

# TYPOLOGICAL APPROACHES TO MODALITY

*Ferdinand de Haan*

*May 13, 2004*

To appear in:

William Frawley, ed. (2005) *Modality*. Berlin: Mouton de Gruyter.

## **TABLE OF CONTENTS**

- 1. introduction**
- 2. terminology**
- 3. expressions of modality**
  - a. modal auxiliary verbs**
  - b. mood**
  - c. modal affixes**
  - d. lexical means**
    - i. modal adverbs**
    - ii. modal tags**
    - iii. modal particles**
  - e. modal case**
- 4. realis and irrealis**
- 5. semantic map of modality**
- 6. modality and tense**
  - a. modality and future**
  - b. modality and past tense**
- 7. modality and negation**
- 8. evidentiality**
  - a. mirativity**

## I. INTRODUCTION

This chapter surveys the typological literature on modality. Studies within the typological tradition place an emphasis on explaining language structure through analyzing their function (hence typology is part of the functionalist approach to linguistics).<sup>1</sup> Typological analyses involve cross-linguistic comparisons and generalizations, for which explanations are sought. Explanations for cross-linguistic generalizations very often are extra-linguistic in nature. That is, explanations can be drawn from language use, cognition, and from sociological factors. In addition, explanations may be drawn from the history of the language (the diachronic dimension). Diachronic explanations play an important role in typological approaches to modality (known as *grammaticalization* studies). However, since they are treated in detail elsewhere (Traugott, this volume), the area of grammaticalization will receive little attention here.

Most of the typological literature on modality is concerned with its semantic aspects. This is in accord with the general philosophy that the linguistic form can best be explained from the language use. Thus, an analysis of the modal *must* in sentence (1) below could involve taking into account the context in which it occurs, the mode of language (for instance, spoken or written), a comparison with cognate verbs in related languages, a comparison with *must* in earlier stages of English, and possibly even the sociological data of the speaker.

- (1) He told MSPs that a thorough investigation of the cause was needed and lessons

**must** be learned for the future.

(BBC online, May 13, 2004)

This does not mean that such detail is always required or even rigorously followed where it would be appropriate. It is merely meant to illustrate the range of phenomena that can be taken into account when one takes a typological approach to linguistics.

Within typology, the area of modality has not received the same level of attention that categories like ergativity or causativity have enjoyed. Although there are some good typological studies of modality that are older than 20 years, Palmer (1986) is the first book-length work that takes a typological outlook. Palmer (1986) and (2001) are standard reference works on modality in a typological perspective. Even though (2001) is nominally the second edition of (1986) there are enough differences in theoretical outlook between the two books to consider them separate works. There are works on modality in functional-typological frameworks, such as Functional Grammar (e.g., Dik 1997) and Role and Reference Grammar (e.g., Foley and Van Valin 1985) but for reasons of space these will not be discussed in any great detail. Other typological surveys are Chung and Timberlake (1985) and Givón (1984). A very good recent work on epistemic modality in a functional-cognitive framework is Nuyts (2001). There are of course numerous studies and monographs of modality in a single language or language family, and they will be mentioned as is warranted in the general discussion.

Because of the relative youth of typological studies on modality, there is as yet no consensus on the proper terminology for modal meanings. For this reason, a section of the paper is devoted to various proposals for developing a consistent and cross-linguistically valid set of terminology.

The next section is devoted to a survey of the ways in which modality can be expressed. As is the norm for typological studies, an element is considered modal if it has modal meanings (like obligation, permission and prohibition). This means that there are quite a number of formal modal elements besides the familiar (from English) modal verb.

The next couple of sections deal with various topics that are currently in fashion, and that also have relevance for other theoretical frameworks. They are a discussion on the status of the irrealis, and interactions of modality and tense and modality and negation.

Another recent development within typology is that of semantic maps, in which the semantic inventory of a given feature in a given language is mapped on an abstract representation of that feature. This has been done for several features (perfect, evidentiality, indefiniteness) and also for modality.

The chapter closes with some remarks on evidentiality, a category which is very often considered to be modal. It is treated as such in Palmer (1986), for instance.

## **II. TERMINOLOGY**

Part of the typological literature on modality is concerned with the proper terminology of modality. Over the last couple of decades several different sets of terminology have been proposed in the literature, often with subtle differences in meaning. This section surveys the most important terminological debates in the typological literature. It is not meant to be an exhaustive listing of all terms used in modality, of which there are many.

Following the logicians (following on from von Wright 1951), the original division is between *epistemic* and *deontic* modality. Epistemic modality, as in *John must have been at home*, refers to the degree of certainty the speaker has that what s/he is saying is true. Deontic modality, as in *John must go to school*, deals with the degree of force exerted on the subject of the sentence to perform an action. This force can come from the speaker but also from an unspecified third source. This division is used in such works as Lyons (1977), Palmer (1979 [1990], 1986 [2001]), Frawley (1992), De Haan (1997), Van der Auwera and Plungian (1998, but see below), Traugott and Dasher (2002), as well as in various grammatical studies of single languages. Generally, this division requires a separate modality, often referred to as *dynamic* modality, to encode ability (and, depending on the author, often volition as well). An example is the sentence *John can swim*, in which the modal verb *can* denotes the subject's ability to swim.<sup>2</sup>

While the status of epistemic modality is not in doubt, scholars have proposed new terminology for the deontic side of the spectrum. The most influential proposal is that of *root* modality. This term makes reference to root (or main) clauses and it is somewhat of a misnomer, since it does not only occur in root clauses. The term has been around since the late 1960s (no doubt owing to the rise of generative grammar and its emphasis on syntax), but the first influential study to employ the notion of root modality appears to be Coates' 1983 corpus study on the English modals. She rejects the term *deontic* on the grounds that this term primarily refers to the logical notions of obligation and permission, while modals such as *must* and *may* have other interpretations as well (1983:20-1). The term *Root* covers both deontic and dynamic modality, as defined above. In Coates' model of modality, modal meanings are gradual, without "arbitrary cut-off

points”. Her view is that modals have core and peripheral meanings, and the terms deontic and dynamic only refer to the core meanings, hence her use of a neutral term, *root* modality. (see Palmer 1986:103-4 for arguments against the term *root*). There is a clear difference between the terms *deontic* and *root* modality. Unfortunately, this difference is somewhat subtle and in many studies the two terms are used interchangeably. However, in coining and using the term *root* modality, linguists can show that there are aspects of modality that lie outside the traditional domain of modality in logic. There are elements of modality in natural language that lend themselves poorly to descriptions in terms of modal logic. The use of a term such as *root* modality highlights this aspect of modality. It is used in such studies as Sweetser (1990).

The next set of terminology is found in such studies as Bybee (1985:166ff) and Bybee, Perkins and Pagliuca (1994). Their work is based on the premise that, in order to understand the range of modal meanings in a language, one must understand the diachronic developments of modal elements. They propose therefore the following division of modality (1994:177ff):

- Epistemic
- Agent-oriented
- Speaker-oriented
- Subordinating

The use of the term *epistemic* is relatively straightforward, since they include possibility and probability among the epistemic meanings. Another epistemic category is *inferred*

*certainty*, which is used when the speaker has good reasons to believe that the statement is true (their example is *There must be some way to get from New York to San Francisco for less than \$600*). *Subordinating* moods refer to the use of modality in subordinate clauses, such as *concessive* (*although ...*) and *purposive* (*so that ...*) clauses. One exponent of subordinating moods is the *subjunctive*, see below.

The term *agent-oriented* modality refers to those cases in which the agent of a clause is influenced in some way in performing the action described in the clause: “Agent-oriented modality reports the existence of internal and external conditions on the agent with respect to the completion of the action expressed in the main predicate.” (1994:177). Some types of agent-oriented modality are *obligation* (there exist external factors that compel the agent to complete the action, as in *All students must obtain written permission from the Dean ...*), *necessity* (there exist physical conditions, as in *I need to hear a good loud alarm in the morning to wake up*), *ability* (there exist agent-internal enabling conditions, see above), and *desire* (there are internal volitional conditions). A very important type of agent-oriented modality is *root possibility*, which is related to ability, but also takes external factors into account. An example of root possibility is *I actually **couldn't** finish it because the chap whose shoulder I was reading the book over got out at Leicester Square*. (Bybee et al. 1994:178, from Coates 1983:114). The use of the modal *couldn't* does not denote an internal inability, but rather an inability caused by the external factor of someone else's leaving.

*Speaker-oriented* modality refers to those cases in which the speaker is the “enabling condition,” i.e., those cases in which the speaker gives someone an order or gives someone permission. This type of modality includes *directives* (a term from Lyons

1977), *imperatives* (the command mood, see below), *prohibitions* (negative imperatives), *optatives* (see below), *admonitions* (warnings), and *permissions*.

In this framework, agent-oriented and speaker-oriented modality roughly divide the area of root modality, or deontic/dynamic modality. The deciding factor in Bybee et al.'s framework is: who or what is the enabling factor? If it is the speaker, then we are dealing with speaker-oriented modality, otherwise it is an instance of agent-oriented modality.

Some scholars have sought to refine this framework while keeping the basic structure intact. Hengeveld (...), cited from Van der Auwera and Plungian (1998), uses the term *participant-oriented* modality instead of *agent-oriented* modality. This is done to include those cases in which the subject of the sentence is not actually an agent (as in *John needs to be left in peace today*, in which the subject, John, has the thematic role of patient).

Van der Auwera and Plungian (1998:80-6) make a distinction between *participant-internal* and *participant-external* modality. Participant-internal modality is more or less identical with dynamic modality as it deals with ability and need (as in *John needs a book*). Participant-external modality is again divided into deontic and non-deontic participant-external modality. In this view, deontic modality is a subtype of participant-external modality. It encompasses permission and obligation (either from the speaker or another source). Non-deontic modality deals with possibility and necessity. It refers to circumstances wholly external to the situation. An example is the sentence *To get to the station, you can take Bus 66* (1998:80). There is then no need for a special category of subject-oriented modality because it is either subsumed under deontic

modality or, in the case of imperatives, optatives etc., because it is not considered to be part of modality. In addition, volition is in their view not part of modality either. As the authors themselves admit (1998:84), the term *participant-oriented* is perhaps too vague, given that a sentence usually has more than one participant. They consider the term *subject-oriented* as an alternative but reject it on the grounds that it is probably not valid cross-linguistically. They cite the case of the Experiencer in Kannada (Bhat 1991), a Dravidian language, which is not a subject but which can be acted upon by a deontic force.

Another distinction frequently made in typological studies is that between real and unreal events, or a realis – irrealis distinction. It has been claimed that there are languages that encode modality this way rather than in a deontic – epistemic way. This is an important observation and a separate section is devoted to this distinction.

This concludes the section on terminology. While there is as yet little agreement about the correct terminology, there seems to be at least a consensus to use terminology which reflects linguistic ways of analysis. This is a welcome development.

### **III. EXPRESSIONS OF MODALITY**

In the typological tradition, categories tend to be defined semantically. This means that a morpheme is classified as modal if it has a modal meaning (epistemic, deontic, etc.). Modal meanings are expressed by various morphological, syntactic, and lexical categories. This section is a survey of the ways in which modality is formally expressed across languages.

## MODAL AUXILIARY VERBS

The best-known means of expression for speakers of English is doubtless the modal verb. In English, modal verbs are auxiliary verbs as they are used in addition to the main verb of the sentence. Examples are:

- (2) a. John must go to school.  
b. John must be at school.

The modal verb *must* is used to denote necessity on the subject (deontic modality) and strong conviction of the speaker (epistemic modality). These are examples of strong modality (Palmer 1986). English also has expressions for weak modality:

- (3) a. John may go to school.  
b. John may be at school.

The modal verb *may* denotes permission (deontic modality) or possibility (epistemic modality).<sup>3</sup>

Thus, modal verbs such as *must* and *may* are ambiguous between epistemic and deontic modality. This is a very frequent cross-linguistic phenomenon.<sup>4</sup> Ambiguity of modal verbs is found, among others, in the Germanic, Slavic, and Romance language families in Europe, as well as in certain languages outside Europe. An example of an

African language with modal verbs is Yoruba, in which the verb *lè* is used for weak epistemic and deontic modality, and the verbs *gbọdọ* and *ní láti* for strong epistemic and deontic modality. An example is:

- (4) Ó gbọdọ wà nílẹ.  
 He must be at.home  
 ‘He must be in.’ (Adewole 1990:80)

## MOOD

The category of *mood* is here defined as a morphological verbal category which expresses the modal value of the sentence. Mood is the grammaticalized expression of modality, just as, say, tense is the grammaticalized expression of time. Mood is therefore an obligatory category in those languages that have it.

The most common moods cross-linguistically appear to be the *indicative* and the *subjunctive*. These moods are found in, among others, the classical languages (e.g., Sanskrit, Classical Greek and Latin) but also in the modern descendants of Latin, in the Slavic languages, and in certain Germanic languages (notably Icelandic and German). It has also been described for Bantu languages like Swahili, Native American languages (for instance, languages from the Algonquian language family) and Australian languages like Gooniyandi and Mangarayi.

As a first approximation we can say that the indicative is used to describe real, factual events, while the subjunctive is used for unreal, hypothetical events (but see



‘I want him to rent a house for me.’

The Indicative in Latin is used for facts (again from Hale and Buck 1903):

- (7) a. quid tac-es?  
 why be.silent-2SG.IND.PRES  
 ‘Why are you silent?’ (p. 293)
- b. quoad potu-it, resist-it  
 as.long.as can-3SG.IND.PERF resist-3SG.IND.PERF  
 ‘As long as he could, he resisted.’ (p. 294)

Given the data from Latin, it is tempting to equate the indicative – subjunctive distinction with the realis – irrealis distinction alluded to in the section on terminology. The indicative is used for real events that have taken or are taking place and the subjunctive is used for those events that may, will, or should take place. This is a frequently held view but there are good reasons for keeping the two sets of terminology distinct. As this problem has attracted some attention in the typological literature a separate section will be devoted to it.

One objection to equating irrealis and subjunctive is the fact that there are languages which distinguish other moods besides indicative and subjunctive. One such mood is the *optative*, which is used to express the semantic categories of wishing and hoping. A separate optative mood has been reconstructed for Proto-Indo-European, and is found in such languages like Classical Greek and Sanskrit. In Latin, the old Optative and



The imperative differs somewhat from the other moods discussed in that it is a restricted mood in the sense that it is limited to a certain speech situation, namely when the speaker addresses the hearer directly and gives a direct command. Therefore, it is performative in nature because by uttering the imperative, the speaker gives the command. It is quite different from deontic modality in this respect. While it may seem as though the sentence *You must go to school* is identical in all respects to *Go to school!*, there are thus important differences. In the construction with the modal verb the command can come from other sources beside the speaker (e.g., a person not present in the discourse or even abstract objects, in this case the law), and can be used on non-second persons (as in *He must go to school*). With an imperative the “commander” is the speaker and the “commandee” the hearer(s). There are syntactic differences as well which to a degree depend on the language in question. In English, an Imperative cannot be in a subordinate clause (*\*I said that go to school!*) but a corresponding sentence with a modal verb can (*I said that you must go to school.*).

## MODAL AFFIXES

In many languages modality is marked by means of affixes on the verb. This is, for instance, the case in Turkic languages, Greenlandic Eskimo, Dravidian languages like Tamil and many Native American languages. In (9), some examples from various languages are shown:

- (9) a. Tamil (Dravidian: Asher 1979:170; *-laam* permission)

avan peeca-**laam**

3SG speak-PERM

‘He is allowed to speak.’

- b. Koasati (Muskogean: Kimball 1991:200; *-sahá:wa* probability)

ó:la-fon ałi:ya-:**sahá:w**-ok ...

town-ALL go-PROB-SS.FOC

‘She must have gone to town.’ [also possibility]

- c. Turkish (Turkic: Lewis 1967:125-7; *-meli* necessitative)

gel-me-**meli**-siniz

come-NEG-OBLIG-2PL

‘You ought not to come.’

The difference between mood and modal affixes is that mood is an obligatory category. That is, a speaker of a language like Italian must choose between the Indicative and Subjunctive, while a speaker of, say, Tamil can choose not to use a modal affix. This is similar to English where speakers always have the option to use a modal auxiliary or not.

It is not always easy to tell if we are dealing with a mood or with a modal affix. Many grammars do not make a distinction between the two categories and indeed many works on modality do not do so either. Nevertheless, the distinction is worth making because a language can have both moods and modal affixes. Such a language is Turkish, which besides the Necessitative morpheme *-meli* also has a Subjunctive mood, see Lewis (1967:132ff). One way of telling moods and modal affixes apart might be to consider the degree of cohesion. If the morpheme can easily be separated from the rest of the verb (as

is the case with the morphemes shown in example (9) above), then we are dealing with an affix. If we are dealing with an inseparable part of the verb (as is the case in the Latin examples (5-6) above) then we are dealing with a mood. This area needs more research. Quite likely, this involves degrees of grammaticalization.

## LEXICAL MEANS

There are also less grammaticalized (and more lexical) means of expressing modality. These means can, for the most part, be exemplified by English, though some types are better known from other languages.

### *Modal adverbs and adjectives*

Modality can be expressed by means of adverbs. In English, typical examples are *probably*, *possibly*, *necessarily*, and *maybe*. This list is by no means exhaustive, obviously. Some examples are:

- (10) a. John is probably at home.  
 b. Maybe John is at home.  
 c. John is supposedly at home.

As mentioned in Perkins (1983:89), modal adverbs in English are primarily epistemic in nature.

In many languages these forms have become more grammaticalized and sometimes they are the only way to express modal notions. The main ways of expressing strong modality in Russian, for instance, are with adjectives (*dolžen*) and adverbs (*nado* or *nužno*). The former has a subject in the Nominative and declines according on the gender of the subject, while the latter takes a Dative subject and is invariant. An example is shown in (11):

- (12) a. ja dolžen idti v voksal  
 I.NOM must.MASC go.INF to station.ACC
- b. mne nado idti v voksal  
 I.DAT must go.INF to station.ACC
- ‘I must go to the station.’

### *Modal tags*

Epistemic modality, the expression of speaker’s confidence, can be expressed with such tags as *I think*, *I guess*, and *I believe*. (Thompson and Mulac 1991). An example is (1991:313):

- (13) It’s just your point of view you know what you like to do in your spare time  
 I think

While tags like *I think* are derived from pure matrix clauses, they behave more like modal adverbs and they show signs of grammaticalization. Corpus research shows that the complementizer *that* is often omitted (in about 90% of the cases). They can occur at various places in the sentence (initially, medially, and finally). A further step on this grammaticalization path would be the use of tags like *I think* as a pure adverb. This seems to have happened in certain Creole languages, such as Tok Pisin (Keesing 1988), where this tag has been simplified to *ating* with the meaning ‘maybe’.

While tags have not received very much attention in the literature, it seems that they are also mostly epistemic in nature. An exception might be volition verbs, like *want*. Although English is not a good example (the verb *want* has a different clause structure), something like this seems to have happened in the history of Greek. The Future particle *tha* in Modern Greek derives ultimately from the construction *thelo na* ‘want-1SG that’. Volition is a dynamic modal category, not an epistemic one, but this example shows that matrix clauses with a 1SG subject have a tendency to become tags, and then adverbs or particles.

#### *modal particles*

Another common means of expression is the modal particle. This is a method not very common in English. Modal particles are rarely found in British English but seem to become more popular in American English. An example is (14) where the words *too* and *so* function as “contrary to what you said/thought” modal particles.

- (14) a. Affective is **too** a word! (from the WWW)  
 b. There is **so** a Santa Claus!

Modal particles are well known from Germanic languages like Dutch and German (Abraham 1991), where they are ubiquitous. Two examples from German are shown in (15). As can be seen from the translations, it is not always easy to give an adequate rendering of modal particles in languages that lack them. Sometimes the best solution is to just leave them untranslated as in (15b) where the particle *doch* can best be translated with a slightly irritated intonation!

- (15) a. Kommt er **denn?** (1991:335)  
 come.3SG.PRES he MP  
 ‘Will he really come?’ or ‘Will he come after all?’
- b. Gib mir **doch** den Löffel! (1991:340)  
 give.IMP me.DAT MP the.ACC spoon  
 ‘Give me the spoon!’

One of the problems is that it is not always clear whether we are dealing with an adverb or a particle. This is a poorly researched area and it is likely that any distinction between the two must be made on a language-by-language basis. For instance, it has been claimed that modal particles in Dutch can never receive stress, nor can they occupy the first position in a sentence (ANS 1984:891). Modal adverbs can have stress and occupy

sentence-initial position. however, given the English data shown in (...), in which the particles do have stress, this may not be a cross-linguistically useful diagnostic.

The nebulous status of modal particles is at least partly due to their origins. German (and Dutch) modal particles can derive from a number of sources: Abraham (1991:332) lists modal particles that come from adverbs, adjuncts, scalar particles, adjectives and interjections. The typological concern addressed by Abraham is why languages like German and Dutch have a multitude of modal particles while closely related English has none (or at least very few). Abraham's (1991) view is that the peculiar syntax of German and Dutch is responsible. Modal particles occur mainly between the verbal elements of the sentence; the verbs *brace* the part of the sentence where modal particles typically occur. English (as well as the Romance languages) lack such a sentence part and hence lack the possibility of developing modal particles. There are problems with this analysis, however. Languages like Russian have a number of modal particles, yet word order is very free. The same goes for the older Indo-European languages (like Classical Greek) which teem with modal particles but whose word order is very free, or at least lack a German-type sentence structure.

Because modal particles have the entire sentence in its scope, they are often found at clause boundaries, and very often clause-finally. Cantonese (Matthews and Yip 1994, chapter 18) is an example of a language with a plethora of sentence-final particles, many of which are modal. An example is given in (16) in which the particle marks a polite request.

(16) Léih bái dō dī sihgaan ngóh lā !

you give more some time me PRT  
 “Give me a bit longer, won’t you?” (p. 351)

## MODAL CASE

The Tangkic language family of Northern Australia provides us with a typologically unusual device for marking modality. In Lardil, Yangkaal and Kayardild modality can be marked on the Noun as a case marker.<sup>7</sup> In Kayardild, (Dench and Evans 1988, Evans 1995, 2003) a (non-subject) noun phrase can have an optional case morpheme, besides a regular case marker, which denotes mood. An example is shown in (17), from Evans (2003:208). The morpheme *-u*, which occurs on all non-subject NPs, is called the Modal Propriative, is used to denote future and potential meanings. This morpheme must co-occur with the Potential suffix *-ju* which is found on the verb.

- (17) *dangka-a*      *burldi-ju*      *yarbuth-u*      *thabuju-karra-ngun-u*  
 man-NOM      hit-POT      bird-M.PROP      brother-GEN-INSTR-M.PROP  
*wangal-ngun-u*  
 boomerang-INSTR-M.PROP  
 ‘The man will/can hit the bird with brother’s boomerang.’

Besides the Modal Propriative, there are several other modal case morphemes, not all of them used for modal categories as we understand the term here; some have tense meanings (like the modal Ablative which, used together with the Past tense morpheme,

denotes anteriority). Evans (1995, chapter 10) is a full discussion of modal case and its development.

#### IV. REALIS AND IRREALIS

We now turn to some recent developments in the typological literature on modality. We will start with the realis – irrealis distinction. This distinction divides the world into real and unreal events and / or situations. It has been claimed that there are languages which encode modality in precisely this way, i.e., there are languages with irrealis morphemes, which mark an action or situation as unreal.

While languages with irrealis morphemes can be found on every continent, they have most prominently been described for New Guinea (Roberts 1990, Bugenhagen 1994). Other languages with irrealis morphemes can be found in North America (Chafe 1995, Mithun 1995).

One of the major problems in discussing irrealis issues is the fact that the term *irrealis* is very vague and can refer to a number of different circumstances (see Palmer 2001:149). Furthermore, the semantic content of irrealis morphemes differs from language to language (even between languages that are closely related). This makes defining a cross-linguistic category of irrealis very hard.

To give but one example, the future.<sup>8</sup> It can be argued that future is a prototypical irrealis category because it refers to events that have not yet happened and are therefore unreal. In languages like Amele and Muyuw the future is indeed an irrealis category.

However, in others it is treated as a realis category. One such language is the Native American language Caddo (Chafe 1995:358), shown in (18). The Future morpheme *-ʔaʔ* occurs not with the Irrealis prefix *t'a-/t'i-* but with the Realis prefix *ci-*:

(18) *cíibáwʔaʔ*

*ci-yi=bahw-ʔaʔ*

1SG.AG.REAL-see-FUT

'I will look at it.'

In yet other languages, the future can be used with either realis or irrealis, depending on the speaker's judgement of likelihood that the event described will actually occur. One such language is Central Pomo, a Californian language (Mithun 1995:378-80). The same goes for other categories that can be considered part of irrealis, including categories like negation, hypothesis, and imperative.

Given these facts, it appears that "irrealis" is a term that is not comparable from language to language and, consequently, it may be asked whether it is a useful or even valid object for typological research. This problem is examined in Bybee *et al.* (1994:236-40) and Bybee (1998). These studies reach the conclusion that the term "irrealis" is too broad to be of real use. This is illustrated with data from the Australian language Maung (Capell and Hinch 1970). In Maung, Irrealis and Realis have the following categories under them:

(19) *Realis and Irrealis in Maung*

	PRESENT:	Indicative present, future
REALIS	IMPERATIVE:	Negative only
	PAST:	Simple and complete past, imperfect
IRREALIS	PRESENT:	Potential, negative present and future
	PAST:	Negative past, Conditional, Imperative

This can be illustrated with the verb *-udba* ‘to put’ (p. 67):<sup>9</sup>

(20)	Realis		Irrealis
	<i>ŋi-udba</i>	<i>I put (pres.)</i>	<i>ni-udba-ji</i> <i>I can put</i>
	<i>ŋi-wan-udba</i>	<i>I shall put</i>	<i>marig ni-udba-ji</i> <i>I don't put</i> or
			<i>I won't put</i>
	<i>juwunji g-udba</i>	<i>don't put it!</i>	<i>da ŋi-udba-ŋji</i> <i>if I put</i>
	<i>ŋi-udba-ŋ</i>	<i>I put it (perfect)</i>	<i>marig ŋi-udba-ŋji</i> <i>I didn't put</i>
	<i>ŋi-udba-ŋ-uŋ</i>	<i>I was putting</i>	<i>g-udba-ŋji</i> <i>put it!</i>

The two problematic cases for treating the morpheme *-ji* as an Irrealis morpheme are the Future and the Imperative. The positive Future is a Realis category while the negative Future is an Irrealis one (it is identical to the negative Present). The positive Imperative is Irrealis while the negative Imperative (or Prohibitive) is a Realis one, at least in the



‘They will kill the pig as it runs out.’

It might be premature to call this joint-marking because the Irrealis morpheme occurs as part of a Different Subject morpheme which would have to be present in any case. A difference between realis and irrealis DS morphemes only shows up in examples like (21) above where the action in both clauses is simultaneous. When the action is not simultaneous or when the subjects in both clauses are identical, there is no difference in realis – irrealis marking on these morphemes.

The second type, called *non-joint*, has irrealis morphemes that do not need other morphemes but function all by themselves. An example is (22) from Muyuw (Bugenhagen 1994:18). The Irrealis morpheme *b(i)-* is used here to denote future. It contrasts with the Realis morpheme *n(i)-* which is used for realis events. Note that it is not necessary to choose between either morpheme. It is perfectly grammatical to omit the Realis and Irrealis morphemes. In that case the modality of the sentence is determined by other means, for instance by using the particle *bo* which denotes certainty.

- (22) yey    **b-a-n**            Lae    nubweig  
       I        IRR-1SG-go    L.        tomorrow  
       ‘I will go to Lae tomorrow.’

This distinction is a useful one if one keeps in mind that they are not absolute. The discussion on Muyuw shows that neither the Realis nor the Irrealis morpheme is obligatory; they can be omitted without impacting the modal status of the sentence. It is



tends to be used as a vague term for anything that is even remotely modal. But it is widely used in grammatical descriptions. In the literature on pidgins and creoles it is the standard way of referring to modal distinctions (cf. Bickerton 1975, Holm 1988:164-6).

It may be that the difficulties sketched in this section are insurmountable and that we will have to give up realis and irrealis as linguistic categories. It may also be possible to retain the terms, and to account for variation within these categories. One way of doing so might be with the use of semantic maps a discussion of which we will now turn.

## V. SEMANTIC MAP OF MODALITY

One of the most powerful models that try to come to grips with the complex interactions of modal meanings in the world's languages is the *semantic map* model: a representation that is the sum total of the semantic possibilities of the category under investigation. An exponent of this category in a given language can then be mapped onto this representation and thus be compared to similar means of expression in other languages.<sup>10</sup>

A further advantage of semantic maps is that they can be used for synchronic as well as diachronic purposes, i.e., they can be used to chart possible changes in meaning. This faculty is put to good use in Van der Auwera and Plungian (1998) which is an application of semantic maps to chart meanings and changes in meaning in modal elements. Their investigation starts with the grammaticization paths of Bybee et al. (1994), which can be considered to be diachronic semantic maps. One such path, the path of development from ability is shown in Fig. 1:

-----  
 Fig. 1, which is Fig. 6.3. in Bybee et al. (1994:240)  
 -----

This path shows the development of a morpheme meaning ability (in Bybee et al.'s terms, this is an ability *gram*). First, it develops into a marker of root possibility (see above) and then into a *gram* meaning permission or epistemic possibility (or both, as in English *may*). After this, the morpheme may take on additional meanings. This path is taken to be universal and *unidirectional*, that is, a morpheme with an ability meaning will follow the path in Fig. 1, and does not skip a step (i.e., an ability morpheme does not change directly into an epistemic possibility *gram*). This does not mean, of course, that an ability *gram* necessarily will take on these additional meanings. Language is not deterministic and an ability *gram* can quite happily remain an ability *gram*. If it is going to take on additional meanings, the first one will be that of root possibility.

Besides ability, Bybee et al. provide other grammaticization paths for the area of mood and modality. The overall conclusion (1994:241) is that the overall path is from agent-oriented modality to the other kinds of modality, with subordinate modality at the end of the path. This corresponds to a development from less to more grammaticalization. The agent-oriented modal *grams* are typically least likely to be bound (i.e., are least likely to be an affix on a verb, but rather tend to be independent morphemes, like English *can*), while subordinate modal *grams* are most likely to be bound (i.e., are most likely to be part of the verb, as is the case with subjunctive morphemes, for instance). There are some problems, as the authors freely admit. For instance, the prediction is that imperatives, a

speaker-oriented modality, is less frequently bound than subjunctives, a subordinate modality. However, the opposite is the case (1994:242). Nevertheless, there is a general correlation between the various modal categories and their boundedness.

Van der Auwera and Plungian (1998) take these grammaticalization paths and develop them into a full-blown semantic map. Fig. 2 below shows their representation of Bybee et al.'s ability path, shown in Fig 1. above.

-----

Fig. 2, which is Fig. 5 in Van der Auwera and Plungian (1998:91)

-----

The semantic map is divided in three parts. The main part is the large enclosed rectangle which shows the modal domain. To the left side of the rectangle is the premodal domain graphically represented by dotted blocks. This contains the lexical sources for the modal domain. To the right, the postmodal domain contains further grammaticalization paths for modal grams, which Van der Auwera and Plungian do not consider part of the modal domain.

Within the modal domain, individual meanings (for terminology, see section 2 above) are represented by ovals, and arrows mark the pathways from one meaning to another. In this representation, meanings can be part of other meanings. As mentioned above, deontic possibility is part of participant-external possibility in this model. The arrow wholly within the oval for participant-external possibility shows the pathway to deontic possibility.

By unifying Bybee et al.'s grammaticization paths plus adding their own material, Van der Auwera and Plungian end up with the following map for modality:

-----

Fig. 3, which is Fig. 19 in Van der Auwera and Plungian (1998:111)

-----

The basic representation is similar to the one shown in Fig. 2 above. Represented are, from left to right, the premodal, modal, and postmodal domains. The premodal and postmodal domains are represented abstractly. As before, grammaticalization paths are denoted by arrows, with the arrows leading from premodal to modal and those from modal to postmodal being abstract representations.

As is readily visible, some meanings are contained in larger enclosures. For instance, the notions of participant-external possibility and participant-external necessity are grouped together into a larger meaning block. This represents the fact that in some languages, these individual meanings are expressed by one and the same morpheme. An example of this is Swedish *få* 'get', which is vague between both types of participant-external modality (*vagueness* refers to the fact that it is impossible to decide on one interpretation in favor of another, often even in context):

(24) **Swedish** (Van der Auwera and Plungian 1998:103, citing Wagner 1976:56)

Lasse får köra bil.

Lasse gets drive car

“Lasse gets to drive the car” > “Lasse may / must drive the car.”

Similarly, there are languages in which all type of participant modality can be expressed with one vague construction. Van der Auwera and Plungian cite the German modal infinitive (sometimes called the *modal passive*, which is similar to English constructions like *We are to meet there at seven*). It is said to be vague between all types of non-epistemic modality.

By combining diachronic changes and synchronic states, one can get a good representation of what is a possible interpretation of a given modal element and what is universally ruled out. The semantic map shown in Fig. 3 is certainly not the last word on the matter. For one, several modal meanings are not included in this map (such as imperatives, which were deliberately omitted by the authors). Nevertheless, it represents an important and fruitful way of looking at a conceptually and linguistically complex category.<sup>11</sup>

## VI. MODALITY AND TENSE

This section deals with the interaction of modality and tense. In the typological literature tense is commonly defined as the grammaticalized expression of location in time (Comrie 1985:9). There are clear interactions between the area of tense and that of modality. An obvious candidate for such interaction is the future. Since events in the future have not (yet) happened, it is easy to see that there is a certain amount of uncertainty surrounding

it, hence there is a connection between future and epistemic modality. The future is sometimes classified as a realis category, and sometimes as an irrealis category.

There is a similar connection between modality and the past tense. This seems at first to be somewhat counterintuitive since the events happened in the past and one can be certain (or, at least, modally neutral) about past events. Nevertheless, the past tense is routinely used to express notions of modality such as hypotheticals and conditionals.

### *Modality and future tense*

The fact that future events can be described both temporally and modally is an observation which goes back a long time. Discussions on the relation between the two areas can be found in Comrie (1985:43-6), Dahl (1985:103ff, 2000b), Bybee (1988), Bybee et al (1991,1994) and Palmer (1986, 2001) among others.<sup>12</sup>

Future is quite often a part of so-called irrealis morphemes, as discussed above. Nevertheless, there is no one-to-one correlation between the two categories as there are many languages (Maung, Caddo, Latin) in which the future is a realis category.

Despite the fact that future events have not come to pass and the speaker cannot know for certain that the event will occur, there are languages in which these considerations either do not play a role or in which there is a choice between various future tense forms to denote various shades of certainty. A language in which the future has been analyzed as a pure tense (that is, it refers to events occurring subsequent to the moment of speech without conveying a modal meaning as well) is the Tibeto-Burman language Manipuri

(Bhat 1999:18-19). In Manipuri, there is a basic future/non-future distinction. That is, there is one morpheme to denote future and one to denote present and past tense. In (25) below, the morpheme *-ŋi* is used for Non-future (a,b) and *-kəni* for Future (c).

- (25) a.     $\eta\text{əsi}$      $\text{no}\eta$      $\text{mə}\eta\text{-}\mathbf{\eta\text{i}}$   
           today rain    cloudy-NFUT  
           ‘It is cloudy today.’
- b.     $\eta\text{əra}\eta$           $\text{no}\eta$      $\text{mə}\eta\text{-}\mathbf{\eta\text{i}}$   
           yesterday    rain    cloudy-NFUT  
           ‘It was cloudy yesterday.’
- c.     $\text{nu}\eta\text{da}\eta\text{wayrəmdə}$      $\text{no}\eta$      $\text{mə}\eta\text{-}\mathbf{gəni}$   
           evening.LOC         rain    cloudy-FUT  
           ‘It will be cloudy in the evening.’

There are languages in which a speaker has a choice between different future morphemes to denote various shades of certainty. In Bybee et al. (1994:247-8) several languages are listed in which there are two or more future morphemes with various levels of confidence. An example is Southern Agaw, a Cushitic language (Bybee et al. 1994:248, data cited from Hetzron 1969). The Future Certainty morpheme *-aGa* is used when the speaker is certain that the action will occur, while *-e* is the Future Possibility morpheme.<sup>13</sup>

- (26) a.     $\text{táq-}\mathbf{áGá}$

know-2SG-FUT.CERT

‘You will [certainly] know [it].’

b. dəngéta      ča      des-é

perhaps      tomorrow      study-FUT.POSS

‘Perhaps tomorrow I shall study.’

Another connection between modality and future is the fact that quite often future morphemes develop from modal (deontic) forms. This has happened in English, of course, where the modal verbs *will* and *shall* were originally modal verbs.<sup>14</sup> The connection between obligation/volition and future is that one can only lay an obligation on someone to do something in the future. Thus, a sentence such as *you must go to school* means that the action of going to school is necessarily subsequent to the moment at which the obligation was uttered. This is a widespread development, accounting for the vast majority of cases in Bybee et al (1994). Similarly, Fleischman (1982:145-8) discusses the French verb *devoir* ‘must, ought to’ which functions in many respects like a marker of future tense rather than obligation.

(27) Je      dois      diner      avec      Joseph      la      semaine      prochaine.

I      must      dine      with      j.      ART      week      next

‘I must / will have dinner with Joseph next week.’      (1982:146)

‘I am to have dinner with Joseph next week.’

Fleischman comments that the modal verb can be replaced by other future forms (the go-future or the synthetic future) without much change in meaning.

*Modality and past tense*

While the past tense is usually taken as a pure temporal category, because it refers to events that are immutable and known. Nevertheless, there is a connection between the past tense and modality. This can be demonstrated even with English. The Past tense morpheme *-ed* is usually a pure tense morpheme, yet in certain environments it can mark various modal meanings, as in the following examples, from Comrie (1985:19):

- (28) a. If you did this, I would be very happy.  
 b. If John was / were here ...  
 c. I just wanted to ask you if you could lend me a pound.

Sentences (28a) and (b) are counterfactuals and hypotheticals, which are usually considered to be modal in nature (they are typical irrealis categories). Sentence (c) is a polite request and refers to a non-actual event as well; a request is a type of wish. The question is: is this an isolated fact of English (in which case examples (a) and (b) could be explained away as an instance of homophony: the irrealis forms in (a) and (b) are homophonous with the regular past tense forms), or is this a cross-linguistic feature (in which case we can look for a principled account of the phenomenon)?

Steele (1975) is an early paper that discusses the relation between past tense and modality. It is a reconstruction of part of the tense-aspect-modality system of Proto-Uto-Aztecan. She reconstructs two morphemes: *\*ta-* as a general Irrealis morpheme, and *\*ta-* as a Past tense morpheme. She then goes on to state that both are actually the same morpheme and that there is one abstract feature that underlies both categories. This feature is called *dissociative* as past tense is dissociated from the present and irrealis is dissociated from reality. Steele suggests that this observation is valid cross-linguistically. This view of past tense as a ‘remoteness’ device has been echoed in other works, such as James (1982) and Fleischman (1989). Palmer (1986:211; 2001:210) considers this line of reasoning circular but does not provide any real alternative. Bybee (1995:513-6), which is a paper concerned with the development of the past tense forms of the modals *should* and *would*, rejects the notion that it is the past tense alone that is responsible for the modal interpretation. In her view, it is the combination of past tense plus some other element, such as a modal verb, the subjunctive or (per Fleischman 1995) the imperfective aspect.

## VII. MODALITY AND NEGATION

This section looks at the literature on the interaction of modality and negation. By itself, negation has been considered by some scholars to be part of modality. The reason for this is that by talking about nonexistent events or states, we are talking about events or states that are not real. Hence, negation can be considered an instance of an irrealis category and it is so classified in some languages as was discussed above. Example (29) shows such an example, from the Native American language Caddo (Chafe 1995:355). When

the negative prefix *kúy-* is present, the Irrealis form of the pronominal prefixes must be used.

(29) *kúyt'áybah*

*kúy-t'a-yi=bahw*

NEG-1.AG.IRR-see

'I don't see him.'

In other languages, the presence of a negation has no influence on the choice of realis/irrealis. Latin has already been mentioned and we can add the Native American language Central Pomo (Mithun 1995:380-2) and the Papua New Guinean language Amele (Roberts 1990) to this list. This difference can occur in very closely related languages. Mithun (1995:383-4) makes mention of two dialects of the Yuman language Diegueño, namely Mesa Grande and Jamul. The former has an obligatory irrealis morpheme whenever a negation is present, in the latter negation plays no role in the choice of irrealis or realis.

The fact that there is a relation between modality and negation as far as scope is concerned has been known for at least 2500 years. Ever since Aristotle philosophers and logicians have been concerned with this issue and it has attracted the attention of linguists as well, but only comparatively recently.

While linguistic issues are different from the issues in modal logic, for ease of reference the symbols from modal logic will be retained, even though it must be kept in

mind that there is no exact correspondence between the two realms (see Karttunen 1972 for the basic facts). In (30) below, the basic symbols are mapped on sample sentences from English.

- (30)
- |    |                              |   |
|----|------------------------------|---|
| a. | John must be a bachelor.     | $\Box p$                                  |
| b. | John may be a bachelor.      | $\Diamond p$                              |
| c. | John must not be a bachelor. | $\Box \neg p$                             |
| d. | John need not be a bachelor. | $\neg \Box p$                             |
| e. | John may not be a bachelor.  | $\Diamond \neg p$ (or $\neg \Diamond p$ ) |

There are two questions that have been addressed in the recent typological literature.

They are:

- How are modal meanings mapped onto the logical possibilities?
- How do languages disambiguate possible scope ambiguities?

*Mapping of modals onto logical structures*

This question deals with the problem of what types of modals exist and can exist. We know there are modals that express the notion  $\Box$  and  $\Diamond$ , because they are found in a

language like English as demonstrated in (30). But are there for instance modals that specifically and uniquely express the notion  $\Box\neg$  or  $\neg\Diamond$ ?

In two important papers Van der Auwera investigated the connection between modality and negation. Van der Auwera (1996) investigated the inadequacy of the well-known *square of oppositions* for dealing with modality in natural language (see e.g., Horn 1989 on this issue). Van der Auwera (2001) is an investigation into which combinations of modality and negation receive a specialized modal element.

In Van der Auwera (1996), a distinction was made between two logical types of possibility. One, known from logic dating back to antiquity, is the familiar  $\Diamond$ , used in such formulae as:

- (31) a.  $\Diamond p = \neg\Box\neg p$  ('it is possible that p = it is not necessary that not-p')
- b.  $\Box p \rightarrow \Diamond p$  ('it is necessary that p entails it is possible that p')

The logical operator  $\Diamond$  is not always directly translatable into real language use. This has been known for a long time as well (As argued by Van der Auwera it fails to differentiate the uses of the modal *may* in (32a). This sentence suggests the equivalency shown in (32b) which is absurd from a logical point of view.

- (32) a. John may be there and he may not be there.
- b.  $\Diamond p = \Diamond\neg p$

Example (32a) illustrates what is called the *contingent* use of the modal *may* and not the *possibility use*. This was argued by Hintikka (1960) and also noted by Horn (1989, among others). Hence a new symbol is necessary to distinguish between the two uses. Van der Auwera suggests a new operator  $\blacklozenge$  which is used for cases like (32a). the relation between  $\diamond$  and  $\blacklozenge$  is one of implicature. When something is possible, it is taken for granted that it is also contingent:

(33)  $\diamond p$  conversationally implicates  $\blacklozenge p$

There appear to be no languages in which a lexical distinction is made between weak modals that express possibility and contingency.

The next question is: are there specialized modals for every possible logical combination of modality and negation. This question is addressed in van der Auwera (2001) and, to an extent, in De Haan (1997). Van der Auwera surveys a number of languages, mostly European ones, and provides us with the following cases, here abbreviated:

(34)

modality	Language	element
$\neg\Box p$	English	<i>need</i>
	Dutch	<i>hoeven</i>

	Kashmiri	
$\Box\neg p$	Bengali	<i>nei</i>
$\neg\Diamond p$	Russian	<i>nel'zja</i>
$\Diamond\neg p$	(English	<i>mightn't</i> )

That is, every possible combination is attested in at least one language, with the possible exception of  $\Diamond\neg p$ . The combination form *mightn't* is very rare and not accepted by many speakers (its equivalent form *mayn't* is even rarer and even less accepted). The reason for the rarity of specialized  $\Diamond\neg p$  forms is possibly the peripheral status of  $\Diamond\neg p$  itself.

An issue is the status of double modals. It is well-known that strong modality can be expressed by means of double negation and a weak modal, conform the logical equivalency of  $\Box p = \neg\Diamond\neg p$ . This can be seen in English constructions like:

(35) John can't not go to school = John must go to school

There are languages in which strong modality is expressed only by means of a weak modal and a double negation. That is, there is no separate strong modal in the language.

In English, the two are stylistic variants, but in the following languages  $\neg\Diamond\neg p$  is the only way to express  $\Box p$ :

- (36) a. Malagasy  
 tsy main-tsy  
 NEG able-NEG (Horn 1989:220)
- b. Classical Tibetan  
 mI-V mthu mI-V  
 NEG-V ability NEG-V (Beyer 1992:247-8)
- c. Japanese  
 Rekishi no hon o yoma-na-kereba nari-masen  
 History GEN book ACC read-NEG-PROV work.out-NEG  
 ‘I have to read history books.’ (Han 1983:341)

There is also an equivalence between  $\neg\Box\neg p$  and  $\Diamond$ , but there seem to be no languages that have a weak modal that is made up of a strong modal and a double negation. Even in languages that allow double negations and strong modals (such as Russian) there are severe restrictions on the use of such constructions.

### *Resolving scope ambiguities*

The next question is how do languages differentiate between, say,  $\Box\neg p$  and  $\neg\Box p$ ? in other words, how do languages resolve scope ambiguity in the area of modality and negation. The basic problem can be illustrated by considering the sentences in (37). In sentence (37a) the negation is in the scope of the modal, while the modal is in the scope

of the negation in (37b). The first instance is a case of *narrow scope* and the second one of *wide scope*, both with respect to the negation.

- (37) a. John must not go to school.  $\Box\neg p$   
 b. John need not go to school.  $\neg\Box p$

As is clear from the example, English in this case uses two different modal verbs to show the difference in scope. The modal verb *must* is used for narrow scope, and *need* is used for wide scope. De Haan (1997) calls this the *Modal Suppletion Strategy* for disambiguation.

Another strategy can be exemplified by a language like Russian, as shown in (38). The modal verb *možet* 'may/can' interacts with the negation *ne* in the following way:

- (38) a. Ivan ne možet rabotat'.  $\neg\Diamond p$   
 Ivan NEG can.3SG work.INF  
 'Ivan is not allowed/able to work.'  
 b. Ivan možet ne rabotat'.  $\Diamond\neg p$   
 Ivan can.3SG NEG work.INF  
 'Ivan is allowed/able not to work.'

In this example from Russian, the modal verb stays the same but the place of the negation changes. This strategy is referred to in De Haan (1997) as the *Negation Placement*

*Strategy*. It iconically shows the scope relation: if the negation precedes the modal, it has wide scope but if it immediately precedes the main verb it has narrow scope.

English uses primarily the first strategy, but there are some verbs that do not have suppletion. As seen in (30d) above, the verb *may* is ambiguous when a negation is present (and also the verb *can*), although there are prosodic differences depending on scope. Another complication is the fact that languages that use the Negation Placement strategy also tend to have the phenomenon of NEG-raising which muddies the iconic relation.<sup>15</sup> Since Russian is one of these languages, sentence (38a) can be interpreted with a narrow scope reading:

- (39) Maša ne dolžna rabotat'                     $\neg \Box p / \Box \neg p$   
 M.    NEG   must.FEM   work.INF  
 'Masha mustn't/needn't work.'

## VIII. EVIDENTIALITY

The category of evidentiality deals with the source of evidence a speaker has for his or her statement.<sup>16</sup> Evidential morphemes mark whether a speaker has been a direct witness to the action he/she is describing or whether s/he has received the information about the action or event from another source. Thus evidentiality can be divided into two subcategories, *direct* evidentiality which marks that the speaker was a witness to the action, and *indirect* evidentiality which marks that the speaker was not a witness but obtained knowledge about the action from another person (hearsay or quotative

evidentiality) or through inference such as tracks in the snow, etc. (inferential evidentiality). This is exemplified in (40), from Evenki (Tungusic; Nedjalkov 1996:239), where the suffix *-re* in (a) is used for direct evidence, and *-che* in (b) for indirect evidence. The Turkish data in (41), from Slobin and Aksu (1982) and DeLancey (1997:37), shows the evidential use of the two past tenses, *-di* to show direct evidence, *-miş* indirect evidence (see also example (44) below for other uses of *-miş*).

- (40) a.    eni    eme-**re**-n  
           mother come-NFUT-3SG  
           ‘mother came.’ (direct evidence)
- b.    eni    eme-**che**-n  
           mother come-PST-3SG  
           ‘mother came.’ (no direct evidence)

- (41) a.    Kemal gel-**di**  
           K.     come-PAST  
           ‘Kemal came.’        (direct)
- b.    Kemal gel-**miş**  
           K.     come-PAST  
           ‘Kemal came.’        (indirect)

Evidentiality has for a long time been considered to be an exotic category, associated with Native American languages primarily. However, it is a category that manifests itself

in languages on every continent, including in some well-studied languages of Europe, such as German. English lacks a grammatical category of evidentiality but it can be expressed with lexical means like *evidently*, *ostensibly* and with verbs like *seem*.

Evidentiality is typically considered to be a modal category in the typological literature. Palmer (1986:51) distinguishes four kinds of epistemic modality; one is the degree of confidence in the truth discussed above and the other three types are evidential categories.<sup>17</sup> This point of view is found in other typological studies of modality or evidentiality, such as Bybee (1985), Bybee et al. (1994), Willett (1988), and Frawley (1992).

The reasoning for treating evidentiality as a modal category is the belief that one is inherently less certain about actions one has not witnessed than about those one has witnessed. This belief is anchored in the fact that when one sees something with one's own eyes, one tends to accept that sight as a true representation of the world while a secondhand report is viewed with more suspicion. Hence, indirect evidentiality presents the action as less certain than does direct evidentiality.

While this is an appealing belief *a priori*, there are some problems with it. First, it is risky to compare grammatical categories across languages, as the typological literature has amply shown. This is especially true for notional categories that are strongly rooted in a subjective environment (that is, an environment that is inherently linked to the speaker), like evidentiality and epistemic modality. It is hard to compare modal verbs in related languages, and comparing English to, say, Evenki, might be extremely hard. To illustrate this, example (42) from De Haan (1999) shows a Dutch sentence with the modal

verb *moeten* which is cognate with English *must* and yet a straightforward comparison is not always possible.

- (42) Het **moet** een goede film zijn.  
 It must a good movie be.
- a. It is bound to be a good movie. (epistemic)  
 b. It is required to be a good movie. (deontic)  
 c. It seems to be a good movie. (evidential)

Sentence (42) has three possible interpretations (as always, out of context), only two of which have *must* as a possible English translation, (a) and (b). In its evidential reading, *must* is not appropriate. Thus, even in two closely related languages there is not always a one-to-one correspondence. Consequently, how can one assess the status of a sentence like (43) from Tuyuca, an Eastern Tucanoan language from the Vaupés River region of Western Amazon (Barnes 1984:257).

- (43) diíga apé-yi  
 soccer play-INFER.3SG.MASC.PAST  
 ‘He played soccer.’ (I have seen evidence that he played,  
 but I have not seen him play)

A not unusual analysis is to equate the Inferential in Tuyuca with the English category of probability (possibly by translating the above sentence as *he must have played soccer*),

and lump the two together and claim that the two are synonymous. Then epistemic modality has been introduced through the backdoor and the ‘analysis’ is done. Note that at no stage of the game has it been shown that the Inferential in Tuyuca is a modal category. Indeed, there is good reason to assume that it is not. Not only does Barnes never use the term ‘(epistemic) modal’ in her description, but there are also true epistemic modals in the language. By translating an evidential with the modal *must* the analysis is prejudiced at best, wrong at worst.

A recent proposal is to analyze evidentiality not as a modal, but as a deictic category (De Haan 2001, 2003, forthcoming). Frawley (1992) already made an attempt in this direction but called it *modal deixis*, so still a modal category. There are also early hints in Givón (1982, 1984). De Haan (forthc.) notes the connection between spatial deictic elements such as demonstratives, temporal deictic elements, such as tense, and evidential elements. In all cases, the morphemes in question denote the distance between the speaker and: an object (spatial), time (temporal) and the entire proposition (evidential). Hence De Haan (forthc.) proposes the term *propositional deixis* for evidentiality.

## MIRATIVITY

An offshoot from the research into evidentiality is *mirativity*. This category refers to the fact that the speaker has received his information from an unexpected source (DeLancey 1997). The connection with evidentiality is that the two categories are usually expressed by the same morphemes in those languages that have both categories. An example is (44)

from Turkish (DeLancey 1997:37, citing Slobin and Aksu 1982, but my translations). It has already been mentioned above that there are two past tenses in Turkish, one for direct information, one for indirect. The Indirect Past tense *-miş* can also be used to mark unexpected information. The context for (44) is that Turkish Prime Minister Ecevit resigned unexpectedly, whereas President Nixon's resignation was widely expected. This warrants the use of the *-miş* Past in (a) and the *-di* Past in (b). Note that this parallels the use of the so-called Hot News Perfect in English (McCawley 1971).

- (44) a. Ecevit istifa            **et-miş**  
       E.     resignation    make-PAST.MIR  
           'Ecevit has resigned!'
- b. Nixon istifa            **et-ti**  
       N.     resignation    make-PAST.DIR  
           'Nixon resigned.'

DeLancey (1997) notes that there is a widespread correspondence between mirative and (indirect) evidential morphemes. He cites examples from Hare Slave, several Tibeto-Burman and Dardic languages, and mentions several more, from all parts of the world.

Although a thorough analysis of the connection between evidentiality and mirativity is still to be done, as indeed an analysis of evidentiality and its proper place in grammar, it is clear that this particular area is currently one of the more exciting ones in all of

linguistics. There is every reason to think that a convincing analysis of evidentiality and related areas will prove important for all areas of linguistics.

## REFERENCES

- Abraham, Werner (1991). The grammaticization of the German modal particles. In Traugott and Heine (eds.), 331-80.
- Adéwolé, Fémi (1990). Gbódò 'Must': Analysis of a Yoruba modal verb. *Journal of West African Languages* 20, 73-82.
- Aikhenvald, A. and R.M.W. Dixon. 2003. *Studies in evidentiality*. Amsterdam: Benjamins.
- Anderson, Lloyd B. (1982). The 'perfect' as a universal and as a language-particular category. In Hopper (ed.), 227-64.
- Anderson, Lloyd B. 1986. Evidentials, Paths Of Change, And Mental Maps: Typologically Regular Asymmetries. In Chafe and Nichols, 273-312.
- Asher, N. (1979). *Tamil*. Lingua Descriptive Series. Amsterdam: North Holland.
- Barnes, Janet. 1984. "Evidentials in the Tuyuca Verb". *International Journal of American Linguistics* 50, 255-271.
- Beyer, Stephan V. (1992). *The Classical Tibetan Language*. Albany: SUNY Press.
- Bhat, D. N. S. (1991). *Grammatical Relations, the evidence against their necessity and universality*. London: Routledge.
- Bhat, D. N. S. (1999). *The Prominence of Tense, Aspect, and Mood*. Amsterdam: Benjamins.
- Bickerton, Derek (1975). *Dynamics of a creole system*. Cambridge: Cambridge University Press.

- Buck, Carl D. (1933). *Comparative Grammar of Greek and Latin*. Chicago: University of Chicago Press.
- Bugenhagen, Robert D. (1994). The semantics of irrealis in the Austronesian languages of Papua New Guinea. In Ger P. Reesink (ed.) *Topics in descriptive Austronesian linguistics*. Leiden: Rijksuniversiteit Leiden.
- Bybee, Joan (1985). *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: Benjamins.
- Bybee, Joan L. 1988. the diachronic dimension in explanations. In John A. Hawkins (ed.) *Explaining Language Universals*. Oxford: Blackwell, 350-79.
- Bybee, Joan L. (1995). The Development of Past Tense Modals in English. In Bybee and Fleischman (eds.), 503-17.
- Bybee, Joan L. (1998). "Irrealis" as a grammatical category. *Anthropological Linguistics* 40, 257-71.
- Bybee, Joan, William Pagliuca and Revere Perkins. 1991. Back to the Future. In Traugott and Heine (eds.), 17-58.
- 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.
- Bybee, Joan L. and Suzanne Fleischman (1995). *Modality in Grammar and Discourse*. Amsterdam: Benjamins.
- Capell, A. and H. E. Hinch (1970). *Maung Grammar: texts and vocabulary*. The Hague: Mouton.

- Chafe, Wallace (1995). The Realis – Irrealis Distinction in Caddo, the Northern Iroquoian languages, and English. In Bybee and Fleischman (eds.), 349-66.
- Chafe, Wallace, and Johanna Nichols. 1986. Evidentiality: the Linguistic Coding of Epistemology. Norwood, NJ: Ablex.
- Chung, Sandra and Alan Timberlake. 1985. Tense, aspect and mood. In Tim Shopen (ed.) *Language typology and syntactic description, vol. III*. Cambridge: Cambridge University Press, 202-58.
- Coates, Jennifer (1983). *The Semantics of the Modal Auxiliaries*. London: Croom Helm.
- Comrie, Bernard (1985). *Tense*. Cambridge: Cambridge University Press.
- Croft, William. (1991). *Syntactic categories and grammatical relations*. Chicago: University of Chicago Press.
- Croft, William (2001). *Radical Construction Grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William (2003). *Typology and Universals, second edition*. Cambridge: Cambridge University Press.
- Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.
- Dahl, Östen, ed. 2000a. *Tense and Aspect in the Language of Europe*. Berlin: Mouton de Gruyter.
- Dahl, Östen. 2000b. The grammar of future time reference in European languages. In Dahl (ed.), 309-28.
- De Haan, Ferdinand. 1997. *The Interaction of Modality and Negation: A Typological Study*. New York: Garland.

- De Haan, Ferdinand. 1999. "Evidentiality and Epistemic Modality: Setting Boundaries". *Southwest Journal of Linguistics* 18, 83-101.
- De Haan, Ferdinand. 2001. "The Place of Inference within the Evidential System". *International Journal of American Linguistics* 67, 193-219.
- De Haan, Ferdinand. 2003. "Visual evidentiality and its origins". Ms. University of Arizona.
- De Haan, Ferdinand (forthcoming a). Encoding Speaker Perspective: evidentials. In Z. Frajzyngier, D. Rood and A. Hodges (eds.) *Linguistic diversity and language theory*. Amsterdam: Benjamins.
- De Haan, Ferdinand (forthcoming b). Semantic distinctions of evidentiality. In Haspelmath (et.al).
- De Haan, Ferdinand (forthcoming c). coding of evidentiality. In Haspelmath (et.al).
- DeLancey, Scott. 1997. Mirativity: the grammatical marking of unexpected information. *Linguistic Typology* 1.33-52.
- Dench, Alan and Nicholas Evans (1988). Multiple case-marking in Australian languages. *Australian Journal of Linguistics* 8: 1-47.
- Dendale, P, and L. Tasmowski, eds. (2001). Evidentiality. Special issue of *Journal of Pragmatics* 33(4).
- Diewald, Gabriele. (1999). *Die Modalverben im Deutschen: Grammatikalisierung und Polyfunktionalität*. Tübingen: Niemeyer.
- Dik, Simon C. (1997). *The Theory of Functional Grammar* (2 volumes). Berlin: de Gruyter.

- Ehrman, Madeleine. (1966). *The meaning of the modals in Present-day English*. The Hague: Mouton.
- Evans, Nicholas D. (1995). *A Grammar of Kayardild, with Historical-Comparative Notes on Tangkic*. Berlin: Mouton de Gruyter.
- Evans, Nicholas (2003). Typologies of Agreement: Some problems from Kayardild. *Transactions of the Philological Society* 101:203-34.
- Fleischman, Suzanne (1982). *The Future in Thought and Language*. Cambridge: Cambridge University Press.
- Fleischman, Suzanne (1989). Temporal distance: a basic linguistic metaphor. *Studies in Language* 13, 1-51.
- Fleischman, Suzanne (1995). Imperfective and Irrealis. In Bybee and Fleischman (eds.), 519-51.
- Foley, William A. and Robert D. Van Valin. (1984). *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Frawley, William. 1992. *Linguistic semantics*. Hillsdale, NJ: Lawrence Erlbaum.
- Geerts, C. et al. (1984). *Algemene Nederlandse Spraakkunst*. Groningen: Wolters Noordhoff.
- Givón, Talmy. 1982. Evidentiality and epistemic space. *Studies in Language* 6.23-49.
- Givón, Talmy. (1984). *Syntax, vol. 1*. Amsterdam: Benjamins.
- Givón, Talmy (1994). Irrealis and the Subjunctive. *Studies in Language* 18.265-337.
- Haarmann, Harald (1970). *Die indirekte Erlebnisform als grammatische Kategorie*. Wiesbaden: Harrassowitz.
- Hale, William G. and Carl D. Buck (1903). *A Latin Grammar*.

- Han, Mieko S. (1983) *Modern Japanese, second edition*. Los Angeles: Institute for Intercultural Studies Press.
- Hansen, Björn (2001). *Das slavische Modalauxiliar. Semantik und Grammatikalisierung im Russischen, Polnischen, Serbischen/Kroatischen und Altkirchenslavischen*. München: Verlag Otto Sagner.
- Haspelmath, Martin (1997). *Indefinite Pronouns*. Oxford: Oxford University Press.
- Haspelmath, Martin (2003). The geometry of grammatical meaning: semantic maps and cross-linguistic comparison. In M. Tomasello (ed.) *The new psychology of language: cognitive and functional approaches to language structure, vol. 2*. Mahwah, NJ: Lawrence Erlbaum, 211-42.
- Haspelmath, Martin, Matthew Dryer, Bernard Comrie and David Gil, eds. (forthcoming). *World Atlas of Language Structures*. Oxford: Oxford University Press.
- Hengeveld, K. (...). Mood and Modality. In G. Booij, C. Lehmann and J. Mugdan (eds.). *Morphology: A handbook on inflection and word-formation*. Berlin: de Gruyter.
- Hetzron, Robert (1969). The Verbal system of Southern Agaw. Berkeley: University of California Press.
- Hintikka, K. J. (1960). Modality and Quantification. *Theoria* 27, 119-28.
- Holm, John (1988). *Pidgins and Creoles. Vol. I, Theory and Structure*. Cambridge: Cambridge University Press.
- Hopper, Paul, ed. (1982). *Tense – Aspect: between semantics and pragmatics*. Amsterdam: Benjamins.
- Horn, Laurence R. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press.

- Hughes, G.E. and M.J. Cresswell (1996). *A new introduction to Modal Logic*. London: Routledge.
- James, Deborah (1982). Past tense and the hypothetical: a cross-linguistic study. *Studies in Language* 6: 375-403.
- Johanson, Lars; Bo Utas, eds. (2000). *Evidentials*. Berlin: Mouton de Gruyter.
- Karttunen, Lauri (1972). Possible and must. In John P. Kimball (ed.) *Syntax and Semantics, vol. 1*. New York: Academic Press, 1-20.
- Keesing, Roger (1988). *Melanesian Pidgin and the Oceanic Substrate*. Stanford: Stanford University Press.
- Kemmer, Suzanne (1993). *The middle voice*. Amsterdam: Benjamins.
- Kimball, Geoffrey D. 1991. *Koasati Grammar*. Lincoln: University of Nebraska Press.
- Krug, Manfred G. (2000). *Emerging English modals: a corpus-based study of grammaticalization*. Berlin: Mouton de Gruyter.
- Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Oxford University Press.
- Lichtenberk, F. (1991). Semantic change and heterosemy in grammaticalization. *Language* 67, 475-509.
- Lyons (1977). *Semantics*. Cambridge: Cambridge University Press.
- Matthews, Steven and Virginia Yip (1994). *Cantonese, a comprehensive grammar*. London: Routledge.
- McCawley, James D. (1971). Tense and time reference in English. In Charles Fillmore and D. Terence Langendoen (eds.) *Studies in Linguistic Semantics*. New York: Holt, Rinehart and Winston, 96-113.

- Mithun, Marianne (1995). On the Relativity of Irreality. In Bybee and Fleischman (eds.), 367-88.
- Nedyalkov (1996). *Evenki*. London: Routledge.
- Nuyts, Jan (2001). *Epistemic modality, Language, and conceptualization: a cognitive-pragmatic perspective*. Amsterdam: Benjamins.
- Palmer, Frank R. (1990 [1979]). *Modality and the English modals*, 2nd edition. London: Longmans.
- Palmer, Frank R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.
- Palmer, Frank (1995). Negation and the Modals of Possibility and Necessity. In Bybee and Fleischman (eds.), 453-72.
- Palmer, Frank R. 2001. *Mood and Modality, second edition*. Cambridge: Cambridge University Press.
- Perkins, Michael R. (1983). *Modal expressions in English*. London: Frances Pinter.
- Rice, Keren D. (1987). *A Grammar of Slave*. Berlin: Mouton de Gruyter.
- Rice, Keren (2000). *Morpheme Order and Semantic Scope: Word Formation in the Athapaskan Verb*. Cambridge: Cambridge University Press.
- Roberts, John R. (1990). Modality in Amele and other Papuan languages. *Journal of Linguistics* 26, 363-401.
- Slobin, Dan I. and Ayhan Aksu (1982). Tense, aspect and modality in the use of the Turkish evidential. In Hopper (ed.), 185-200.
- Steele, Susan (1975). Past and Irrealis: Just what does it all mean? *International Journal of American Linguistics* 41, 200-17.

Sweetser, Eve E. 1990. From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure. Cambridge: Cambridge University Press.

Thompson, Sandra A. and Anthony Mulac (1991). A quantitative perspective on the grammaticization of epistemic parentheticals in English. In Traugott and Heine (eds.), 313-29.

Traugott, Elizabeth and Bernd Heine, eds. (1991). *Approaches to Grammaticalization*. Amsterdam: Benjamins.

Traugott, Elizabeth Closs and Richard B. Dasher (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Van der Auwera, Johan (1996). Modality: the Three-layered Square. *Journal of Semantics* 13:181-95.

Van der Auwera, Johan (2001). On the typology of negative modals. In Jack Hoeksema, Hotze Rullmann, Víctor Sánchez-Valencia and Ton van der Wouden (eds.) *Perspectives on Negation and Polarity Items*. Amsterdam: Benjamins, 23-48.

Van der Auwera, Johan; Vladimir Plungian. 1998. "Modality's semantic map". *Linguistic Typology* 2, 79-124.

Von Wright, E. H. (1951). *An essay in modal logic*. Amsterdam: North Holland.

Wagner, Johannes (1976). Eine kontrastive Analyse von Modalverben des Deutschen und Schwedischen. *International Review of Applied Linguistics in language Teaching* 14, 49-66.

Westney, Paul (1995). *Modals and periphrastics in English: an investigation into the semantic correspondence between certain English modal verbs and their periphrastic equivalents*. Tübingen: Niemeyer.

Willett, Thomas L. 1988. "A Cross-Linguistic Survey of the Grammaticization of Evidentiality". *Studies in Language* 12, 51-97.

Xrakovskij, V.S. and A.P. Volodin. 1986. *Semantika i tipologija imperativa*. Leningrad: Nauka.

Young, Robert and William Morgan (1987). *The Navajo Language: Grammar and Dictionary*. Albuquerque: University of New Mexico Press.

## ABBREVIATIONS

ABL	ablative	MIR	mirative
ACC	accusative	MP	modal particle
AG	agent	M.PROP	modal proprietive
ALL	allative	NEG	negation
AOR	aorist	NFUT	non-future
ART	article	NOM	nominative
CERT	certainty	OBLIG	obligation
COMP	complementizer	OPT	optative
DAT	dative	PAST	past tense
DS	different subject	PERF	perfect
DIR	direct evidential	PERM	permissive
FEM	feminine	PL	plural
FOC	focus	POSS	possibility
FUT	future	POT	potential
GEN	genitive	PRES	present tense
IMP	imperative	PROB	probability
IND	indicative mood	PRT	particle
INF	infinitive	REAL	realis
INFER	inferential evidential	SG	singular
INSTR	instrumental	SIM	simultaneous
IRR	irrealis	SS	same subject
MASC	masculine	SUBJ	subjunctive mood

## NOTES

<sup>1</sup> An excellent recent book on typology is Croft (2003).

<sup>2</sup> Many languages make a distinction between physical and mental ability. For instance, German uses the verb *kennen* for mental ability and *können* for physical ability.

<sup>3</sup> In recent years, the field of *corpus linguistics* has produced several important studies that deal with modal verbs, ranging from studies that deal with a single modal verb to those that deal with the entire modal system. English is the language best represented here, as should come as no surprise. An early study is Ehrman (1966), which probably has the honor of being the first corpus linguistic study overall. Other significant studies are Palmer (1979), revised as Palmer (1990) and Coates (1983) which are still cited. More recent studies are Westney (1995) and Krug (2001), the latter of which deals mostly with semi-modals. A study of modality as a whole is Perkins (1983). Other languages are lagging behind but are catching up. For German, there is Diewald (1999).

<sup>4</sup> This ambiguity has a diachronic origin as one meaning (epistemic) develops from the other (deontic). See Traugott (this volume) for discussion.

<sup>5</sup> In this paper I will follow the typological tradition of making a typographical distinction between language-specific categories (capitalized) and typological categories (not capitalized). An example is: the English Past Tense morpheme does not always indicate past tense.

<sup>6</sup> The best typological study on the imperative and related areas is Xrakovskij and Volodin (1986), in Russian.

<sup>7</sup> Modality can be marked on the verb as well.

<sup>8</sup> See also section 6 below for more discussion on modality and future.

<sup>9</sup> Maung has several verb classes, each with its own allomorphs. The above table is therefore not typical for all verbs.

<sup>10</sup> For a full discussion on the usefulness of semantic maps in typology see Croft (2003:133ff). An easy introduction is Haspelmath (2003). This technique has been applied to various categories. A small sample of (non-modal) studies is: the perfect (Anderson 1982), evidentiality (Anderson 1986), voice (Kemmer 1993), case (Croft 1991), coming and going (Lichtenberk 1991), and indefinite pronouns (Haspelmath 1997). In addition, semantic maps play a prominent role in Radical Construction Grammar (Croft 2001).

<sup>11</sup> Hansen (2001) takes these findings and applies them quite successfully to modality in the Slavic language family.

<sup>12</sup> There are several important studies on the use and development of the future in individual languages and language families. Palmer (1990) and Coates (1983) include sections on the modal verbs with future reference and Fleischman (1982) is an in-depth study on the future in Romance.

<sup>13</sup> Hetzron calls these two morphemes the Imperfect Definite and Imperfect Indefinite, respectively. The morpheme *-aGa* has only the one function exemplified in (26a) but *-e* has other meanings besides future possibility.

<sup>14</sup> The development of modals is beyond the scope of this paper, see Traugott (this volume) for extensive discussions.

<sup>15</sup> NEG-Raising refers to the process in which the negation associated with an embedded verb is moved (“raised”) to a verb higher in the sentence. An example from English is the sentence *I don’t think that he will come* in which the negation has been raised from the original *I think that he will not come*. See Horn (1989) for details of this process.

<sup>16</sup> Evidentiality currently enjoys an upsurge in popularity, with several important studies being released. An early study (in German) is Haarmann (1970). The classic in the field is Chafe and Nichols (1986) which is a collection of articles on evidentiality in all parts of the world. More recent books are Johanson and Utas (2000) and Aikhenvald and Dixon (2003), which are both collections of articles in the vein of Chafe and Nichols (1986). There is a special issue of the *Journal of Pragmatics* devoted to evidentiality (Dendale and Tasmowski (2001). An early typological study is Givón (1982). Willett (1988) is in need of updating but still useful. Other important articles are De Haan (1999, forthcoming a) which explore the limits of evidentiality and epistemic modality. De Haan (forthcoming b, c) are studies for the World Atlas of Language Structures (Haspelmath *et al.*, forthc.) on the existence of evidential morphemes in over 400 languages. Barnes (1984) is an article on the evidential system of the Tucanoan language Tuyuca which

---

continues to feature prominently in the field. In addition, there are many individual studies on “evidentiality in language X” in the literature, which do not always deal with evidentiality proper.

<sup>17</sup> In Palmer (2001) evidentiality and epistemic modality are themselves categories of the hypercategory of *propositional* modality.