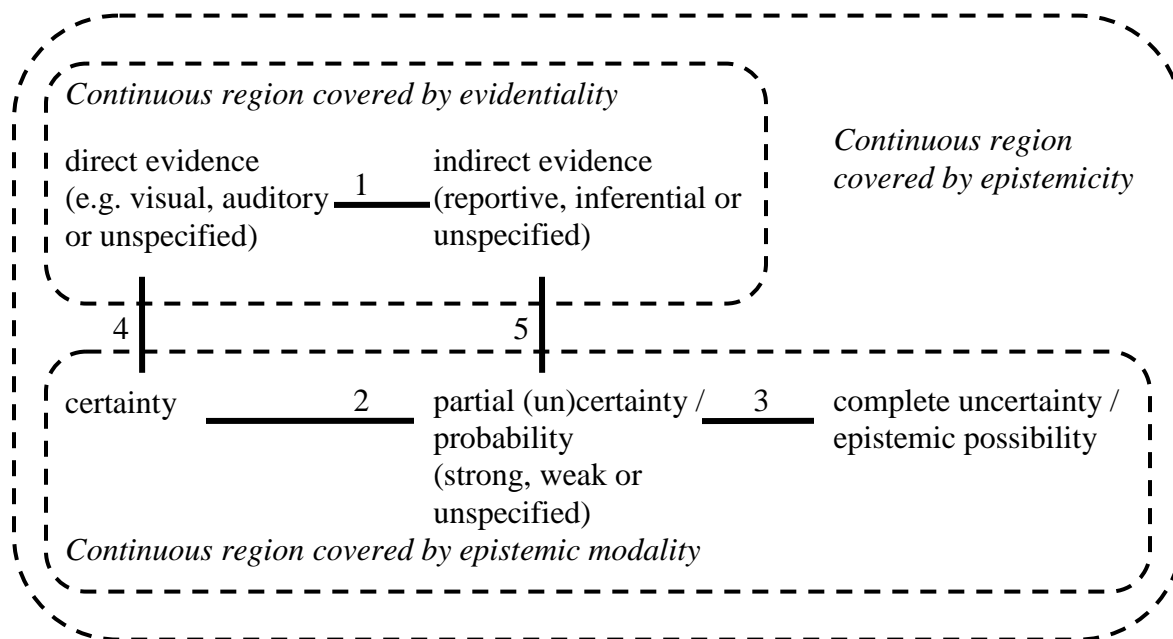

Figure 6: The cross-linguistic categories of evidentiality, epistemic modality, and epistemicity mapped onto the semantic map of epistemic expressions

What disqualifies semantic-map continuity as a sufficient criterion for categorial status and category membership is that a one-to-one correspondence between the range of meanings defined by notional coherence and the range of meanings defined by semantic-map continuity cannot be expected (cf. Section 3). First, the fact that two meaning regions are notionally similar and together make up a continuous region of a semantic map does not entail that a category covers *only* the two regions at hand. For instance, the epistemic modal meaning regions certainty and probability can both be defined in terms of the notion of degree of certainty, and together they make up a continuous region of the semantic map of epistemic expressions (cf. Figure 6), but this does not entail that epistemic modality covers only these two regions and not the region of epistemic possibility as well. Second, the fact that two meaning regions together make up a continuous region of a semantic map does not entail that they are notionally coherent. For instance, expressions of indirect evidence may develop from, among other things, perfects, resultatives, and past tenses (Aikhenvald 2004:279-281). Accordingly, any semantic map that covered the meaning of indirect evidence as well as resultative meaning, perfect meaning, and past tense meaning would have to include connecting lines between the latter meanings and the former. But clearly, the meaning of indirect evidence is notionally distinct from its three source meanings, and to my knowledge, it has never been claimed to belong to the same category as any of them. Semantic-map continuity may cut across category borders then. Connecting lines are not necessarily category internal, but may be category external.

If we want to be able to qualify essentially subjective claims about notional coherence into empirically based claims about cross-linguistic generic categories, the criterion of semantic-map continuity will not do the job alone. We need other criteria. An obvious candidate would draw on the fact discussed in Section 2 that some language-specific systems are actually notionally coherent. The criterion might be formulated something like this:

> *Morphosyntactic criterion for status as and membership in a cross-linguistic category*
> Meanings comprised by a cross-linguistic category are frequently found with all members of language-specific morphosyntactically delimited systems of expressions.

As we need other criteria anyway, it is natural to ask whether we need the criterion of semantic-map continuity in the first place. It is not a sufficient criterion, but is it a necessary criterion for categorial status and category membership? I suggest that it is. However, the answer is ultimately dependent on an exact definition of cross-linguistic generic categories. Cross-linguistic generic categories are taken here to refer to linguistically significant meaning generalizations. The question as to whether or not the semantic-map continuity criterion needs to be met in order for something to be identified as a category or as a member of a category is a question of whether or not "linguistic significance" is taken to necessarily include semantic-map continuity. Semantic-map continuity is a nontrivial fact which can be used to reinforce claims about notional coherence into claims about linguistically significant notional significance, but it is not the only such fact.

7. Conclusion[8]

As noted in Section 1, there is a remarkable lack of criteria for what counts as a cross-linguistic generic category, and categories like evidentiality, tense, aspect, modality, person, and number tend to be identified and distinguished from each other mainly on the basis of notional coherence. One consequence of this is that discussions about such categories are often hampered by lack of consensus as to what the categories cover (discussions about modality are a notorious case in point). Notional generalizations are essentially subjective, and every scholar is free to come up with his or her own definition. Another consequence is that, as enlightening and intuitively sound as they may be, claims about the functional or cognitive motivations for the categories are fundamentally hollow. As long as we have no means for assessing whether a notional generalization is linguistically significant or not, it might seem a complete waste of time to try to motivate it or to draw conclusions about communicative or cognitive phenomena on the basis of it. A third consequence, finally, is that the notion of cross-linguistic generic categories itself becomes exposed to suspicions that it is superfluous. Unless we come up with criteria for distinguishing categories as linguistically significant notional generalizations from purely notional generalizations, we cannot blame scholars like Bybee for dismissing the former and sticking with the latter.

In this paper I have argued that cross-linguistic generic categories cannot be dispensed with in linguistic theory, and that semantic maps can be used to give substance to claims about them. Focusing on the category of evidentiality and its relation to the category of epistemic modality, I have tried to show that semantic maps may provide us with an empirically based criterion for categorial status and category membership as well as with a falsifiable basis for identifying relations between different categories. However, I have also argued that semantic maps are not an unequivocal means for identifying categories and relations between them, and that the criterion of semantic-map continuity thus needs to be supplemented with additional criteria.

Cross-linguistic generic categories should not be conceived of as pre-established in the sense of Haspelmath (2007). Rather, they are the result of what Lazard (2001:365) calls "the only way out of the difficulty of linguistic typology (especially grammatical typology)". As outlined by Lazard, this way starts with the formation of hypotheses and intuition-based notional generalizations, and it goes on with empirical evaluation of these hypotheses and generalizations until they are ultimately validated. Not only can semantic maps help us pass the starting line. They can also take us a good deal further.

Abbreviations

Abbreviations used in this paper and not standardized in the Leipzig Glossing Rules: AG agent, EVID evidence, INFEVID inferential evidence, PAT patient, RECPST recent past, REPEVID reportative evidence, SENSEVID sensory evidence, VISEVID visual evidence.

References

Aikhenvald, Alexandra Y. 2004. Evidentiality. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. and R.M.W. Dixon (eds.). 2003. Studies in Evidentiality. Amsterdam: Benjamins.

Akatsuka, Noriko. 1985. Conditionals and the epistemic scale. Language 61/3.625-639.

Anderson, Lloyd B. 1986. Evidentials, paths of change, and mental maps: Typologically regular asymmetries. Chafe and Nichols 1986, 273-312.

Bache, Carl. 1997. The study of aspect, tense and action: Towards a theory of the semantics of grammatical categories. Frankfurt am Main: Lang.

Botne, Robert. 1997. Evidentiality and epistemic modality in Lega. Studies in Language 21/3.509-532.

Boye, Kasper. 2006. Epistemic meaning: A cross-linguistic study. Ph.D. thesis. Copenhagen: University of Copenhagen.

Boye, Kasper and Peter Harder. 2009. Evidentiality: Linguistic categories and grammaticalization. Functions of Language 16/1.9-43.

Bybee, Joan L. 1985. On the nature of grammatical categories. Proceedings of the second eastern states conference on linguistics, ed. by Soonja Choi, Dan Devitt, Wynn Janis, Terry McCoy and Zheng-sheng Zhang, 17-34. Buffalo, NY: The State University of New York.

Bybee, Joan L., Revere D. Perkins and William Pagliuca. 1994. The evolution of grammar: Tense, aspect and modality in the languages of the world. Chicago: University of Chicago Press.

Chafe, Wallace and Johanna Nichols (eds.). 1986. Evidentiality: The linguistic coding of epistemology. Norwood: Ablex.

Cole, Peter. 1982. Imbabura Quechua. Amsterdam: North-Holland.

Croft, William. 2003. Typology and universals, 2nd edition. Cambridge: Cambridge University Press.

Davis, Philip and Ross Saunders. 1975. Bella Coola deictic usage. Rice University Studies 61.13-35.

de Haan, Ferdinand. 1999. Evidentiality and Epistemic Modality: Setting boundaries. Southwest Journal of Linguistics 18.83-101.

Dendale, Patrick. and Liliane Tasmowski (eds.). 2001a. On Evidentiality. Journal of Pragmatics 33/3 (special issue). Amsterdam: Elsevier.

-----. 2001b. Introduction: Evidentiality and related notions. Dendale and Tasmowski 2001a, 339-348.

DeLancey, Scott. 2001. The mirative and evidentiality. Dendale and Tasmowski 2001a, 369-382.

Derbyshire, Desmond. C. 1979. Hixkaryana. Amsterdam: North-Holland.

Donaldson, Tamsin. 1980. Ngiyambaa. Cambridge: Cambridge University Press.

Fortescue, Michael. 2003. Evidentiality in West Greenlandic: A case of scattered coding. Aikhenvald and Dixon 2003, 291-306.

Givón, Talmy. 1982. Evidentiality and epistemic space. Studies in Language 6/1.23-49.

Gordon, Raymond G., Jr. (ed.). 2005. Ethnologue: Languages of the world, 15th edition. Dallas, TX: SIL International. Online version: http://www.ethnologue.com/.

Guentchéva, Zlatka (ed.) 1996. L'énonciation médiatisée. Paris: Peeters.

Guentchéva, Zlatka and Jon Landaburu. (eds.) 2007a. L'énonciation médiatisée, vol. 2: Le traitement épistémologique de l'information: Illustrations amérindiennes et caucasiennes. Louvain: Peeters.

-----. 2007b. Introduction. Guentchéva and Landaburu 2007a, 1-19.

Haspelmath, Martin. 1997. Indefinite pronouns. Oxford: Clarendon Press.

-----. 2003. The geometry of grammatical meaning: Semantic maps and cross-linguistic comparison. The new psychology of language: Cognitive and functional approaches to language structure, ed. by Michael Tomasello, vol. 2, 211-242. London: Erlbaum.

-----. 2007. Pre-established categories don't exist: Consequences for language description and typology. Linguistic Typology 11/1.119-132.

Heine, Bernd, Ulrike Claudi and Friederike Hünemeyer. 1991. Grammaticalization: A conceptual framework. Chicago: University of Chicago Press.

Hengeveld, Kees. 1989. Layers and operators in Functional Grammar. Journal of Linguistics 25.127-157.

Ifantidou, Elly. 2001. Evidentials and relevance. Amsterdam: Benjamins.

James, Deborah, Sandra Clarke and Marguerite MacKenzie. 2001. The encoding of information source in Algonquian: Evidentials in Cree/Montagnais/Naskapi. International Journal of American Linguistics 67/3.229-263.

Jendraschek, Gerd. 2003. La modalité épistémique en basque. Munich: Lincom.

Kornfilt, Jaklin. 1997. Turkish. London: Routledge.

LaPolla, Randy J. 2003. Evidentiality in Qiang. Aikhenvald and Dixon 2003, 63-78.

Lazard, Gilbert. 1999. Mirativity, evidentiality, mediativity, or other? Linguistic Typology 3/1.91-110.

-----. 2001. On the grammaticalization of evidentiality. Dendale and Tasmowski 2001a, 359-367.

McLendon, Sally. 2003. Evidentials in Eastern Pomo with a comparative survey of the category in other Pomoan languages. Aikhenvald and Dixon 2003, 101-130.

Merlan, Francesca. 1982. Mangarayi. Amsterdam: North-Holland.

Nuyts, Jan. 2001. Epistemic modality, language, and conceptualization. Amsterdam: Benjamins.

-----. 2006. Modality: Overview and linguistic issues. The expression of Modality, ed. by William Frawley, 1-26. Berlin: Mouton de Gruyter.

Oswalt, Robert L. 1986. The evidential system of Kashaya. Chafe and Nichols 1986, 29-45.

Oyharçabal, Bernard. 2003. Tense, aspect and mood. A grammar of Basque, ed. by José. I. Hualde and Jon Ortiz de Urbina, 249-284. Berlin: Mouton de Gruyter.

Palmer, Frank R. 1986. Mood and modality. Cambridge: Cambridge University Press.

-----. 2001. Mood and modality, 2nd edition. Cambridge: Cambridge University Press.

Papafragou, Anna. 2000. Modality: Issues in the semantics-pragmatics interface. Oxford: Elsevier.

Slobin, Dan. I. 1997. The origins of grammaticizable notions: Beyond the individual mind. The crosslinguistic study of language acquisition, ed. by Dan I. Slobin, vol. 5: Expanding the contexts, 265-323. London: Erlbaum.

van der Auwera, Johan and Vladimir A. Plungian. 1998. Modality's semantic map. Linguistic Typology 2/1.79-124.

Whorf, Benjamin L. 1956a [1938]. Some verbal categories of Hopi. Language, thought, and reality: Selected writings of Benjamin Lee Whorf, ed. by John B. Carroll, 112-124. Chicago: MIT Press.

Willett, Thomas. 1988. A cross-linguistic survey of the grammaticization of evidentiality. Studies in Language 12/1.51-97.

Woodbury, Anthony C. 1986. Interactions of tense and evidentiality: A study of Sherpa and English. Chafe and Nichols 1986, 188-202.

Author's contact information:
Kasper Boye
Department of Scandinavian Studies and Linguistics
University of Copenhagen
Njalsgade 120
2300 Copenhagen S
Denmark
kabo@hum.ku.dk