Aspect, tense and modality: theory, typology, acquisition
Boland, J.H.G.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 6

TMA in English

6.1 Introduction

The TMA system of English has been the subject of an enormous number of studies. There is no other TMA system to which so many monographs, book chapters, and articles have been devoted. Yet it cannot be claimed that the definitive description has yet been achieved. The domain of TMA is very complex. Inevitably, this chapter cannot cover all the literature on this subject. The focus here will be on the role of scope in English TMA, and many other interesting topics will be ignored. The question is: is the Scope Hierarchy reflected in English?

In Chapter 5 hypotheses were derived on the basis of the Scope Hierarchy with regard to diachrony, frequency, synchronic configurations, expression form and expression order. In this chapter, these hypotheses will be explored to see if they make the correct predictions for English. As a first step, an inventory is made of the grammatical TMA expressions in English. Criteria are given as to which expressions should be considered grammatical and belong to the set of TMA operators. For each operator in English, the semantic functions and scope will be discussed (6.2). The inventory of TMA expressions serves as the basis to examine whether English indeed shows the predicted linguistic patterns.

The hypotheses derived in Chapter 5 are tested for English. Section 6.3 discusses the diachronic development of English TMA expressions and will test H1 (see 5.4.1). In section 6.4 the frequency of TMA expressions is examined. H2a (see 5.4.2) about token frequency is tested in corpus data of adult conversations and H2b (see 5.4.2) about the size of operator classes is tested on the basis of the inventory presented in 6.2. Not only is the Scope Hierarchy expected to be reflected in diachrony and frequency, the possible synchronic configurations of a TMA system are also supposed to be restricted by the Scope Hierarchy. Hypotheses H3a and H3b (see 5.4.3) will be tested in 6.5. The inventory of TMA expressions forms the basis for testing H3a and the semantic paths that result from the diachronic survey for H3b. In section 6.6 the English expression forms of TMA expressions will be examined to see whether they are in line with H4 (see 5.4.4). Finally, section 6.7 is dedicated to the expression
order of TMA expressions, which is expected to reflect scope as formulated in H5 (see 5.4.5).

6.2 INVENTORY OF TMA EXPRESSIONS

Which TMA functions are grammatically encoded in English? Before the configuration of TMA functions in English can be determined, the expressions to be considered grammatical have to be defined. This will be the subject of 6.2.1.

In 6.2.2-6.2.6 the specific expressions and functions within the different TMA domains are discussed. As a functional model of language should account for actual language use (in contrast to sentences made up by linguists), the inventory of TMA expressions is almost entirely based on corpus data. It relies heavily on earlier work on TMA in English, in particular the corpus-based studies on modality of Coates (1983) and Palmer (1990) and the general corpus-based grammar of English of Biber, Johansson, Leech, Conrad & Finegan (1999). Examples from these sources are referred to with C for (Coates 1983), P for (Palmer 1990) and B for (Biber et al. 1999), followed by the page number. Additional data were collected from the Collins’ Wordbanks Online English Corpus (titania.cobuild.collins.co.uk/form.html), abbreviated as CWO, and from the World Wide Web (indicated by the URL). Every attempt was made to use examples from conversational data, or else, from informal registers. Although the corpus examples formed the majority of data, a few examples had to be made up in order to illustrate a certain use or interpretation. Most of the literature and corpus data available are on British English. However, as the data in this chapter will later be compared to child American English (Chapters 8 and 9), some specific properties of American English will be discussed.

6.2.1 Criteria for grammaticality

Before the TMA system of English can be defined, it must be made clear what the limitations of this study are. Firstly, this study is restricted to the domains of aspect, tense, and modality and the related domains quantification, irrealis and evidentiality, such as defined in Chapters 3 and 4. This excludes, for example, imperative and negation and constructions like try to, happen to and dare to.

Secondly, this thesis is restricted to grammatical expressions, as opposed to lexical expressions. Section 5.2 showed that the boundaries between grammatical and lexical elements are not clear-cut. However, for this research, it has to be decided which items should be included and which excluded. Clear lexical items are adverbs, nouns and adjectives, such as for example the words probably, maybe, expectation, capacity, certain, and possible. Clear grammatical items
are the tense inflections and irregular forms for past and present tense. More
difficult constructions, however, are periphrastic constructions (periphrases),
with two verb-like elements, a finite verb and a non-finite verb, such as an
infinitive with or without to, a past participle (-ed) or a gerund (-ing). Examples
are must go, be able to see, appear to snow, stop running, be closed and have talked. Which
of these expressions are grammatical and which are lexical?

Olbertz (1996: 26) defined a grammatical periphrastic construction as
consisting of the inseparable combination of an auxiliarized finite verb with a
specific non-finite verb in which the finite verb agrees with the first argument
of the non-finite verb.1 There are two important aspects to this definition: the
finite verb is an auxiliarized verb, which means that syntactically it functions as
a modifier of the main, non-finite verb. When the construction is lexical, the
finite verb is the main verb and the non-finite verb functions as part of an
argument or adjunct of the main verb. The second important feature is that a
periphrasis functions as an inseparable construction: its meaning is not identical
to the interpretation of its components. Because of both properties, a
grammatical periphrasis shows different uses in language than a lexical
combination. This makes it possible to distinguish between the two.

Firstly, an important characteristic of a lexical verb (in contrast to an
auxiliary) is its potential to occur on its own, without a complement.
Constructions in which the finite verb with non-finite complement can be
replaced by the finite verb only (b-sentences), or by a finite verb and a nominal
complement (c-sentences) are therefore considered lexical (Olbertz 1996: 34).
Consider (1) and (2):

(1) a. I stopped working at eight.
   b. I stopped at eight.
   c. I stopped it.

(2) a. I was going to work at eight.
   b. *I was going at eight.
   c. *I was going it.

In (1), stop may occur on its own, without the complement working (1b) or with
a nominal complement it (1c). It is therefore a lexical element. In contrast, be
going to in (2) cannot occur without a complement (2b) or with the nominal
complement it (2c) without changing its meaning and it should therefore be
considered a grammatical construction. Many English verbs that express an
aspectual notion, for example, stop, start, begin, finish, continue, quit, and give up, can

1 This definition excludes causatives, since in those constructions the finite verb does not agree
with the first argument of the non-finite verb (cf. Olbertz 1996: 31). For example: I make
him go, He let her sleep, They got the car fixed.
occur on their own, as full lexical, intransitive verbs. They are excluded from this research.

A second reason to consider an expression lexical is if part of the construction can be modified by an adverb. If this is possible, it indicates that the syntagmatic position of the separate elements is not completely fixed and that at least the element that is modified has meaning of its own. This holds for example for the modal idioms (ha)d better and (would) rather, and also for be willing to and be able to. Consider the examples of natural speech (3)-(6):

(3) Oh, you’d much better go alone. (CWO)
(4) You ’d much rather play with shared shared memory or shared memory programming. (CWO)
(5) The farmers were only too willing to go back to a price system. (CWO)
(6) How well will they be able to use this? (CWO)

The modal expressions in (3)-(6) can be modified partly. They are considered lexical and will therefore be excluded from this research.

Thirdly, a specific type of construction that looks like a periphrasis, but is not, is a raising construction. In a raising construction the subject of an embedded clause is raised to the subject position of a matrix clause (Dik 1997b: 344-46). An example is Jack seems to work, which is an instance of raising related to the construction it seems that Jack is working. Jack, the subject of the embedded clause with the main predicate work, is raised to the subject position of the matrix clause with the main predicate seem. The relation between the two predicates seem and (to) work is therefore a matter of subordination and not of periphrasis, in which the relation is one of modification. Because of this criterion, the constructions seem, appear and turn out are considered lexical and they are excluded from this research.

Fourthly, passive constructions must be distinguished from periphrases (cf. Olbertz 1996: 29). An active counterpart and the possibility to add a by-phrase are clear indicators of passive constructions. This holds for expressions like be allowed to, be obliged to, or be expected to, as shown in (7). These expressions are therefore considered lexical.

---

2 A verb that is mainly used intransitively is much less likely to grammaticalize than a verb that is mainly or only used transitively. Only a transitive construction in which a verbal phrase functions as a complement could be reanalyzed as a periphrastic construction.

3 An expression is not necessarily lexical if the construction can be modified as a whole. For example, the construction be about to can only be modified in its entirety. In the sentence He is just about to go the adverb just does not modify about, but be about to.
TMA in English

(7)  a. I'm allowed (by him) to go.
    b. He allowed me to go.

An expression that looks like a passive, but is not, is be supposed to. It is impossible to combine it with a by-phrase (8a) or to make it active (8b).

(8)  a. He's supposed (*by them) to go
    b. *They supposed him to go.

Whether be supposed to should be considered grammatical is dependent on the final criterion, selection restrictions.

In addition to the formal criteria mentioned above, there is a fifth, semantic, criterion that is used here to distinguish lexical from grammatical constructions, i.e. selection restrictions. Since grammatical elements have general meanings, they will freely combine with different lexical elements, that is, they hardly impose selection restrictions on the elements with which they may co-occur, in contrast to lexical elements (Hengeveld 1992: 30f). However, since grammatical elements are not meaningless, they are also not compatible with every other lexeme of the language. For example, present tense inflection does not combine with a time adverb that indicates past time, an ingressive is rare with a punctual event and a modal marker expressing ability will in general not combine with an inanimate subject. On the basis of their general meaning, it is thus predictable for grammatical elements what kind of selection restrictions occur, again in contrast to the selection restrictions of lexical elements.

Which selection restrictions may be expected within the different domains? Firstly, aspect operators select part of the internal temporal structure of an event to focus on. Since the internal temporal structure of stative and punctual events is rather undefined compared to other event types, it may be expected that not all aspect operators combine freely with these types of events, or that, if they do, it results in a specific interpretation. On the contrary, it is not expected that an aspectual operator imposes selection restrictions on the first argument, since the first argument does not interfere with the aspectual semantics. The same holds for whether the predicate designates a controlled or a non-controlled event. In a controlled event the first argument has the power to determine whether or not the event will obtain, such as in John opened the door or John was sitting in the garden, versus the non-controlled events in the substance was red or the tree fell down (Dik 1997a: 112). There is no reason to assume that aspectual operators interfere with the semantic feature of control. In conclusion, only aspectual expressions that can combine with both animate and inanimate first arguments, and occur in controlled and non-controlled events are considered grammatical. The aspectual expressions that apply to this
criterion are be –ing, be going to, keep –ing, be about to, have –ed and be –ed. Not all of these expressions combine freely with every event type—*be knowing, *be jumped—but they do not show unexpected selection restrictions on the arguments or on the feature controlled /non-controlled events. Both with regard to their formal behavior and their semantic behavior, these expressions can be considered grammatical and they are included in this research. Most other aspectual verbs in English (stop, start, finish, continue, begin) were already excluded based on their independent main verb use, but moreover, it holds for most of these verbs (except for begin) that they also apply selection restrictions on their infinitival complement or first argument.

Secondly, expressions of event quantification may be expected to impose selection restrictions on the event type with which they occur as the event type has to be compatible with a notion of ‘repetition’, which is not the case for permanent states or for specific telic events (?be died repeatedly). Furthermore, the notion of habituality seems in principal to be restricted to animate arguments: only animate arguments can be involved in habitually repeated events. Other expressions of event quantification are not expected to impose selection restrictions on their arguments. In English, there are three expressions of habituality, would, will, and used to: these expressions have grammaticalized to such an extent that they are compatible with all kinds of event-types and with both animate and inanimate arguments (see 6.2.4.1 for examples and further discussion of the habitual in English). The construction keep –ing, which is later explained to be an expression for continuative aspect and for frequentative event quantification (see 6.2.2 and 6.2.4.2), cannot combine with every event-type (the event has to be repeatable), but it does not impose any selection restriction on the arguments. All these constructions can therefore be considered grammatical.

Tense operators and operators of irrealis (hypothetical and counterfactual) are not expected to impose any selection restrictions on the predicate or argument since their semantics are compatible with any type of event or argument. This is in fact the case for the past and present tense inflection and for the future tense markers will/shall. It also holds for the hypothetical constructions ‘past modal + V^c’ and ‘past tense form’ and the counterfactual constructions ‘past modal + have –ed’ and ‘past perfect form’.

For expressions of modality, the matter of selection restrictions is more complicated. Event-oriented or proposition-oriented modality should neither impose selection restrictions on the predicate nor on the arguments. However, participant-oriented modality, in particular modality that ascribes the source of the modality to participant-internal characteristics, is expected to impose certain restrictions on the arguments. The notions of (weak) obligation, desirability, permission, ability, internal need (x is forced because of internal characteristics)
and volition (will) are restricted to animate beings and, except for the latter two, restricted to controlled events. Most modal expressions in English have event-oriented or proposition-oriented uses besides their participant-oriented uses and they are clearly grammaticalized. This holds for all the central modals—must, will, would, shall, should, can, could, may, might—and the constructions have to, have got to, got to, need (to), ought (to) and be supposed to. In their participant-oriented use, these constructions do impose selection restrictions, as is to be expected, but they have also much more general uses, characteristic for wider scope operators. All these expressions are therefore considered real grammatical markers. The so-called marginal modal dare is excluded as it has a questionable modal meaning. Moreover, it is very infrequent and its use is mainly limited to fixed expressions, which are both characteristics of lexical items.4 Both the constructions dare (to) and try to can only combine with animate first arguments and controlled events: their meaning is still very specific.

A construction that needs some closer inspection is want to. The modal use of this construction is confined to expressing will or volition. It combines only with animate first arguments, although it can be combined with controlled and non-controlled events. The question is whether want to is a lexical predicate with a clausal argument or a modal operator (still) restricted to participant-oriented modality. Krug (2000) has carried out an in-depth analysis of the historical development of want and want to in English. He gives several arguments why want to, or at least the contracted construction wanna V, has grammatical status in (American) English. First, the frequency of the verb want has tremendously increased from Middle English to present day American English (Krug 2000: 119), in particular the string want to is much more frequently used (it makes up about 60% of all instances of want in present day written American English) (p.134). In spoken language the frequency will probably be even much higher, which is also the case in spoken British English compared to written British English (p.136). Second, the string want to is often phonologically reduced to wanna, particularly in American English and by younger speakers (up to 45%) (p.161). The phonological reduction indicates that speakers analyze the construction no longer as a two-verb phrase [want] + [to V], in which want is a lexical verb, but rather as a one-verb phrase [wanna [V]], in which wanna functions as an operator. An indication that speakers apply the latter analysis is that ellipsis of to V, as in you can wait outside if you want is less common than ellipsis of only the verb: …if she wants to, she can get in touch with you. (p.156). A third argument that want to is grammaticalizing is its syntactic behavior, which in certain respects resembles more the syntactic behavior of the central modals

---

4 Martine Taeymans, personal communication.
Table 6-1. English TMA expressions included in this research

<table>
<thead>
<tr>
<th></th>
<th>–ed /irregular</th>
<th>be going to</th>
<th>might</th>
<th>will</th>
</tr>
</thead>
<tbody>
<tr>
<td>–ø /-s</td>
<td>be supposed to</td>
<td>must</td>
<td>would</td>
<td></td>
</tr>
<tr>
<td>be –ed</td>
<td>can</td>
<td>need (to)</td>
<td>wanna</td>
<td></td>
</tr>
<tr>
<td>be –ing</td>
<td>could</td>
<td>ought (to)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>have –ed</td>
<td>have to</td>
<td>shall</td>
<td>and possible</td>
<td></td>
</tr>
<tr>
<td>keep –ing</td>
<td>(have) got to</td>
<td>should</td>
<td>combinations</td>
<td></td>
</tr>
<tr>
<td>be about to</td>
<td>may</td>
<td>used to</td>
<td>thereof</td>
<td></td>
</tr>
</tbody>
</table>

(can, could, may, must, etcetera) than that of lexical verbs with a similar token frequency as want to. It appears that ‘full verbs in general combine far more often with modals than the emerging modal want to, the difference being even more drastic when this occurs in its contracted form.’ (p.158). Furthermore, wanna is sometimes used with third person, i.e., without a third person inflection –s (p.160). Finally, the past tense of want to can be used as a polite form, such as in I wanted to ask you a question, where it does not indicate the pastness of want (Bybee 1995). This is similar to the past forms of the central modals that also yield a polite reading, whereas past tenses of lexical verbs do not get this interpretation (Krug 2000: 155). To conclude, Krug suggests that we are possibly ‘witnessing the very initial stages of a functional split between modal use and nonmodal use, very similar in type to historical changes leading to modalization.’ (p.153). In this thesis, all instances of contracted wanna are analyzed as grammatical while all other instances of want (to) are considered lexical.

In sum, Table 6-1 presents all the TMA expressions in English that are considered grammatical and that are included in this research. In the next sections, the semantic functions of these expressions will be discussed. Corpus examples will illustrate the functions of each expression form. As the definitions of specific TMA categories were already discussed in Chapters 3 and 4, the discussion here will be limited to the characteristics specific for English. The final inventory of the TMA expressions in English and the functions for which they are used serves as the basis for testing the hypotheses.

6.2.2 Aspect
Which aspecual functions are grammatically encoded in English? The constructions be –ing, keep –ing, be going to, be about to, have –ed and be –ed are used
to express progressive, continuative, prospective, immediate prospective, perfect and resultative aspect (see 3.3.3).

The construction *be* –*ing* expresses progressive aspect. It limits the predicated property of the participant to the inner part of an event and can be used in contexts of past time reference (9), present time reference (10) or future time reference (11):

(9) That’s why I was thinking I might hang on to the Volvo. (B460)
(10) What’s she doing? (B470)
(11) But she’s coming back tomorrow (B471)

A continuative selects the initial boundary and the internal part of the event, but not the terminal boundary. In English the construction *keep* –*ing* has continuative as one of its meanings. Especially in combination with atelic events, this construction is interpreted as a continuative. *Keep* –*ing* in combination with telic events may encode either a continuative or a frequentative meaning. The latter meaning will be discussed in 6.2.4.2. Clear examples with continuative interpretations are presented in (12) and (13):

(12) I don’t want to keep living with my mum. (CWO)
(13) She kept thinking while she was smiling. (CWO)

The construction *be going to* marks prospective aspect. It limits the predicated property to the pre-state of an event. See (14)-(17):

(14) Well I was going to say just about the same as that. (CWO)
(15) I was going to be called Kate if I was a girl. (B456)
(16) And he’s going to see it. (B456)
(17) When are you going to see your parents? (P145)

*Be going to* is by some authors considered a future tense marker in English, whereas I have classified it as aspect. This issue will be discussed below.

The expression *be about to* marks immediate prospective. Like *be going to* it selects the pre-state of an event, but it also indicates that the pre-state is very close in time to the initial boundary of the event. See (18) and (19):

(18) Now my colleagues were about to throw that one out as irrelevant. (CWO)
(19) This was about to melt. (CWO)
The expression *have* –*ed* limits the predicated property of the participant to the post-state of the event. This construction marks perfect aspect, see (20)-(23).

(20) We’ve been to a lot of seminars too. (B464)
(21) He’s gone home. (B465)
(22) They’ve done so much. (B465)
(23) We’ve had it since last January. (B468)

The perfect is mainly restricted to present time reference, especially in conversations. Except for *had been*, past perfect forms are very rare in conversation (Biber et al. 1999: 468). The perfect in English selects a post-state, but does not indicate by itself whether the event and the post-state of this event are close in time (*I have just arrived*) or not (*I have been in Paris once*) and it also does not specify the length of the post-state or of the event leading up to the post-state (*I’ve lived here all my life*). This is a matter of contextual interpretation.

The final expression in English that marks aspect is *be* –*ed* for resultative aspect. Resultative indicates that an ‘action in the past produces a state that persists into the present.’ (Bybee et al. 1994: 318). Like a perfect, a resultative selects a post-state, though the focus on the post-state seems to be stronger and the implication of a preceding event that led up to the post-state seems to be much weaker. Examples are presented in (24)-(26):

(24) The door was closed.
(25) He is gone.
(26) The stick is broken.

The resultative construction and the passive are equal in form in English, but only the passive construction can be combined with a *by*-phrase. If there is no *by*-phrase, then the two constructions can be distinguished using an adverb of time: with a resultative construction, the time adverb refers to the time of the resultant state, whereas with a passive, the time adverb refers to the time of the event leading up to the result (Nedjalkov & Jaxontov 1988: 49). This can be shown by an example from Jespersen (1924: 247) in (27):

(27) When I came at five, the door *was shut*, but I do not know when it *was shut*.

The first *was shut* is a resultative construction; at five, the door was in the state of being shut. The second *was shut* on the other hand is a passive construction: *when* refers to the time of the action of shutting the door.
Notice that all aspectual forms in English are in fact combined tense-aspect markers. *Be, keep* and *have* bear the tense marking, whereas the non-finite verb form expresses the aspectual component. The auxiliaries *be* and *have* are in spoken English mostly contracted to, for example, *'m* or *'s* or to *'ve* or *'s*.

The classification of *be going to* as prospective aspect is not uncontroversial. Biber et al. (1999) consider *be going to* as one of the possible markers for future time in English. However, in using this expression, the focus is on a current activity or state or on a present decision or intention leading to a future event. According to Bohnemeyer (personal communication) prospectives crosslinguistically are associated frequently with the expression of intentions, since intention is one of the many possible pre-states of an event: intention is, however, not entailed by the prospective. Intentional readings in English admit the specification of the time of the event, so that the argument can be described as being in the pre-state of 'doing something at a specific time'. The temporal specification of the event is thus inside the scope of the aspectual operator and there is no entailment of realization of the event at that time, contrary to a real future tense. It is this possibility of specifying the event-time that is responsible for the tense-like appearance of *be going to*. Quirk et al. (1985: 214) show that *will* in the main clause of a conditional can in most cases not be replaced by *be going to*, which proves that the prospective is not a future tense marker:

(28) If you leave now, you'll never regret it.
(29) *If you leave now, you are never going to regret it.

The classification of *have –ed* as perfect aspect is not uncontroversial either. For example, Elsness (1997) claims that this construction marks indefinite past tense. In my view, however, the indefiniteness of the temporal location of the event is the result of the fact that it is not the event that is located in time, but the post-state of the event. It is an implicature that the event occurred prior to the post-state, but its temporal location is not specified and therefore indefinite.

In sum, both the prospective and the perfect in English have temporal sequentiality as an implicature, whereas future and past tense encode temporal sequentiality. When the notion of current relevance were in most cases not required for the use of *be going to* or *have –ed*, then the constructions would express the temporal sequentiality and they would have grammaticalized into tense-markers. But since in English the notion of current relevance is obligatory for both constructions, the encoded meaning is here considered aspectual and not temporal.
6.2.3 Tense

In most contexts, tense is obligatorily marked on the main verb in English. Only imperative and non-finite clauses are not marked for tense. The inflection on the finite verb and the auxiliary will distinguish between past, present and future tense. This will be further discussed below. The relation between tense and modality and between tense and irrealis will be discussed in 6.2.5 and 6.2.6.

6.2.3.1 Past

There are several expressions in English that mark past tense. The so-called simple past or preterit is the suffix \(-ed\) on the main verb or irregular verb forms. Examples are presented in (30)-(35):

(30) I saw him yesterday. (B467)
(31) She just shrugged her shoulders. (B460)
(32) Well I rang them up yesterday. (B453)
(33) Hey, did you read through this yet? (B463)
(34) We already gave him a down payment. (B463)
(35) Then they said: ‘Well,’ and then I realized that it was Fennite. (B467)

In non-verbal predicates the past form of the copula \(be\) marks past tense and in periphrastic constructions, the past forms of the auxiliaries \(be\), \(have\) or \(keep\) or of modal auxiliaries mark past tense. Past tense indicates that the relevant part of the event occurred anterior to the moment of speech (ST). It is used to refer to events in the immediate and in the remote past. The simple past is also used in contexts when a past event has led to a present resultant state, especially in American English conversation (see (33)-(34)) (Biber et al. 1999: 463).

The simple past in English makes no explicit selection of the event structure, i.e., it is aspectually neutral and the exact interpretation depends on the event type with which it occurs. In combination with temporary events, past tense leads in general, but not necessarily, to a perfective interpretation: the entire temporal structure is ascribed to the argument(s) and located anterior to ST. In combination with permanent states, the past tense only locates the part that is relevant to the conversation in time with respect to ST (see 3.4.3).

6.2.3.2 Present

The morphologically unmarked (zero) verb form in English, except for the third person singular that has a suffix \(-s\), indicates present tense\(^5\). The general

---

\(^5\) Strictly, the present tense in English is a non-past tense as in certain contexts it may be used to refer to future events. However, since future reference of the present tense is very limited I will speak of present tense.
meaning of the present tense is that the relevant part of the event overlaps the moment of speech. It is claimed, among others by Comrie (1976: 66), that the present tense is essentially imperfective. But is this indeed the case? In the proposed account it is assumed that the simple present, like the simple past in English, does not make a selection of the temporal structure of the property or relation. In other words, the simple present is aspectually neutral. Which part of the event is located in time is dependent on the topic time and the event structure. As discussed in 3.4.2 topic time is the time interval to which an assertion is confined. It determines which part of the state of affairs is relevant to the discourse, and it is only this part of the event that is located in time. The interpretation of the simple present therefore depends on the relation between topic time, event type and speech time. What are the possible interpretations of the simple present in English?

First, present tense can describe a state existing at the speech time (Biber et al. 1999: 452-54), as in (36)-(38):

(36) I want a packet of crisps. (B453)
(37) I know her very well.
(38) I think you might be wrong. (B453)

In this interpretation, the topic time interval overlaps speech time and the interval of the event overlaps topic time. In Figure 6-1, the relation between topic time (TT), speech time (ST) and event structure (solid line on the time line) is depicted. As states are conceived of as having no temporal boundaries, the event is represented as a solid line without an initial or terminal boundary.

![Figure 6-1. Representation of simple present in combination with a stative event in English](image)

The present forms in (36)-(38) make no selection of the temporal structure, but language users know that only the part of the state that is relevant to the discourse is located in time. The topic time interval selects the relevant part of the state and as this part overlaps ST, the present tense is used. It is a matter of
pragmatic knowledge that not the complete state is located in time, for indicating the time location of an entire state would be virtually impossible or at least pragmatically very inconvenient. Logically, it holds before, at and after the moment of speaking, but to continuously encode this fact—*I knew, know and will know her very well*—is communicatively irrelevant. Language users infer that the state may extend past the boundaries of the speech time and even the topic time interval. There is no reason to assume that the present tense marks imperfective aspect in this use.

A second use of the present tense in English is to describe present habitual behavior (39) or generic statements (40)-(41):

(39) She’s vegetarian but she eats chicken. (B453)
(40) When the prey is dead, the tiger drags it to a safe place and eats it.⁶
(41) Tigers eat deer, wild pigs and wild cattle.⁷

These readings generally arise with dynamic events. The time interval for which the speaker wants to assert something is in these cases much larger than ST. Since most events do not overlap the entire topic time the interpretation is that the state of affairs is repeated, such that it fills the topic time. The interpretation may be either habitual / frequentative or distributive which depends on the number and specificity of the argument(s). For example, (39) is interpreted as habitual, since a specific individual is involved; (40)-(41) are interpreted as distributive, since the first argument is generic. The state of affairs is ‘repeated’ because it holds for several individuals. Consider Figure 6-2 for a representation of the habitual or distributive⁸ interpretation of the present tense:

![Figure 6-2. Representation of habitual or distributive interpretation of simple present in English](https://www.kidcyber.com.au/topics/tigers.htm)

* For a proper representation of a distributive a three dimensional picture is necessary, since several events occur simultaneously or overlap partly.
In this interpretation the boundaries of the temporal structure of one event are included in topic time, but at the same time, the boundaries of the repeated event-sequence are not included in topic time. Note that it is not necessarily the case that one of the repeated events actually takes place at the same time interval as ST. The resulting interpretation is that the situation of the repeated events holds at an interval that overlaps the speech time interval. A habitual or distributive interpretation therefore resembles a permanent state. The relation between topic time and event structure is very similar to the one in Figure 6-2 and it does indeed yield a ‘stative’ interpretation: it is a general statement, a situation that holds at all times. The topic time determines that the sequence of events that overlaps ST is relevant to discourse, hence the present tense. The difference with a real state is that in a habitual or distributive reading there are time intervals at which the state of affairs is not actually taking place.

A third interpretation of the present is a report on an ongoing action. Examples are presented in (42)-(44):

(42) Here comes your mother. (B454)
(43) Situation: talking about a toddler.
   Oh, my goodness. There he goes. Look at him walk. (B454)
(44) Spreads it wide. (football commentary, from Mackenzie (2005))

In this interpretation the topic time is more or less confined to ST. This interpretation is possible if the interval of the state of affairs indeed more or less coincides with ST and this is only the case for momentaneous events. The effect is a reportative interpretation, very common in sports reports. In Figure 6-3 the reportative interpretation is represented:

Figure 6-3. Representation of reportative interpretation of simple present in English

A fourth possible interpretation of the simple present in English is future reference in cases of scheduled events, such as in the train leaves at five tomorrow but also in temporal adverbial phrases such as when he comes here, I’ll tell him (from Comrie (1976: 68)). In both structures the present tense cannot be interpreted as habitual. Why is the simple present used in these cases? I follow Klein’s
(subm.) analysis that the topic time stretches from the present into the future and includes the time of the complete temporal structure. Apparently, the speaker wants to assert something about a time interval for which the state of affairs does not (yet) hold up to the interval for which the event holds in its entirety. The complete temporal structure of the event including the preceding temporal span is relevant to discourse and located on the time axis. Since this interval overlaps speech time, it is possible to use a present tense. The resultant reading has a sense of a scheduled or planned event. See Figure 6-4 for a representation.

![Figure 6-4. Representation of scheduled event interpretation of simple present in English](image)

The counterpart of the scheduled event interpretation is that the topic time extends into the past. The speaker wants to assert something for a time span that stretches from some point in the past up till the moment of speech. This may account for the use of the simple present for narrating past events, the so-called historic present (also Klein’s analysis, subm.). The simple present occurs in narratives to create a more vivid style. It is especially common with speech-act verbs. See (45):

(45) No. He says, are you going home tonight? He thought I was going home to my parents. (B455)

The possible interpretations of the present tense do not support the view that the present tense in English is imperfective. In combination with a punctual event, a reportative interpretation is most plausible, which is similar to a perfective view. The most common use, however, is that the topic time extends ST and that the state of affairs is either a state or a repeated dynamic event. In these uses the simple present has a sense of imperfective aspect, since only a part of the event is located in time. However, in the case of habitual and

---

9 This is not identical to a pre-state, since the event is not in any way the consequence or purpose of the preceding time interval.
distributive, it also has a sense of perfective aspect, since the temporal structure is in its entirety ascribed to the argument(s). Note that the simple present does not mark one of the above interpretations: it only marks present tense. The different aspectual interpretations result from implicatures or inferences. I have shown that the relation between topic time, the event structure and speech time result in different interpretations of the simple present.

The difference between the possible interpretations of simple tenses in English and an imperfective and perfective marker is that the latter explicitly select a subinterval or the entire temporal structure, whereas with simple tenses, no selection is explicitly encoded. And exactly because there is no specific selection, different interpretations are possible, depending on the event type.

6.2.3.3 Future

English has a specific marker for future tense, which indicates that the relevant part of the event takes place posterior to the moment of speech. It is expressed by will or by shall for first person. In the case of indirect speech with the main clause marked for past tense, future tense is expressed by would or should (for first person). Consider (46)-(49):

(46) Well, I’ll be back tomorrow. (C179)
(47) It will be lovely to see you. (C179)
(48) My babe-in-arms will be fifty-nine on my eighty-ninth birthday … the year two thousand and fifteen when I shall be ninety. (P137)
(49) The judge in the mail train robbery # eh trial # said today that it was unlikely that the jury would be able to retire to consider their verdict until late next Tuesday. (C210)

Will and would can also be used to indicate volition or prediction. These meanings will be discussed in 6.2.5.2.1 and 6.2.5.2.3. Will and would are furthermore used for the expression of habituality (6.2.4) and would for the expression of irrealis (6.2.6).

6.2.4 Event quantification

6.2.4.1 Habitual

English has three markers that specifically encode habituality: used to and would are past habitual markers and will is a present habitual marker, in one of its uses. Note that the present tense can yield a habitual interpretation, but it is not a specific marker of habituality. In 6.2.1 it was discussed that habituans are in principle expected to impose selection restrictions on their first argument and
the predicate, for a habit can only be ascribed to entities that can have habits (*Houses have the habit of being big) and habits are only properties or relations that are controlled (*He has the habit of being tall). Are there still selection restrictions of this kind to the expression of used to? See (50)-(53) for some possible uses of used to:

(50) I used to read a lot. (CWO)
(51) Everyone used to call her smelly. (CWO)
(52) The flags used to be in the colors of the “Home” Club.\(^\text{10}\)
(53) There used to be a playground here.\(^\text{11}\)

The examples show that there are no selection restrictions on the arguments or predicates with which used to can combine: it combines with animate and inanimate subjects and with controlled and uncontrolled properties. It may therefore be concluded that used to is highly grammaticalized. The specific ‘habitual’ sense of used to in English is marginalized and the expression has become closer to a more general frequentative. It is probably a matter of world knowledge or inference that if a specific animate entity is involved in a repeated and controlled event, the most likely interpretation is that the repetition of the event is due to the habit or propensity of this entity. The representation of a past habitual in combination with an event with boundaries is presented in Figure 6-5 (see also 3.5.3):

![Figure 6-5](image)

**Figure 6-5.** Representation of a past habitual in combination with an event with boundaries in English

The examples in (52) and (53), however, show that used to can also combine with permanent states, such as be in the colors of the “Home Club” and be here. In Klein’s analysis (1994: 47) of a habitual a speaker chooses to speak about a series of topic times when using a habitual marker, and for all of these topic

\(^{10}\) [www.kenaston.org/KenAston/KenAston48.htm](www.kenaston.org/KenAston/KenAston48.htm)

\(^{11}\) [www.meliorism.com/coh/sucks/8.htm](www.meliorism.com/coh/sucks/8.htm)
times, the event holds. It is thus not necessarily the case that the state of affairs indeed recurs, but the time interval relevant to the discourse is a sequence of intervals in the past. When the event is a (semi-)permanent state, such as in (52) and (53), the state of affairs does not recur, but the state holds at all the topic time intervals. This is represented in Figure 6-6, in which the solid line indicates a state, that apparently ended before ST. Note that the event holds at all times in the past, but since the topic time intervals determine which parts of the state are relevant to the discourse, it creates a sense of repetition. I assume that this use of a habitual expression is later in diachronic development and inferred from the interpretation in Figure 6-5. Only a generalized habitual expression may combine with states.

![Figure 6-6. Representation of a past habitual in combination with a state with a terminal boundary in English](image)

In English, it is not necessarily the case that the state of affairs is actually recurring when a habitual expression is used: in that case, the time interval relevant to the discourse is a sequence of intervals.

Would and will also freely combine with different types of predicates and arguments. These expressions are, however, less frequently used than used to. Examples of would for expressing habitual or frequentative are presented in (54)-(55):

(54) His blue eyes gazing seriously through a wisp of fair hair which would keep falling across his eyes. (C207)
(55) The thing was he would # he would concoct anecdotes and he’d tell them to me over and over again you know obviously not realising that he’d told them to me before. (C209)

Finally, will can be also used with a sense of habituality, or typical behavior (56):

(56) So one kid will say to another, one kid will make a suggestion to another, he’ll say the moon's further away from the earth than the sun. (P136)
6.2.4.2 Frequentative

A second subcategory of quantification, the frequentative, is in English expressed by *keep –ing*. A certain event is repeated several times, not habitually and not necessarily on one occasion. The frequentative reading may arise when the auxiliary is combined with a telic event, as in (57)-(60):

(57) I keep saying to him Oh hang on a bit don’t go to XXX. (CWO)
(58) Gizzmo keeps trying to persuade me to go with her. (B741)
(59) She kept running out of the room. (B364)
(60) He kept bringing it up so in the end I said … (CWO)

In section 6.2.2 it was already shown that *keep –ing* is used as a marker of continuative aspect (\(\pi_1\)). This is the most common interpretation when the construction is combined with a state, compare the examples earlier given in (12)-(13). When the construction is combined with a dynamic event, however, it is often ambiguous as to which interpretation is intended. The two interpretations of *keep –ing* can be distinguished by adding an adverb of specified duration, such as *for an hour*, which is preferred with the continuative interpretation, but dubious with the frequentative reading. The frequentative, on the contrary, can be easily combined with an adverbial phrase of repetition, such as *over and over again*. Consider the following examples:

(61) He kept running for an hour /?over and over again.
(62) He kept bringing it up ?for an hour / over and over again.

*Keep running* in (61) cannot be interpreted as a frequentative, but only as a continuing action. *Kept bringing it up* in (62) is not completely impossible with *for an hour*, but then the repeated activity is continuous and uninterrupted. In particular momentaneous or punctual events in combination with *keep –ing* are ambiguous between a continuative or a frequentative reading. They allow both interpretations. Consider (63) and (64):

(63) His brake lights kept flashing on for an hour / over and over again. 
(modified example from B746, *keep > kept*)
(64) I kept doing garlic burps for an hour / over and over again. (modified example from B361, *keep > kept*)

The expression *keep –ing* can thus express a continuative aspect, in which case it selects part of the temporal structure to ascribe to the argument(s). It can also express a frequentative, in which case it indicates that the state of affairs was repeated, not necessarily on one occasion or with regular intervals. In that
case it indicates that the speaker does not refer to one event in the real or imaginary world, but to a range of events. The interpretation is dependent on the situation type: telic events may be interpreted as frequentative, if the intervals between the repetitions are quite long.

6.2.5 Modality

English has an elaborate modal system. As a consequence, this section is fairly long, even though the discussion is restricted to the most important aspects. This section is structured according to the subcategorization for modality discussed in 4.2. There, it was shown that the domain of modality can be described by the parameters of sense, scope and source. The different functions of the modal expressions in English are primarily discussed on the basis of their basic sense: potentiality, disposition, weak necessity or necessity. For each sense, it will be discussed which specific meanings arise according to the scope of the modal expression. If the scope is the predicate (participant-oriented modality, \( \pi_1 \)), the modal marker specifies the relation between the predicate and the argument. If the scope is the predication (event-oriented modality, \( \pi_2 \)), the actuality of the event is expressed. If the scope is the proposition (proposition-oriented modality, \( \pi_3 \)), the commitment or attitude of the speaker to the propositional content is expressed. Within each scope type, examples of different sources of modality are presented. For participant-oriented modality, the source of the modality can lie in participant-internal characteristics or in external circumstances that are either non-deontic or deontic. For event-oriented modality, the source of the modality can lie in general or deontic external circumstances, or in general (objective) knowledge about events. For proposition-oriented modality, the source can lie in the (subjective) knowledge of the speaker or in the attitude of the speaker (the speaker as a deontic source). For each specific semantic function, examples will be presented. If the example contains different modal expressions, then the modal expression(s) that illustrates the semantic function is underlined. In general, English has a few expressions for the same functions, but the dominant contexts of use differ for each expression. Distributional properties of each expression are not discussed in detail.

A general point to discuss in advance is the relation between tense and modality. The past modal forms, could, might, would and should, started out as past tense forms of can, may, will and shall. In non-epistemic contexts, these forms can still express past tense, but it is also possible that they do not refer to past time anymore, but rather express a more tentative (and therefore more polite) sense than their present modal counterpart or indicate irrealis in combination with the semantics of their present modal counterpart. The same development is taking place in the past tense forms of the newer modals, had to, wanted to,
et cetera. They are still used to express past time reference beside their modal meaning, but they are also used to express irrealis or tentativeness. In epistemic and irrealis contexts, modal markers are tenseless. Past modal forms in general carry the meanings of the present forms, but they may have developed uses that the present form lacks. For example, *should* is used epistemically, whereas *shall* is not, and *could* is used freely as an uncertainty marker whereas *can* has a very restricted use as a marker of certainty (see 6.2.5.1.3).

6.2.5.1 Potentiality
The first sense to be discussed is potentiality, ‘not precluded from’. The English forms used to express this sense are *can* / *could* and *may* / *might*. The combination of potentiality with the parameter scope, results in three subcategories, which will be discussed in the following sections. The application of the parameter of source leads up to a further refinement of these modal subcategories.

6.2.5.1.1 Participant-oriented potentiality
When the scope of the forms *can*, *could*, *may* and *might* is the predicate, the resultant meaning is that the argument x is not precluded from PREDing. The reason why x is not precluded from PREDing, may be because of participant-internal characteristics, which raises the specific meaning of ability. *Can* and *could* are used to express ability. Core examples concern animate subjects with dynamic verbs, (65) and (66), but inanimate subjects (67), and combinations with cognition (68) and perception verbs (69) can also be interpreted as expressing ability.

(65) I can only type very slowly as I am quite a beginner. (C92)
(66) They can’t speak a word of English, of course, not a word, but, you know, they can say what they like. (P85)
(67) The plane has a built in stereo tape recorder which can play for the whole four hours it will take to fly to Majorca. (C92)
(68) They asked me and I just couldn’t refuse. (B493)
(69) I couldn’t feel my hand. (B493)

A second source for x not being precluded from PREDing may lie in participant-external non-deontic circumstances. This creates the meaning of root-possibility: x is not precluded from PREDing (mainly) because of external circumstances. *Can* and *could*, and less frequently *may* and *might* as more formal variants, encode root possibility. Consider (70)-(73):

(70) You can always say it’s just not your style. (P85)
(71) Can you pick your own trousers up? (C94)
(72) I could have gone straight there but I just couldn’t get there. (C121)
(73) I am afraid this is the bank’s final word. I tell you this so that you may make arrangements elsewhere if you are able to. (C141)

The third possible source of participant-oriented potentiality are participant-external, deontic circumstances. This leads to the specific meaning of permission: x is not precluded from doing something because of a deontic source. Permission is expressed by all four potentiality markers, although can and could are most frequent. See (74)-(77):

(74) Can I pinch a ciggie? – Course you can. Would you like a menthol or a plain? (P71)
(75) Could I go just briefly to describe the other two areas? (C118)
(76) If you want to recall the doctor, you may do so. (P71)
(77) May one taste? You said I might. (C156)

6.2.5.1.2 Event-oriented potentiality

When the scope of potentiality markers is the predication, the potential occurrence of an event is expressed: the event e is not precluded from occurring. The first possible source for event-oriented potentiality are general external circumstances. This category is distinct from the participant-oriented counterpart (π1) in that it concerns a non-specific or generic participant or the participant is not mentioned at all, such as in bare passives. The potentiality does not apply to a specific argument; it applies to the entire event, independently of the involved arguments. It is mainly expressed by can, but may and might can also express this meaning, as in (78)-(82). No examples were found of could in this interpretation.

(78) I mean, you can travel from Belgium to France with much less palaver than you can travel from the North to the South of Ireland. (P84)
(79) National pressure groups cannot exist without full time staffs and a regular income. (P91)
(80) Well, I’ll see what can be done and give you a ring. (P84)
(81) … and will examine ways in which this may be more effectively safeguarded. (C144)
(82) In 1814 the completion of the Mons-Conde canal increased the ease with which Mons Coals might be sent to Nord. (C156)

The second possible source for event-oriented potentiality are external circumstances that are deontic. Deontic modality with scope over the
predication is distinct from deontic modality with scope over the predicate in that the deontic source is not an individual, but a general law or rule of conduct, that is independent of the participants involved. *Can*, *could*, *may* and *might* express that e is not precluded from occurring because of external circumstances, more specifically a general deontic source. Deontic event-oriented modality mainly combines with non-specific or generic participants, or the participants are not mentioned at all, as in bare passives. Consider (83)-(85):

(83) In the library you *can* take a book out and keep it out for a whole year unless it is recalled. (P103)
(84) Men cannot enter.
(85) No vehicle *may* be left in the University grounds during vacation. (C132)

The general law can also apply to specific participants, which was discussed in section 4.2.3.2.

The third possible source of event-oriented potentiality is general knowledge (objective epistemic modality). Within the domain of potentiality *can* or *could* and *may* or *might* express epistemic possibility: the occurrence of e is not precluded because it is logically not excluded. It is objectively possible that e occurs or that e is the case. Examples of epistemic possibility are presented in (86)-(90):

(86) …and this can mean, it doesn’t always mean, it can mean, that the students are restructuring their learning … (P108)
(87) You can get quite lost in that, I think, you see. (P84)
(88) Everything they said *could* be quite true and yet it *could* still remain a good book. (P101)
(89) They *may* or *may* not come and connect the television on Saturday. (C134)
(90) On the other hand he may say ‘My dear fellow, of course we understand this problem and we would arrange it this way’. (P52)

The difference between epistemic event-oriented possibility and root-possibility was discussed in 4.2.2.

### 6.2.5.1.3 Proposition-oriented potentiality

When the scope of potentiality markers is the proposition, the general meaning arises that the speaker S does not preclude the truth of the propositional content. The source may lie in the attitude of the speaker towards the propositional content (deontic modality) or in the knowledge of the speaker
(subjective epistemic modality). Proposition-oriented deontic modality expresses the attitude of the (reported) speaker towards the proposition. The distinction between proposition-oriented deontic modality and participant-oriented deontic modality was discussed in 4.2.3.2. The principal difference is that with proposition-oriented deontic modality the participants are not in fact influenced by the deontic source (mostly the speaker). No directive speech act is performed; no permission is actually applied, no rule of conduct is active. It is only the attitude of the (reported) speaker towards the propositional content that is expressed. This use is restricted to the potentiality marker can in negative contexts. It is used to express that S does not appreciate that p is true. S precludes the truth of p because of his values and opinions. See (91):

(91) You can’t have just given up painting completely, not if you had that kind of talent. (C101)

A specific use of might and could can be considered proposition-oriented deontic modality. In particular with deontic modality, expressions can have a stronger or weaker basic sense than the same expressions in other uses (cf. Van der Auwera & Plungian 1998). In (92), might is used to express proposition-oriented deontic modality: no actual permission is applied and the speaker only expresses his attitude towards the proposition:

(92) I don’t mind getting pin money for typing someone’s thesis but they might tell me so beforehand. (C162)

The interpretation of (92) is, however, different from what is to be expected in that the speaker does not so much permit that the proposition would be true, but actually desires that it should be true. This use of a potentiality marker is ironic: it is an indirect speech-act to express obligation in a mitigated way. A similar use is possible with could.

When the source of the potentiality lies in the knowledge of the speaker, it expresses that S does not preclude p’s truth on the basis of his or her personal knowledge, or, in other words, the speaker is uncertain about the truth. The distinction between proposition-oriented epistemic modality and event-oriented epistemic modality was discussed in 4.2.3.3. The potentiality markers may, might and could encode the speaker’s uncertainty or doubt about the truth of the proposition. See (93)-(97) for some examples:

(93) As chairman of our court he may perhaps sometimes feel that… (C138)
(94) I may have put it there # out of the way. (C137)
(95) I think I might have walked out too, from all accounts. (P64)
144  CHAPTER 6

(96) I’m not very good on mechanical language, but one of the cylinders might be just missing a bit from the sound of it. (C150)
(97) Mambo, from Peter Ashworth’s Treadwell stables, could be the one to give King’s Probity most trouble. (C165)

When epistemic potentiality is used with negation, it indicates that the S precludes p from being true, he or she is certain that p is not true. This can be expressed by can’t (98) or couldn’t (99):

(98) I almost phoned them up and said come a bit later – and then I thought oh they’ve probably left by now – so I didn’t and – twelve thirty, now that …. can’t be them, and it was. (Coates 1995: 63)
(99) Well, it couldn’t have been in April, Milord. (Pl62)

The preclusion of the truth of the proposition is based on inference, so that it might be considered as the negative counterpart of the certainty marker must (Coates 1983: 101), that will be discussed in 6.2.5.4.3.

A specific use of may is the concessive “speech-act” use, as in (100):

(100) On the ferries I was always being told # you may have a degree but you’re no better waiter than I am. (C137)

This use falls outside the scope of this thesis as may does not function as a π1-, π2- or π3-operator, but rather as a marker of illocution, which belongs to the interpersonal level in FG (see 2.4). Table 6-2 presents the different semantic functions with the sense of potentiality.

Table 6-2. Subtypes of potentiality, expressed by can, could, may or might

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>Participant-internal</th>
<th>Participant-external</th>
<th>Epistemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOURCE</td>
<td></td>
<td>Non-deontic</td>
<td>Deontic</td>
</tr>
<tr>
<td>Participant-oriented (π1)</td>
<td>ability</td>
<td>root-possibility</td>
<td>permission</td>
</tr>
<tr>
<td>Event-oriented (π2)</td>
<td></td>
<td>root-possibility</td>
<td>permission</td>
</tr>
<tr>
<td>Proposition-oriented (π3)</td>
<td></td>
<td>permission</td>
<td>uncertainty</td>
</tr>
</tbody>
</table>
6.2.5.2 Disposition

The second general sense is disposition, ‘disposed to’. Forms that express this sense are wanna, will, shall and would. In the next sections, the different subcategories are discussed along the lines of the parameters of scope and source.

6.2.5.2.1 Participant-oriented disposition

When the scope of disposition expressions is the predicate, the meaning is that the argument x is disposed to PRED. The source for this disposition in general lies in participant-internal characteristics, which results in the meaning of volition. It is primarily expressed by wanna, and sometimes by will or would. Especially won’t can be combined with inanimate beings, by metaphor. See (101)-(105) for some examples:

(101) I wanna go home.  
(102) Psst, wanna have a better puppy?  
(103) I mean I don’t think the bibliography should suffer because we can’t find a publisher who will do the whole thing. (C170)  
(104) The trouble is # oh for god’s sakes # the key won’t go in the lock. (C173)  
(105) I rang up and said oh I’ve finished now# would you come and get the machine? (C213)

Disposition can also have its source in a deontic authority, expressed by will. However, in these uses, its sense is closer to obligation than to desirability. First, it may be used in questions as an indirect expression of deontic modality. Consider (106):

(106) Will you listen to me and stop interrupting? (C172)

In its literal meaning (106) questions the volition of the addressee. However, by convention, it functions as an indirect expression of obligation. Like might in (92) the basic sense of the form is indirectly used to express a stronger sense.

Furthermore, will can be used deontically in declarative sentences with second person. In this use, it also does not express a mild sense of desirability, but a strong sense of obligation. See (107):

(107) You will listen to me, do what I say, and keep quiet!

---

12 www.imc.org/atom-syntax/mail-archive/msg15517.html  
13 www.cyberpet.com/cyberdog/articles/general/rotq5.htm
In my view, deontic *will* might have originated from its use as a future tense marker. To describe an event in which the addressee is involved as (definitely) taking place in the future, implies that the addressee has no choice, hence a strong deontic sense of obligation. At the present time, though, I have no concrete evidence to support this idea.

### 6.2.5.2.2 Event-oriented disposition

Event-oriented disposition expresses that the event e is disposed to occur. As argued in 4.3.1 this is nearly similar to the meaning of future tense, expressed by *will* (or *shall* for first person). Examples of this use were discussed in 6.2.3.3.

A special use intermediate between volition (participant-oriented) and future tense (event-oriented) is intention. When the first person singular is used in combination with *will*, it may at the same time express that x is disposed to PRED and that e is disposed to occur. See (108)-(110):

(108) I’ll keep an eye open for you (C174)
(109) But I will bring you more today I promise (C174)
(110) I’ll take the old car and then you’ve got the other one (C175)

According to Coates (1983: 173), with intention uses, the focus is not on the ‘subject’s state of mind’, but on the predication. Accordingly, intention uses are in this thesis analyzed as future tense, although, strictly, they may be considered in between modality and future tense. There are no event-oriented modal uses of *will* and *would*, except for its future tense use.

### 6.2.5.2.3 Proposition-oriented disposition

Finally, *will* and *would* can express proposition-oriented disposition. This indicates prediction: S is disposed to conclude p. Future tense can be distinguished from prediction in that future tense indicates that an event will take place in the future, whereas prediction indicates that a proposition will turn out to be true in the future, whereas the event itself can take place in the present or future, or even in the past, in which case it is marked for past by the perfect, (114) and (115). *Would* indicates a more tentative prediction than *will*. Consider (111)-(115):

(111) Tell him Professor Cressage is involved – he will know Professor Cressage. (P57)
(112) They’ll probably be bored with me anyway. (C179)

---

(113) He wouldn’t know exactly where it came down, but he might well have a rough idea. (P68)

(114) And my mother was not [drunk]. Several people in the house will have said that to you. (C179)

(115) Let me see # when would he have been born. (C217)

When the event itself takes place in the future, utterances can be ambiguous between a future and a prediction interpretation. See (116) and (117). This ambiguity is comparable to event-oriented and proposition-oriented modality: does the speaker present the information as an objective or subjective statement? The correct interpretation can only be given in the context.

(116) He’ll burn himself out if he goes on at this rate. (C227)

(117) It is a fairly safe bet that one of the guests will want to take the empty flask home; they make delightful lamp bases. (C178)

Table 6-3 summarizes the different semantic functions with the basic sense disposition.

**Table 6-3. Subtypes of DISPOSITION, expressed by will, would, or wanna**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>Participant-internal</th>
<th>Participant-external</th>
<th>Epistemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOPE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-deontic</td>
<td>Deontic</td>
<td></td>
</tr>
<tr>
<td>Participant-oriented (π₁)</td>
<td>volition</td>
<td>-</td>
<td>“desirability”</td>
</tr>
<tr>
<td>Event-oriented (π₂)</td>
<td>-</td>
<td>-</td>
<td>(future)</td>
</tr>
<tr>
<td>Proposition-oriented (π₃)</td>
<td>-</td>
<td>-</td>
<td>prediction</td>
</tr>
</tbody>
</table>

**6.2.5.3 Weak necessity**

The third basic sense that is discussed is weak necessity. This sense could be paraphrased as ‘expected to’. English has different forms to express this sense: should, ought to, and be supposed to are the most important ones. One use of shall may also be covered under the sense of weak necessity.
6.2.5.3.1 Participant-oriented weak necessity

When the scope of weak necessity is the predicate, it expresses that x is expected to PRED. The source for this expectation can only lie in participant-external circumstances. So English has no expressions for ‘weak internal need’. The external circumstances can be deontic or non-deontic, but especially within weak necessity, this is not always clear. Examples (118)-(121) can be interpreted as non-deontic circumstances; x is expected to PRED because of general external circumstances:

(118) A: What you been doing? B: Well, I shouldn’t be here. I ought to be on holiday today. (C82)
(119) We ought to return the cassette some time fairly soon you see. (C71)
(120) I left a note in my pigeon hole which should be in the student’s pigeon hole, so I’d be grateful if you could move it …(C65)
(121) She’s supposed to be coming in. (B500)

In particular for expressions of weak necessity it is not always clear whether the external circumstances should be considered a deontic source or not. In the above examples (118) and (120) the participants seem to be forced by a sort of organizational plan or schedule, which is closely related to a general deontic source, such as a rule of conduct or a law (which is analyzed as event-oriented deontic modality, see 6.2.5.3.2). However, the specific participants in (118) and (120) cannot be replaced by a non-specific participant, so that the modal expressions are analyzed as participant-oriented.

Examples (122)-(125) below can be interpreted as expressing weak obligation: x is expected to PRED because of a deontic source. This category is intermediate between permission and obligation. Dependent on the status difference between the deontic source and the participant—thus not on the expression chosen—the strength of the desirability runs from a mere advice or invitation to a weak or even rather strong obligation. Compared to the obligation markers in the necessity domain, this type of obligation is a mitigated form. When it is used with speaker-reference the speaker expresses self-exhortation.

(122) You should walk round the ramparts of the old city too. (C58)
(123) A: If I put a Marvel [tin] in his wastepaper basket will he think it unaesthetic.
   B: I shouldn’t worry. (C222)
(124) So perhaps I ought to ask you some further questions. (P123)
(125) You really ought to be buying something a bit more modern and a bit more expensive. (C72)
6.2.5.3.2 Event-oriented weak necessity

When the scope of weak necessity is the predication, it comes to mean that the
event e is expected to occur. The source of the necessity may first of all lie in
general non-deontic circumstances. This may apply to utterances with non-
specific subjects or passives, such as ought to in (126) and should in (127),
although both examples leave open the possibility of a deontic interpretation.
Be supposed to can also be interpreted in the sense that e is expected to occur
because of general, unspecified circumstances. It means something like ‘it is
meant to’. Consider (126)-(129):

(126) The job here ought to be finished in a matter of days. (C73)
(127) This should be done before the pollen is ripe. (P123)
(128) My class is supposed to start in fifteen minutes. (CWO)
(129) Are you supposed to be going with him on holiday? (B500)

The source of event-oriented weak necessity may also be a general deontic
source, such as a rule of conduct or a law. Examples are presented in (130)-(133):

(130) I just insisted very firmly on calling her Miss Tillman but one should
really call her President. (C59)
(131) I don’t think you ought to … I’ve a sort of feeling you shouldn’t ask.
(C81)
(132) But they can’t decide yet whether the new man should be appointed by
Freeman (…) or whether the appointment should be made by whoever
takes his place. (C59)
(133) As Americans, we believe that one is supposed to vote and participate
in the democratic process.\footnote{foucault.info/Foucault-L/archive/msg08657.shtml}

With first person subjects it is not always clear whether the speaker is the
deontic source and the utterance is an instance of self-exhortation or whether
the speaker cites a general rule of conduct, which also applies to him/herself.
This is illustrated in (134) and (135):

(134) I ought to be ashamed to say so, but I can’t. (P123)
(135) I think I ought to say something about my step father. (C71)

The final possible source of event-oriented weak necessity is general
knowledge. Should, ought to and supposed to express probability: e is expected to

\footnote{foucault.info/Foucault-L/archive/msg08657.shtml}
occur or to be the case because of general knowledge. Objectively, the occurrence of the event is probable. Consider (136)-(138):

(136) Lad says if it is on the delivery note it should have been delivered. (C226)
(137) It ought to be I suppose in those two # those double grey filling cabinets you remember. (P74)
(138) But medical supplies are supposed to be getting through. (CWO)

6.2.5.3.3 Proposition-oriented weak necessity
Finally, when weak necessity modifies the proposition, it expresses that S expects p's truth or expects p to be true. The first possible ground for expecting the truth of the proposition is that the speaker infers the propositional content on the basis of his or her personal knowledge (subjective epistemic modality). This is mainly expressed by should, as in (139)-(142):

(139) Have sent off my diary a couple of days ago – you should get it soon. (C58)
(140) So he should be around sort of between half past two and half past three. (P60)
(141) The trip should take about sixteen days. (C64)
(142) That shouldn't be difficult. What do you want to know? (P62)

A second ground for expecting the truth of the proposition are the speaker's own values and opinions about how the world should be. This is proposition-oriented deontic modality. The weak obligation is not actually applied to the participants; it is only the personal opinion of the speaker that the proposition should be true (4.2.3.3). Consider (143)-(144):

(143) The unemployed, they should be made to do some some work and not scrounge off the state. (C59)
(144) I think people ought to be better informed about what marriage entails. (P123)

A use of should which falls outside the scope of this thesis is its quasi-subjunctive reading, as in (145)-(146):

(145) It is most necessary that we should have the funeral bill. (P68)
(146) It was inevitable that Peter Ustinov should join the exclusive four-star club by writing, producing, directing and starring one film. (P68)
This use belongs to the interpersonal level in FG (see 2.4) and is further excluded.

To resume, the specific semantic functions with the basic sense of weak necessity are presented in Table 6-4.

Table 6-4. Subtypes of *weak necessity*, expressed by *should*, *ought to*, or *be supposed to*

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>Participant-internal</th>
<th>Participant-external</th>
<th>Epistemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOPE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant-oriented ((\pi_1))</td>
<td>-</td>
<td>weak necessity</td>
<td>weak obligation</td>
</tr>
<tr>
<td>Event-oriented ((\pi_2))</td>
<td></td>
<td>weak necessity</td>
<td>weak obligation</td>
</tr>
<tr>
<td>Proposition-oriented ((\pi_3))</td>
<td></td>
<td>weak obligation</td>
<td>weak certainty</td>
</tr>
</tbody>
</table>

6.2.5.4 *Necessity*

The final basic modal sense is necessity, expressing ‘forced to’. There are several expressions in English with this basic sense: *need* (to), *have to*, *have got to*, *must* and *shall*.

6.2.5.4.1 *Participant-oriented necessity*

When the scope of necessity is the predicate, it expresses that the argument \(x\) is forced to \(PRED\). The source of this force may be internal characteristics of the argument, which expresses the meaning of internal need: \(x\) is forced by internal/intrinsic factors to \(PRED\). It is expressed by *need to*, *have to* and *have got to*. See (147)-(150):

(147) Boris needs to sleep ten hours every night for him to function properly. (Van der Auwera & Plungian 1998: 80)
(148) I need to be left in peace today. (Van der Auwera & Plungian 1998: 83)
(149) I’ve got to eat something, otherwise I’ll just die.16
(150) I’m going to make this short since it’s way past midnight and I really have to sleep.17

More commonly than purely by internal characteristics, it is expressed that \( x \) is forced to PRED (mainly) by external circumstances. These circumstances may be non-deontic. This use is illustrated by (151)-(153):

(151) I’ve got to be at London airport at fourish. (P114)
(152) We had to make a special trip down to Epsom to collect the bloody thing. (P120)
(153) I may need to stay a couple of nights at Minna before I can find transport for the last 60 miles or so. (P127)

It is also possible that the external circumstances are deontic: \( x \) is forced by a deontic source to PRED. The necessity markers \( \text{have to} \), \( \text{have got to} \), and sometimes \( \text{must} \) or \( \text{need} \) express obligation.\(^{18}\) Consider (154)-(157):

(154) Me mum says I have to wait. (Krug 2000: 92)
(155) Have we got to go on this hike? (C52)
(156) You must tell me how to get to it. (C34)
(157) “I’m very grateful to you.” “You needn’t be. I told you. I’m glad to do it.” (C50)

English makes a distinction in deontic expressions that indicate that the speaker is the deontic source or that it is some other deontic source. In \( \text{You must come and see me tomorrow} \) (Palmer 2001: 75), it is the speaker who is the deontic source, whereas with \( \text{you have to come and see me tomorrow} \) there is an obligation that is independent on the speaker. The same distinction holds for the opposition between \( \text{should} \) and \( \text{ought to} \) on the one hand, where the speaker is the responsible deontic source in contrast to \( \text{be supposed to} \), where the speaker is not responsible for the weak obligation.

Finally, the expression \( \text{shall} \) can be used as a deontic expression of obligation. See the example in (158):

(158) You shall go, I insist on it. (Sweetser 1990: 55)

The notion of necessity or obligation is one of the oldest uses of \( \text{shall} \) and it grammaticalized to a future meaning (see 6.3.3). \( \text{Shall} \) can also be used deontically in questions such as (159), in which case the addressee is the deontic

---

\(^{17}\) www.infektia.net

\(^{18}\) Contrary to Bybee et al. (1994: 320) I use the term obligation only if the external circumstances are identified as a deontic source and not for all types of external circumstances.
source. However, in these constructions its sense is much weaker, and closer to weak obligation or desirability than to strong obligation.

(159) What shall I choose for Ann if she wins the raffle? (CWO)

These uses are the only remnants of the modal uses of shall. Its past counterpart should developed to a marker of weak necessity that is fairly frequent.

6.2.5.4.2 Event-oriented necessity
When the scope of necessity is the predication, it means that the event e is forced to occur. It may be forced to occur because of general external circumstances. It occurs in contexts with non-specific participants or passives. Consider (160)-(163):

(160) Clay pots must have some protection from severe weather. (C35)
(161) A really healthy effective opposition which you’ve got to have if you’re going to shake the government. (C53)
(162) He’s going on the 7.40 tomorrow morning and everything must be packed tonight. (C35)
(163) He fully understands that the thing has to be reprogrammed every year. (C55)

It is also possible that e is forced to occur because of a deontic source, such as a general rule or law. This is in general expressed by have to, have got to or must: e is forced to occur because of a general deontic source. Examples are presented in (164)-(168):

(164) If you commit murder, Charlotte, you must be punished. (C34)
(165) In the whole of southern Africa one has to drive on the left side of the road.19
(166) You haven’t got to park on double yellow lines – it’s against the law. (Perkins 1983: 62)
(167) You have to listen carefully and look for clues from three sources.20
(168) There is already a great imbalance between what a student has to pay if he’s in lodgings and what he has to pay if he is in a hall of residence. (C55)

Shall can also be used in this sense, although it is infrequent. See (169):

19 www.namibia-travel.net/namibia/a_z.htm
20 www.pinkmonkey.com/electricdesk/studysmart/ssmart5.asp
(169) (The law decrees that) all citizens shall constantly carry violet parasols from 3/9/83 on. (Sweetser 1990: 55)

The third possible source of event-oriented necessity can be knowledge, although this use is restricted to negated necessity: e is not forced to occur. It seems communicatively irrelevant to encode that the event is necessarily actual, since that is the unmarked communicative situation (see 4.2.3.3). Negated epistemic necessity is semantically the opposite of epistemic possibility: ‘not necessarily e’ is similar to ‘possibly not e’. This meaning can be expressed by need not or don’t have to. See (170)-(171):

(170) Oh gosh, getting married is an awfully complicated business. [other speakers argue] Actually it needn’t be – it can be very straightforward. (C50)
(171) He doesn’t have to be at home: he could have gone straight to Caroline.

6.2.5.4.3 Proposition-oriented necessity

Finally, when the scope of necessity is the proposition, it means that the speaker S is forced to conclude the propositional content p. This is most commonly expressed by must, but also by have to or have got to. S may be forced to conclude p because of the available premises. In that case it expresses certainty about the truth of the proposition. Consider (172)-(175):

(172) She must have been such a pain in the neck to her Mum and vice versa. (C44)
(173) This has to be the biggest ant-hill ever seen. (Perkins 1983: 61)
(174) You’ve got to be joking. (P56)
(175) It had to be there – there wasn’t anywhere else it could have been. (P65)

Note that event-oriented epistemic necessity is always negated (see (170)-(171) and 4.2.3.3), whereas proposition-oriented epistemic necessity on the other hand is always positive. The speaker states that from the available premises, he is forced to conclude p. If the speaker wants to state that he is certain that p is not true, can’t is used as in (98) above.

It is also possible that S may be forced to conclude p because of his personal opinions and values. In that case the speaker expresses that he is certain that p should be true. In this interpretation the modal expresses proposition-oriented deontic modality (4.2.3.3). There is no obligation actually laid on the
participants: it is only the speaker’s attitude towards the proposition that is expressed. S states that p must be true in his opinion. See (176) and (177):

(176) The government must act. It must make up its mind about priorities – offices or houses, housing estates or luxury buildings. (P106)
(177) This I think is something on which universities have got to begin now to take a stand on. (C53)

Table 6-5 presents an overview of the different functions that can be expressed with the basic sense of necessity.

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>Participant-internal</th>
<th>Participant-external</th>
<th>Epistemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOPE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participant-oriented (π1)</td>
<td>internal need</td>
<td>root necessity</td>
<td>obligation</td>
</tr>
<tr>
<td>Event-oriented (π2)</td>
<td>root necessity</td>
<td>obligation</td>
<td>(no necessity)</td>
</tr>
<tr>
<td>Proposition-oriented (π3)</td>
<td>obligation</td>
<td>certainty</td>
<td></td>
</tr>
</tbody>
</table>

6.2.6 Irrealis

English has no simple morpheme that marks irrealis. There are, however, specific syntactic constructions that indicate irrealis. The construction to mark hypothesis consists of a subordinate conditional clause (if-clause) and a main clause. The conditional clause contains the condition-marker if and a verb with past tense-morphology: the past form does however not mark past tense, but indicates irrealis in this construction. The condition is often explicated, but can also be left implicit. The main clause contains a secondary or ‘past’ modal, which also does not indicate past tense, but irrealis. The most neutral one is would: in most cases it only marks irrealis, but sometimes it also has a sense of volition. The other secondary modals—could, should, might—can also be used. Apart from marking irrealis, they bear the semantics of their primary modal
counterparts: *can*, *shall*\(^{21}\) or *may*, discussed in the previous section. Examples of the hypothetical construction are presented in (178)-(181):

(178) If I acted like that in front of him I # I mean that would be # that would be inexcusable. (C213)
(179) If he came to live with us then she would have to give up work. (C229)
(180) Anybody could produce at least two test teams of English writers who would be much better worth celebrating than Burns. (C215)
(181) That would apply to Swift too # wouldn’t it? (C217)

Counterfactual constructions in English consists of a conditional clause with a main verb with *past perfect*-morphology, and of a matrix clause that contains a secondary modal plus a main verb with perfect-morphology. Once again, *would* is the most neutral modal, in most cases only indicating irrealis, whereas the other secondary modals—*should*, *might*, *could*—also add the semantics of their primary modal counterpart. Consider (182)-(184):

(182) In fact, I *would have said* that it looks as though London would be worth going through. (P173)
(183) I *would have been* very surprised really if you *had got* a British Academy award. (C214)
(184) If I *could have thought* of that quotation I *would have used* it. (C214)

6.2.7 Summary

In Chapters 3 and 4 the possible semantic functions within different TMA domains and their scopes were established. Here it was examined which English forms cover which semantic functions. It appeared that there are several expressions that are used for different functions and that similar functions are often expressed by different expressions. Although it is sometimes difficult to distinguish between semantic functions and although certain utterances may allow different interpretations, this is not an argument for rejecting the distinct semantic functions all together. On the contrary, it explains how the semantics of a specific linguistic item—as understood by language users—can change or differentiate and why processes of grammaticalization are possible. Only by virtue of implicatures and ambiguity

---

\(^{21}\) The overlap of the semantics of *shall* and *should* is not so neat as the overlap of the semantics of *can* and *could* or of *may* and *might*. Both *shall* and *should* have uses that the other form cannot have. There already are considerable differences between *sceall*/*scealt* and *sceold* in Old English (Goossens 1985).
between related meanings can language users change the use and meaning of a linguistic item.

This section provides a summary of the discussion by giving a list of the English operators that currently belong to the TMA system. In Table 6-6 all the TMA operators with narrow scope ($\pi_1$-operators) are presented: the operators of aspect and operators of participant-oriented modality. Table 6-7 presents the TMA operators with medial scope ($\pi_2$-operators): operators of tense, event quantification, event-oriented modality and irrealis. Finally, Table 6-8 presents the operators with wide scope ($\pi_3$-operators), expressing proposition-oriented modality. These tables serve as the basis for examining the research questions in the next sections.

Table 6-6. English TMA expressions with narrow scope ($\pi_1$-operators)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Semantic function</th>
<th>Expression(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASPECT</td>
<td>Progressive</td>
<td>be –ing</td>
</tr>
<tr>
<td></td>
<td>Continuative</td>
<td>keep –ing</td>
</tr>
<tr>
<td></td>
<td>Prospective</td>
<td>be gonna / be going to</td>
</tr>
<tr>
<td></td>
<td>Immediate prospective</td>
<td>be about to</td>
</tr>
<tr>
<td></td>
<td>Perfect</td>
<td>have –ed, have + irregular form</td>
</tr>
<tr>
<td></td>
<td>Resultative</td>
<td>be –ed, be + irregular form</td>
</tr>
<tr>
<td>PARTICIPANT-ORIENTED MODALITY</td>
<td>Ability</td>
<td>can, could</td>
</tr>
<tr>
<td></td>
<td>Root possibility</td>
<td>can, could, may, might</td>
</tr>
<tr>
<td></td>
<td>Permission</td>
<td>can, could, may, might</td>
</tr>
<tr>
<td></td>
<td>Volition</td>
<td>wanna, will, would</td>
</tr>
<tr>
<td></td>
<td>Weak root necessity</td>
<td>should, ought to, be supposed to</td>
</tr>
<tr>
<td></td>
<td>Weak obligation</td>
<td>should, ought to, be supposed to</td>
</tr>
<tr>
<td></td>
<td>Internal need</td>
<td>need (to), have to, have got to</td>
</tr>
<tr>
<td></td>
<td>Root Necessity</td>
<td>must, have to, have got to</td>
</tr>
<tr>
<td></td>
<td>Obligation</td>
<td>must, have to, have got to, need to, shall</td>
</tr>
</tbody>
</table>
Table 6-7. English TMA expressions with medial scope (π₂-operators)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Semantic function</th>
<th>Expression(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TENSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>–ed, irregular form</td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>–o, –s</td>
<td></td>
</tr>
<tr>
<td>Future</td>
<td>will, shall, would</td>
<td></td>
</tr>
<tr>
<td>EVENT-ORIENTED MODALITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Root possibility</td>
<td>can, could, may, might</td>
<td></td>
</tr>
<tr>
<td>Permission</td>
<td>can, could, may, might</td>
<td></td>
</tr>
<tr>
<td>Weak Necessity</td>
<td>should, ought to, be supposed to</td>
<td></td>
</tr>
<tr>
<td>Weak Obligation</td>
<td>should, ought to, be supposed to</td>
<td></td>
</tr>
<tr>
<td>Necessity</td>
<td>must, have to, have got to</td>
<td></td>
</tr>
<tr>
<td>Obligation</td>
<td>must, have to, have got to, shall</td>
<td></td>
</tr>
<tr>
<td>Epistemic possibility</td>
<td>could, may, might</td>
<td></td>
</tr>
<tr>
<td>(negated) epistemic necessity</td>
<td>need not, don’t have to</td>
<td></td>
</tr>
<tr>
<td>Probability</td>
<td>should, be supposed to</td>
<td></td>
</tr>
<tr>
<td>IRREALIS</td>
<td>Hypothetical</td>
<td>–ed, irregular form, past modal</td>
</tr>
<tr>
<td>Counterfactual</td>
<td>had –ed, past modal + –ed</td>
<td></td>
</tr>
<tr>
<td>EVENT QUANTIFICATION</td>
<td>Habitual</td>
<td>used to, would, will</td>
</tr>
<tr>
<td></td>
<td>Frequentative</td>
<td>keep –ing</td>
</tr>
</tbody>
</table>

Table 6-8. English TMA expressions with wide scope (π₃-operators)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Semantic function</th>
<th>Expression(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROPOSITION-ORIENTED MODALITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertainty / certainty that not</td>
<td>could, may, might / can’t, couldn’t</td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td>will, would</td>
<td></td>
</tr>
<tr>
<td>Weak certainty</td>
<td>should</td>
<td></td>
</tr>
<tr>
<td>Certainty</td>
<td>must, have to, have got to</td>
<td></td>
</tr>
<tr>
<td>“Permission” / no permission</td>
<td>might / can’t</td>
<td></td>
</tr>
<tr>
<td>Weak obligation</td>
<td>should</td>
<td></td>
</tr>
<tr>
<td>Obligation</td>
<td>must, have to, have got to</td>
<td></td>
</tr>
</tbody>
</table>
6.3 **Diachrony**

This section will examine whether English shows the hypothesized pattern of diachronic development: on the basis of inference, speakers will widen the semantic scope of operators to more abstract functions. The expected diachronic development of TMA expressions is formulated in H1:

H1: Diachronically, operators show an increase in scope, and develop in the direction from $\pi_1$-operator to $\pi_2$-operator to $\pi_3$-operator.

A brief discussion will be presented of the diachronic development of each TMA expression in English. Inevitably, this discussion involves strong simplifications and the reader is referred to more elaborate works on the history of modal auxiliaries and English syntax in general (Hogg 1992-2001; Lightfoot 1979; Plank 1984; Traugott 1972; Visser 1963-1973).

Reference will be made to the different stages of English that may not be familiar to all readers. The earliest stage recognized as ‘English’ is Old English, from approximately AD 450 till 1150. The Middle English period extends from about 1150 till 1500, Early Modern or New English from about 1500 till 1750 and Late Modern English or Present Day English from circa 1750 onwards. In this section, asterisks are used to indicate that a historical form is reconstructed.

### 6.3.1 Tense Inflection

The tense inflection and irregular forms are the oldest TMA operators in English. The strong, irregular past forms in English stem from regular Proto-Indo-European ablaut formation (root vowel alternation). The origin of the past tense inflection –ed is not known (Bammesberger 1992: 59; P. Ramat 1998: 405; Shields 1992: 43). It has a long history, since all Germanic languages, in contrast to other Indo-European languages, have a dental preterit marker.

The present tense inflections, -o and -s, are the remnants of a more elaborate system of present tense person/number inflection in Proto-Indo-European. In Old English it was a general non-past marker, being the default marker for present and future time reference, whereas in Present Day English reference to future time by the simple present tense is very limited. The history of both tense markers cannot be tested to the hypothesis, as there is not enough information available.

### 6.3.2 Aspectual Constructions

The resultative be –ed and perfect have –ed have their origin in the Old English lexical constructions *wesan/ beon* (‘be’) + adjectival participle of intransitive verbs and *habban* (‘have’) + adjectival participle of transitive verbs. The adjectival
participle agreed with the subject (intransitive) or object (transitive) and the verbs still had their lexical meaning of existence ('be') and possession ('have'). This raised meanings like 'they were in the state of having come' or 'I had him in a state of being bound.' More and more often, the inflection on the participle was dropped which made it possible to analyze the construction as an auxiliary with a past participle. Furthermore, a word order change took place so that formerly 'has the fish caught' gradually became 'has caught the fish.' Finally, have takes over more and more from be as perfect auxiliary and in the 19th century it has become the default perfect auxiliary (Fischer 1992; Traugott 1972: 91-94). In sum, the developmental paths are presented in Figure 6-7:

![Diagram](image)

**Figure 6-7.** Developmental paths to resultative and perfect

The history of the resultative and perfect constructions is in accordance with the hypothesized order: lexical constructions have developed into aspect markers.

The origin of the progressive be –ing is not completely clear. According to Traugott (1992: 187-89) two separate constructions are underlying the later progressive. Old English had a construction wesan/beon + present participle – ende, as in be wæs huntende and a locative construction of a preposition on / an in combination with a nominal derived from a verb, ending in -ung/-ing, as in be wæs on huntunge. In several contexts, the constructions were equivalent in that they created the same meaning. Furthermore, the inflectional endings began to be confused in late Old English/Early Middle English and possible the two constructions became very similar in Middle English, i.e., *He was hantung(e) and He was on/an/in/a hantung(e). In the end they may have coalesced to what is the Present Day English progressive construction. Although this hypothesized development is plausible, the actual language data do not prove it (Fischer 1992). Nevertheless, what is important to this thesis is that the progressive construction be –ing has its immediate origin in a lexical construction. See Figure 6-8:
This path supports the hypothesis that narrow scope operators do not derive from wider scope operators, but rather from lexical constructions.

The prospective construction *be going to* started as the progressive construction of the directional lexical verb *go* and a purposive clause ‘*to V*’, meaning *going in order to*. This led to the common inference that the subject intended to *V* at a later time. In the next stage, the directionality sense of *go* was demoted and the inference of intention of event realization at a later time was promoted. In the stage thereafter, the purposive sense was also demoted and *be going to* could refer to any pre-state of a coming event or state and no longer applied selection restrictions on the subject or the predicate with which it could combine (Hopper & Traugott 1993: 61, 81-83). In sum, the development is:

This path supports the hypothesis again: the aspect operator directly derives from a lexical construction.

There was hardly any information found about the history of the two other aspectual constructions, *keep –ing* and *be about to*. The latter construction occurs from Late Middle English (Fischer 1992), refers to planned action and has aspectual implications from the start (Rissanen 1999: 233). In all likelihood, both constructions derive from lexical constructions, not from other grammatical operators. This may be considered evidence in support of the hypothesis.

### 6.3.3 Central modals

For modal expressions, in most cases there is no indication in the literature whether a root- or deontic sense can be combined with specific and non-specific participants (the distinction between participant-oriented and event-oriented modality) and whether an epistemic meaning should be considered expressing an objective epistemic stance or the speaker’s commitment. If there is not enough information available, the root- and deontic meanings will be
conceived of as \( \pi_1 \)- and \( \pi_2 \)-operators and the epistemic meanings as \( \pi_2 \)- and \( \pi_3 \)-operators.

The forms *will* and *would* go back to the Old English verb *willan* (present *wile*, past *wolde*), which was the lexical verb for ‘want’. Late Old English *willan* developed into a grammatical marker of volition. It is occasionally used to express typical or general occurrences, close to a habitual or epistemic interpretation. Sometimes, *wile* is used to express senses of futurity, but this is restricted to contexts of generalizations, prophetic or inevitable future or relative future tense. These uses occur in impersonal or passive contexts. Gradually, during Middle English *will* (and *shall*) becomes a real future marker, and replaces the present tense in this use (Traugott 1989: 38-43; 1992: 195). The sense of desire is lost in the development of Early Modern English into Present Day English and the sense of volition has become much less frequent (Warner 1993: 181) The prediction use of *will* and *would*, in the sense of ‘the speaker is disposed to conclude …’, occurs only in the 19th century. (Traugott 1989: 43). In sum, the development is:

![Figure 6-10. Developmental path of will](image)

in which the symbol ‘†’ indicates that this use has died out. The development in Figure 6-10 shows an increase in scope and is therefore in accordance with the predicted development.

The origin of *shall* and *should* lies in the Old English *sculan* (present *sceal*, past *sceolde*), a lexical verb meaning ‘owe, be necessary’, both in moral and financial respect (Traugott 1992: 195). In Old English, it developed into an auxiliary marking obligation and necessity of various kinds. At the end of Old English it acquired a sense of intention. According to Traugott (1992: 197) and Warner (1993: 170) *sceal* was used occasionally for future of arrangement or inevitable future, especially in impersonal constructions. During the Middle English stage, the forms *shal* and *wil* become the normal markers for future time reference. *Shal* is the predominant form, but *wil* is more typical in informal usage. *Will* as a future marker gradually replaces *shall* in Early Modern English with third and later second person subjects. In Present Day English *shall* is entirely restricted to use with first person and nowadays, *will* has even replaced *shall* in first person contexts in certain dialects or in informal speech (Warner 1993: 181). The development of *shall* supports the hypothesis. It is presented in Figure 6-11:
Some specific developments took place in the use of the past form, *swold*. *Should* develops at first similar to *shall*, although it grammaticalizes a bit later as a future marker, in Early Modern English (Rissanen 1999: 235). In contrast to *shall*, *should* develops further into a marker of epistemic probability in Present Day English (Bybee et al. 1994: 200). There is no detailed information available about the scope of the epistemic modality. Consider Figure 6-12:

This development is consistent with the hypothesis. There is a developmental order from lexical to π₁-operator, to π₂-operator to π₃-operator.

*Should* furthermore had a special use in Old English in indirect speech. It indicated hearsay (evidential use). This use existed into Modern English (Traugott 1989: 41-42; Warner 1993: 171, 78). It is not known how this use came into being. Therefore, it will not be used to test the hypothesis. It is however clear that this is an independent development from the one described in Figure 6-12.

The form *must* has its origin in the Old English verb *motan* (present *mōt*, past *mōste*). It was an auxiliarized expression marking permission and root-possibility. However, in late Old English, it developed a meaning of root-necessity and obligation, which quickly became the predominant use (13th century). The contradictory interpretations of *mōt* may have been caused by the basic sense of *mōt* ‘to be allotted’, which allows both interpretations. Another possibility is that the sense change took place in negative contexts, as ‘not possible’ means ‘necessarily not’ (Warner 1993: 160). The permission meaning gradually lost ground during the Middle English period and *may* took over the function of *mōt* in those contexts. During late Middle English the form *must* became the default form, which may have its origin in the subjunctive preterite of Old English, *moste*, and in the second person singular present form of Old English, *mōst*. The present form *must* disappears in the 16th century.
In Middle English *moot* begins to be used in contexts of inevitable or expected future, with a sense of epistemic necessity. Strong epistemic interpretations are only possible if *moot* is accompanied by epistemic adverbs. From around the 17th century, *must* is established as a common marker of epistemic certainty (Traugott 1989: 42). Consider Figure 6-13:

![Figure 6-13. Developmental path of must](image)

The development of *must* is clearly in accordance with the hypothesis: it follows the direction of the Scope Hierarchy.

*May* and *might* stem from the Old English verb *magan* (present *mæg*, past *meahte*), expressing physical ability or power. Toward the end of this stage it developed a general ability meaning. During Middle English *may* takes over from *mōt* in contexts of deontic uses of permission and root-possibility. The use of *may* for expressing physical power was lost in the 15th century. During the 17th century the sense of general ability is also lost (Rissanen 1999: 237; Warner 1993: 180-81).

The development of epistemic meaning is not very well documented. In Old English, *may* is occasionally used with marginal epistemic senses (Traugott 1992: 197). In Middle English this use becomes more firmly established but even in Early Modern English it is less commonly used epistemically than with a general root-possibility sense. (Rissanen 1999: 237). Epistemic uses of *may* seemed to have developed before the permission uses of *may*; however, the permission meaning did not derive from the epistemic meaning, but directly from the ability meaning (Bybee & Pagliuca 1985). It is not clear from which meaning the epistemic meaning has derived, the ability or root-possibility sense. It is furthermore not clear whether the first epistemic uses were event-oriented or proposition-oriented. I assume that they were event-oriented as Traugott states that pre-modals in Old English ‘with possible epistemic meanings concern possibility in a world independent of the speaker and can be considered to be only weakly subjective.’ (1989: 42, emphasis mine). As for today, *may* and especially *might* are losing their deontic and dynamic uses and become more and more restricted to epistemic use, especially in American English (Denison 1998: 165; Fischer 2003: 23-24). In sum, the two developmental paths are:
Both paths agree with the predicted order: they develop from narrow scope operator to wide scope operator.

Can developed from an Old English lexical verb cunnan (present can, past cūđe) that expressed ‘know’/ ‘have the mental ability’ (Warner 1993). In Middle English it has become an auxiliary expressing general ability and root-possibility and this is still the predominant use in Early Modern English. In the 16th and 17th century it outstrips may in these functions. The meaning of permission is late, not before the 19th century (Rissanen 1999: 237; Warner 1993: 177). Epistemic meanings of can are very infrequent before the Modern English period and can and could are probably used epistemically from the 19th century, especially in negations (not possible). Can gradually loses the meaning of ‘know’ in the period from Early Modern English to Present Day English (Rissanen 1999: 237). As with may, epistemic uses of can seem to have developed before the permission uses of can and the permission meaning has directly derived from the ability meaning, not from the epistemic meaning. They are parallel, independent semantic changes (Bybee & Pagliuca 1985). The use of can as a marker of uncertainty is very restricted in Present Day English (see 6.2.5.1.3), which suggests that it developed later than objective epistemic possibility. See Figure 6-15:
Like the other modals, the semantic changes of *can* and *could* support the hypothesis.

### 6.3.4 Other modals

This section will discuss the diachronic development of all other modal markers. They have gained grammatical status much more recently. *Be supposed to* is the first expression to be discussed. The source of this construction is the French verb *supposer* meaning ‘believe’ or ‘hypothesize’ and from the 15th century also ‘expect’. In the early stages in which the construction with *be* was used, it occurred in statements describing time-stable truths or generic situations, most frequently in the constructions *it’s supposed to be X* or *it is supposed to have –ed*. Ziegeler (2003: 60) claims that these uses are evidentials of hearsay. In a later stage there has been a change from mainly expressing belief or hypothesis of an unnamed source to intention or expectation imposed on the subject (Ziegeler 2003).

If Ziegeler’s reconstruction of the history of *be supposed to* is correct, then there has been an evidential use (π3) before a weak obligation or weak root-necessity use (π1). In fact, this development would semantically be plausible: If the general opinion is that something is the case, then it may be inferred that there are probably (external) circumstances that cause this something to be the case. What is however unclear from Ziegeler’s description is whether the evidential reading should be considered a grammatical construction. It seems to have been the normal passive form of the active lexical verb, which is, according to my criteria (6.2.1) not a grammaticalized construction. In the weak obligation or necessity reading, however, the construction has no longer the resultant meaning of its components and this is probably the first real grammatical use. The development would then look like Figure 6-16, which is in line with the predicted order:

![Figure 6-16. Assumed developmental path of supposed to](image)

The construction *have to* stems from the lexical verb *habban* ‘have’. In Old English this verb underwent the following generalization in meaning: ‘to hold in hand’ developed into ‘to have in one’s immediate personal possession (physically present)’ which developed into ‘to have or own as a possession (not physically present)’ which in turn came to mean ‘to have as an abstract
possession, such as time, an idea, an education, a debt’ (Bybee & Pagliuca 1985: 72).

Additional steps in the grammaticalization process are described in Krug (2000: 55), based on Heine (1993: 41f): I have a letter, expressing possession became I have a letter to mail, expressing possession and purpose, which in turn developed into I have a letter to write in which the possessive meaning of have to has bleached out. This further developed into the meaning I have to write a letter, in which have to expresses obligation and this finally came to mean I have to write, in which the object can be omitted. The sense of obligation or necessity derives from the semantics of have itself. In particular the gradually more frequent construction have to say/do (something unpleasant) seemed to have given rise to the inference of obligation or necessity (Krug 2000: 97ff) In Middle English the first interpretations as obligation or necessity arise, although it remains rare until about 1850 when there is a dramatic frequency increase of have to. Epistemic senses only occurred from the 20th century, but these uses are still rare (Krug 2000: 74, 89). See Figure 6-17 for the complete developmental path:

![Figure 6-17. Developmental path of have to](image)

Related to have to is have got to, originally the perfect construction of get, that combined only with a noun phrase. It entered the English language in the 16th century and expressed possession at first. The rise of the construction have got to could have been motivated by the fact that have was very often reduced to the contracted forms ’ve, ’s or ’d. Got may have been inserted in cases where have was reduced to make the expression perceptually more salient. In the 19th century, the necessity and obligation reading emerged and came to be used very frequently in the 20th century. Have got to did not go through all the stages that have to did. It seems that have got to could become used with obligation and necessity meanings by analogy to the construction have to (Krug 2000: 61ff). In Present Day English have got to, often reduced to got to, can be used epistemically. This use must be very recent and is probably rather infrequent. The developmental path of have got to is presented in Figure 6-18.
Both Figure 6-17 and Figure 6-18 correspond to the hypothesis.

In 6.2.1 it was argued that the contracted construction wanna is grammaticalized. The verb want (wont) is first attested in early Middle English, meaning ‘lack’. This remains the core sense until the end of Early Modern English (Krug 2000: 127, 41). However, already in Middle English there are a few examples in which a sense of desire can be inferred in contexts of want, for a human being often desires what he lacks (p.129). This volitional sense gradually changed from an implicature to the basic sense of want (p.142). At the same time, it became possible to express the desired entity not only by a noun (I want a car) but also by an infinitive (I want to go) (p.144). Since Middle English there has been a general increase in infinitival complements. In the 19th and especially the 20th century the frequency of the verb want dramatically increases, both with a nominal and with a verbal complement (p.131). In Present Day English (American and British) examples are attested in which wanna is used with other senses than volition, even in epistemic uses. Consider (185) and (186), in which wanna indicates desirability, or (187) where it is used to express probability (from Krug 2000: 147ff):

(185) You’ve got tooth ache? You wanna see a dentist.
(186) You wanna turn right at the next corner.
(187) Customer: Do you have Coolers?
       Assistant: Coolers? They wanna be on one of the topshelves somewhere. They only arrived this morning.

The development of wanna is summarized in Figure 6-19. It is in accordance with the hypothesis.

The final modal constructions to be discussed are ought (to) and need (to). Ought to comes from the Old English verb agan (present ah, past ahbe), which meant ‘have’. In Middle English the verb changed to owe and it developed a past
form *ought* (Traugott 1992; Warner 1993). In late Middle and Early Modern English it develops auxiliary properties and combines with an infinitive (with and without *to*). The past form has no longer past time reference and becomes the normal form (Rissanen 1999: 232). As well as *ought* (*to*), *need* (*to*) develops auxiliary properties in late Middle and Early Modern English. In the 17th century it often combines with an infinitive without *to* and the third person –*s* is often not realized (Rissanen 1999: 232). Neither construction can be used for testing the hypothesis, since there is not enough information available on their exact semantic development.

6.3.5 Habitual

As discussed in 6.3.3, the forms *will* and *would* go back to a lexical verb *willan* meaning ‘want’ that developed into an auxiliary expressing volition. Already in Old English, *will* and *would* are occasionally used to express characteristic or habitual behavior in the present and past, respectively. This use is closely related to the earlier meaning of these verbs, volition, as one can infer that ‘what one wants to do, one is inclined or disposed to do’ (Bybee et al. 1994: 157). The developmental path is presented in Figure 6-20. This path is independent of the development of *will* and *would* as irrealis and future marker.

![Figure 6-20. Developmental path of will and would](Image)

This development is in line with the expected order. There is increase in scope.

The habitual expression *used to* is semantically close to its origin, the lexical verb *use*, which had as one of its meanings ‘to follow a usage or custom’. Around 1400 it became very frequent in a construction with a *to*-infinitive, which raised the meaning ‘be accustomed to.’ The construction was at first restricted to human subjects but from around 1600 it could also combine with inanimate subjects. It has furthermore extended to the combination with states (Bybee et al. 1994: 155-56). Still in early Present Day English, *use* *to V* could be used to express habituality in present and past. Probably in the 19th century the use as a present habitual marker died out (Denison 1998: 175). See Figure 6-21.
This development is in accordance with the hypothesis regarding the direction of change. Its first grammatical use, however, is immediately with medial scope. This is possible probably because the lexical construction is so close in meaning to the grammatical operator and the change is more a change in formal status than in semantic status.

6.3.6 Irrealis

In Old English, conditional clauses have the same general structure as in Present Day English, a subordinate if(gif)-clause (protasis) and a main then(ponne)-clause (apodosis). In general, the verbs have indicative inflection. However, the past subjunctive may be used to express imaginary and unreal conditions, in both the main and the subordinate clause (Traugott 1992: 257). In Middle English, the subjunctive becomes more frequent in conditional clauses instead of the indicative and the past subjunctive becomes the rule when the condition is hypothetical. The subjunctive pluperfect is used to indicate a counterfactual event. In the main clause, the subjunctive is often substituted by a past modal auxiliary (Fischer 1992: 350). The same situation holds in Early Modern English (Rissanen 1999: 308). Because of the loss of inflection, the past subjunctive forms become identical to the past indicative forms, except for be. Subjunctive past and pluperfect are still fairly common in the apodosis to encode hypothetical and counterfactual but more frequently a periphrasis with the modal auxiliaries should or would in combination with the indicative past and pluperfect are used to indicate irrealis. The replacement of the past occurs earlier than of the pluperfect (Rissanen 1999: 228-30). In Present Day English the unreal conditional apodosis—the then-clause—requires a past modal verb. As a modal verb is followed by an infinitive, past tense can only be marked by a perfect construction (Denison 1998: 137, 300). In American English, modal verbs are also commonly used in the protasis, such as in (188):

(188) If he would have left him he wouldn’t have gone through all that trouble.22

22http://www.storiestogrowby.com/stories/kids_say_elephpit.html children’s spelling and grammar is supposed to be revised.
The semantic changes of the components of the irrealis constructions, past tense forms, perfect forms and past modal forms, are not easy to describe, since they did not change their meaning in isolation, but within the construction of conditional clauses. The past and perfect forms replaced the subjunctive past and perfect that indicated irrealis, but to suggest that the meanings past tense or perfect aspect developed into the meanings of irrealis would be an oversimplification. The same holds for the irrealis senses of *would* and *should*, of which it cannot be claimed that they developed out of future senses: according to Bybee (1995) their irrealis sense can only arise because of the combination of past tense and modality, because this offers 'two areas of vagueness: (i) whether or not the predicate event was completed; and (ii) whether or not the modality remains in effect.' (p.506). A past modal may thus imply that certain conditions on carrying out the desired predicate were not met in the past and, as an inference, may not be met at all. This results in a hypothetical reading if the construction is used in a conditional context or in a polite reading if the conditions are implicit. As the genesis of the irrealis constructions is such a complex matter, in which several TMA expressions are involved, it will not be used as a test case for the hypothesis.

### 6.3.7 Discussion and conclusion

There is overwhelming support for the hypothesis. In as far as the diachronic developments of TMA expressions are known, they are in accordance with H1: Diachronically, operators show an increase in scope, and develop in the direction from $\pi_1$-operator to $\pi_2$-operator to $\pi_3$-operator. The aspectual expressions all stem directly from lexical constructions, the habitual markers originate in a lexical source and from participant-oriented modality (volition), and modal and future markers have begun their grammatical career as participant-oriented modal markers, climbing up to event-oriented modality or future tense and finally to epistemic, proposition-oriented modality. The Scope Hierarchy has thus made the correct predictions for English. TMA expressions show an increase in scope. They become more abstract and complex in time.

With respect to modal constructions, the combination with impersonal constructions, passives or generic subjects appears to have been a crucial step in the development of wider scope. Warner (1993: 172-73) suggests that modal markers in these contexts impose no selection restrictions on its subject and can therefore act as sentence modifiers, i.e., with wider scope. In a similar vein, Bybee et al. (1994: 287) claim that the first epistemic inferences of *may* occurred in contexts where the participant is non-specific (general pronouns) or absent (passive and stative sentences) as 'in these cases, the agent is practically devoid of semantic content'. Finally, Ziegeler (2003: 42) claims that there is a clear relationship between generic statements and prediction: a generic statement
may form the evidential justification for a predictive statement. All these claims support the view presented in Chapter 4 that event-oriented root- and deontic modality, which combine with non-specific participants, are important links to epistemic interpretations in the conceptual space. Especially in impersonal or generic contexts, epistemic inferences can arise which make the ‘metaphorical leap’ from the sociophysical to the mental world much more comprehensible.

6.4 FREQUENCY
This section will investigate the frequency of TMA operators in American English, more specifically, the questions 2a: What is the token frequency of TMA expressions?, and 2b: What is the size of different operator classes?

6.4.1 Token-frequency
The expected token frequency of operators is formulated in H2a:

H2a: The token frequency of $\pi_1$-operators is higher than or equal to the frequency of $\pi_2$-operators and the frequency of $\pi_2$-operators is higher than or equal to the frequency of $\pi_3$-operators.

It is assumed that wider scope operators are less relevant to communication than narrower scope operators and therefore, there discourse frequency will be lower.

6.4.1.1 Data-selection
Question 2a is examined in spontaneous speech of adults addressed to other adults in normal day conversation. The data are collected from the Santa Barbara Corpus of Spoken American English (Dubois 2000). The speech of eight different speakers in everyday conversation was selected (see Table 6-9). I have used several criteria for the selection of speakers from the corpus. Only speakers were selected who took part in conversations that were characterized as face-to-face conversation, preferably recorded in private homes. Only Phil is recorded in an office, but it is an informal face-to-face conversation. Every speaker was a monolingual and a native speaker of Standard American-English. No more than five participants took part in a single conversation. I furthermore tried to make the sample as representative as possible for adult Standard American English. Therefore, an equal number of female and male speakers were selected with a more or less equal average age in both sex groups. As a result of these criteria, only those speakers are selected whose age was known.
Table 6-9. Selected data for the adult conversations sample (SBCSAE, Dubois 2000)

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Name</th>
<th>Sex</th>
<th>Age</th>
<th>File(s)</th>
<th>No. of utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001</td>
<td>Lenore</td>
<td>f</td>
<td>30</td>
<td>Sbc0001 + Sbc0006</td>
<td>237</td>
</tr>
<tr>
<td>0003</td>
<td>Lynne</td>
<td>f</td>
<td>19</td>
<td>Sbc0001</td>
<td>585</td>
</tr>
<tr>
<td>0007</td>
<td>Pete</td>
<td>m</td>
<td>36</td>
<td>Sbc0002 + Sbc0003</td>
<td>370</td>
</tr>
<tr>
<td>0009</td>
<td>Marilyn</td>
<td>f</td>
<td>33</td>
<td>Sbc0003</td>
<td>475</td>
</tr>
<tr>
<td>0016</td>
<td>Darryl</td>
<td>m</td>
<td>33</td>
<td>Sbc0005</td>
<td>201</td>
</tr>
<tr>
<td>0028</td>
<td>Phil</td>
<td>m</td>
<td>30</td>
<td>Sbc0010</td>
<td>448</td>
</tr>
<tr>
<td>0044</td>
<td>Marci</td>
<td>f</td>
<td>50</td>
<td>Sbc0013</td>
<td>311</td>
</tr>
<tr>
<td>0046</td>
<td>Kevin</td>
<td>m</td>
<td>26</td>
<td>Sbc0013</td>
<td>268</td>
</tr>
</tbody>
</table>

Preferably only one speaker was selected from each conversation so that the topic of the conversations would not have too much influence, but this was not always possible because of the other criteria. The data of Lynn and Lenore and of Marci and Kevin were as a result taken from the same conversation. All speakers in the sample happen to be Caucasian, partly as a result of the distribution in the corpus, since the majority of monolingual speakers is Caucasian. The age of the few African American monolinguals was not available.

Table 6-9 presents the selected data from the corpus, specific information about the speakers and the number of utterances for each speaker. The complete sample consists of 2895 utterances but the number of utterances for each speaker is different.

6.4.1.2 Coding of utterances

The samples were coded for TMA expressions. Each TMA expression that was interpretable in the context was coded for its semantic function and scope, according to the classifications discussed in 6.2. Several utterances were coded for two or more TMA expressions. For example, aspectual periphrases also contain tense marking on the auxiliary be, have or keep, so that both aspect and tense are coded. The same holds for many (non-epistemic) modal expressions and the past habitual forms. Past modals that do not refer to past time but instead express politeness, irrealis or epistemic modality were not coded for tense. In non-verbal predicates (nominal, adjectival, adverbial or phrasal) the copula be was coded as a tense marker.
6.4.1.3 Results

The complete sample contains 2502 TMA expressions. As the samples of the different speakers are not equal in size, the number of TMA expressions for each speaker is also different (range: 167-630). In order to give each speaker an equal weight, the average percentages are not based on the overall absolute token frequency but rather on the relative frequencies within each speaker’s sample. The data from each speaker contribute one-eighth to the average distribution.

It appears that in American English adult-adult conversations the proportion of $\pi_1$-operators (16%) is much smaller than the proportion of $\pi_2$-operators (83%). The proportion of $\pi_3$-operators (1%) is smallest. This distribution is in contradiction to the prediction that $\pi_1$-operators will be more frequently used within a specific language than $\pi_2$-operators.

Which categories account for the percentages? In Table 6-10 the classes of operators are divided according to the broad TMA domains. From this table it appears that the frequency of $\pi_2$-operators is for the greater part the result of the predominance of tense operators in English, whereas other $\pi_2$-operators are far less common than $\pi_1$-operators. As discussed in 6.2.3 tense is an obligatory category in English, and many non-tense markers in English obligatorily co-occur with a tense marker. For example, an aspectual expression cannot occur without tense, whereas a tense marker can occur without an aspect marker. The obligatory expression of tense, in contrast to most other categories (except perhaps for progressive and perfect aspect) is such a strong factor of influence on the frequency, that the possible influence of scope is obscured. A second analysis was therefore performed on categories that are more similar to each other. Modal markers were selected for this, since the occur with all scopes, are semantically related, have similar expression forms

<table>
<thead>
<tr>
<th>TMA domain</th>
<th>$N = 2502$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect ($\pi_1$)</td>
<td>10.4</td>
</tr>
<tr>
<td>Participant-oriented modality ($\pi_1$)</td>
<td>5.6</td>
</tr>
<tr>
<td>Tense ($\pi_2$)</td>
<td>78.5</td>
</tr>
<tr>
<td>Event-oriented modality ($\pi_2$)</td>
<td>1.6</td>
</tr>
<tr>
<td>Event quantification ($\pi_2$)</td>
<td>0.4</td>
</tr>
<tr>
<td>Irreals ($\pi_2$)</td>
<td>2.4</td>
</tr>
<tr>
<td>Proposition-oriented modality ($\pi_3$)</td>
<td>1.1</td>
</tr>
</tbody>
</table>
and are not obligatory. For this category the influence of scope on frequency should probably be more directly visible. It appears that for this restricted comparison, the hypothesis does hold: of all modal expressions \((N = 228)\), participant-oriented modality accounts for 67%, event-oriented modality for 20% and proposition-oriented modality for 13%.

In sum, the hypothesis is too strong in that obligatory expression of a certain category has an enormous impact on the token frequency of such a category. However, when the comparison is restricted to expressions with related semantics, related expression form and related syntactic status (+/- obligatory), then the hypothesis does hold and the token frequency of operators with narrower scope is higher than of operators with wider scope.

### 6.4.2 Size of classes

The second question that relates to frequency of TMA markers is whether the size of operator classes correlates with scope. It is expected that:

\[ H2b: \text{The class of } \pi_1\text{-operators has more members than or an equal number of members as the class of } \pi_2\text{-operators and this class has more members than or an equal number of members as the class of } \pi_3\text{-operators.} \]

There are two reasons to expect that an increase in scope correlates with a decrease in the size of an operator class. During the process of grammaticalization (increase in scope), semantic paths tend to converge because their meanings generalize. This causes competition between different items with overlapping functions and probably leads to the use of expressions dying out. Secondly, it is expected that operators with narrower scope have more basic communicative functions. Because it is communicatively more important to make descriptive modifications (\(\pi_1\)) than to make modifications in situating the event (\(\pi_2\)) or modifications in the presentation of the content (\(\pi_3\)) there should be more different devices for doing so.

Whether this correlation between size of operator class and scope indeed exists for English can easily be seen from Table 6-6, 6-7 and 6-8. There are two ways in which this question can be answered: it is possible either to compare the number of semantic functions within an operator class or to compare the number of different expressions within an operator class. Both counts were made. In the analysis of expression forms, contracted forms were not counted as separate expression forms, but past forms of the central modals were considered as separate forms, since the present and past forms have sometimes developed different uses. Different tense or irrealis forms (irregular, regular, \(do\)-
did, forms of be, etcetera) are considered as one expression form, as their use is not dependent on semantic grounds, but purely on morphosyntactic grounds. Expression forms that have different semantic functions within the same domain are counted as one expression form. For example, ability, root-possibility and permission can as narrow scope modality are counted as one expression form. Similar expression forms in different domains (would as habitual, future and hypothetical marker) are counted as different forms. Table 6-11 presents the number of expression forms and semantic functions for each operator class.

Table 6-11. Number of expression forms and semantic functions per operator class in English

<table>
<thead>
<tr>
<th>Operator</th>
<th>π₁</th>
<th>π₂</th>
<th>π₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression forms</td>
<td>20</td>
<td>23</td>
<td>10</td>
</tr>
<tr>
<td>Semantic functions</td>
<td>15</td>
<td>16</td>
<td>6</td>
</tr>
</tbody>
</table>

It appears that H2b is not completely supported by the data. The class of π₂-operators is slightly larger than of π₁-operators. The class of π₃-operators is on the other hand smaller than the other two classes, which is in accordance with the hypothesis. This suggests that the communicative relevance and/or the cognitive complexity of the function of π₂-operators—modifying the relation of the event to the real world—is about equal to the function of π₁-operators—modifying the description of the event—whereas the function of π₃-operators—modifying the presentation of the propositional content—is in fact communicatively less relevant and cognitively more complex. In Chapter 7 this idea will be further examined to see if it is crosslinguistically valid.

6.5 SYNCHRONIC CONFIGURATIONS

What are possible synchronic configurations for TMA expressions? H3a predicts that the presence of more marked operators implies the presence of less marked operators.

H3a: The presence of operators with wider scope depends on the presence of operators with narrower scope according to the hierarchy: π₁-operator ⊂ π₂-operator ⊂ π₃-operator.
The Scope Hierarchy predicts that wider scope operators only occur in a language if there are also narrower scope operators. In other words, it is only possible to modify the presentation of the content by grammatical means (π3-operators) if it is also possible to modify the situating of the event by grammatical means (π2-operators), and this in turn is only possible if it is also possible to modify the description of the property or relation by grammatical means (π1-operators). As Tables 6-6 to 6-8 show, the configuration in English is clearly in accordance with this hypothesis, as there are operators of all three types.

The second hypothesis on the possible synchronic configurations of a TMA system concerns polysemous items. According to H3b polysemous expressions can only cover adjacent regions in semantic space.

H3b: A single expression can only cover adjacent regions in semantic space. As a result, a polysemous or portmanteau expression will only have semantic functions with similar or adjacent scopes.

The steps in the diachronic paths discussed in 6.3 are the presumed adjacent regions in semantic space. Along these paths it was indicated which meanings have become extinct, and in all cases, these extinct meanings were one or two at the beginning of the path. It was never the case that the meaning of an English TMA expression in the middle of the path became extinct with the earlier meanings being retained.

With respect to scope, the hypothesis implies that a particular item can only be used with similar or adjacent scopes, i.e., with scope over the predicate (π1) and over the predication (π2) or with scope over the predication and over the proposition (π3), but not with scope over the predicate and over the proposition, but not over the predication. All the TMA expressions in English that are used with different scope interpretations are presented in Table 6-12. Uses that are very infrequent are presented in between brackets.

The table shows that the predicted correlation between scope and polysemy holds. There is no grammatical TMA expression in English that may have scope over the predicate and the proposition, but not over the predication.

With respect to synchronic TMA systems, hypotheses H3a and H3b make correct predictions for English. Polysemous items cover adjacent meanings on a semantic path, which are assumed to be reflections of underlying universal conceptual relations. English has π1-operators for the least marked function of description, π2-operators for the more marked function of situating, and π3-operators for the most marked function of expressing the speaker's evaluation of the propositional content.
Table 6-12. Combinations of scopes of polysemous TMA expressions in English

<table>
<thead>
<tr>
<th>Expression</th>
<th>π1</th>
<th>π2</th>
<th>π3</th>
</tr>
</thead>
<tbody>
<tr>
<td>be supposed to</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>can</td>
<td>X</td>
<td>X</td>
<td>(X)</td>
</tr>
<tr>
<td>could</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>have to</td>
<td>X</td>
<td>X</td>
<td>(X)</td>
</tr>
<tr>
<td>have got to</td>
<td>X</td>
<td>X</td>
<td>(X)</td>
</tr>
<tr>
<td>keep—ing</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>may</td>
<td>(X)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>might</td>
<td>(X)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>must</td>
<td>(X)</td>
<td>(X)</td>
<td>X</td>
</tr>
<tr>
<td>need to</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ought to</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>shall</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>should</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>will</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>would</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

6.6 Expression Form

The fourth question—what is the expression form of TMA expressions?—will be examined in this section. The expected relation between form and scope is formulated in H4:

H4: Operators with wider scope show a higher or equal degree of formal grammaticalization than operators with narrower scope.

Roughly speaking, inflection is the most grammaticalized form, auxiliaries are less and periphrases the least grammaticalized (Bybee et al. 1994: 40). However, especially within the group of periphrases, finer distinctions are possible, such as the degree of reduction and of fusion of the component parts of the construction. In general, periphrases that contain bound elements are more grammaticalized than periphrases that only contain non-bound elements, and
periphrases of which the different components are phonologically reduced to one form are more grammaticalized than periphrases where this is not the case. As these finer grades of grammaticalization are hard to establish objectively, they will only be taken into account with great caution. Consequently, the question will only be examined in an exploratory way.

With respect to English, the TMA expressions with the highest degree of formal grammaticalization are the past and present tense inflection. The central modals have the second highest level of formal grammaticalization as they have become real auxiliaries. All other expressions are considered periphrases and are least grammaticalized. However, within this group, the aspectual constructions progressive, perfect and resultative are considered more grammaticalized than other constructions as they contain a suffix and their auxiliary is very often reduced. Furthermore, some of the periphrases are frequently contracted, such as be going to to gonna, have got to to gotta, have to to hafta, used to to usta, and want to to wanna. In the reduced forms, these constructions are closer to auxiliaries than to periphrases. However, as they are not in all contexts and not by all speakers reduced to such a high extent (see Krug 2000), they are still treated as (highly grammaticalized) periphrases and not as auxiliaries. Constructions with the lowest degree of formal grammaticalization are probably be about to, keep—ing and be supposed to.

Does scope indeed correlate with expression form? Table 6-13 shows the scope with which each expression is used, in relation to its degree of formal grammaticalization. For polysemous items, the least frequent use is put in between brackets. The table shows that in general, there is a tendency for wider scope operators to have a higher degree of grammaticalization, but there are quite some exceptions to this tendency. Most importantly, the items that are formally the most grammaticalized—tense inflection—function as operators with medial scope, whereas many auxiliaries or even periphrases function as operators with wide scope. Furthermore, two items that are formally the least grammaticalized can have medial scope (be supposed to, keep—ing), whereas some more grammaticalized items only have narrow scope.

A provisional conclusion may be that the hypothesis is too strong. When, however, only related meanings are compared, the picture becomes different. Within the related domains of tense, aspect and event quantification, the markers with the widest scope (tense inflection, habitual used to, will and would) have a higher or equal degree of formal grammaticalization than the markers with narrow scope (the aspectual periphrases). The only exception is keep—ing: it has the lowest degree of formal grammaticalization but can get a frequentative interpretation with scope over the predication (πι). Within the domain of modality, it is in general the case that operators that can have wider scope are
Table 6-13. Relation between expression form and operator type in English

<table>
<thead>
<tr>
<th>Degree of formal grammaticalization</th>
<th>Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \pi_1 )</td>
</tr>
<tr>
<td>High (Inflection)</td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>X</td>
</tr>
<tr>
<td>Present</td>
<td>X</td>
</tr>
<tr>
<td>Medium (Auxiliary)</td>
<td></td>
</tr>
<tr>
<td><em>may</em></td>
<td>(X)</td>
</tr>
<tr>
<td><em>might</em></td>
<td>(X)</td>
</tr>
<tr>
<td><em>must</em></td>
<td>(X)</td>
</tr>
<tr>
<td><em>will</em></td>
<td>(X)</td>
</tr>
<tr>
<td><em>would</em></td>
<td>(X)</td>
</tr>
<tr>
<td><em>should</em></td>
<td>X</td>
</tr>
<tr>
<td><em>could</em></td>
<td>X</td>
</tr>
<tr>
<td><em>can</em></td>
<td>X</td>
</tr>
<tr>
<td><em>shall</em></td>
<td>(X)</td>
</tr>
<tr>
<td>Medium-Low (Periphrasis)</td>
<td></td>
</tr>
<tr>
<td><em>ought to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>used to / usta</em></td>
<td>X</td>
</tr>
<tr>
<td><em>have to / hafta</em></td>
<td>X</td>
</tr>
<tr>
<td><em>have got to / gotta</em></td>
<td>X</td>
</tr>
<tr>
<td><em>need to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>wanna</em></td>
<td>X</td>
</tr>
<tr>
<td><em>be going to / gonna</em></td>
<td>X</td>
</tr>
<tr>
<td><em>be –ing</em></td>
<td>X</td>
</tr>
<tr>
<td><em>have –ed</em></td>
<td>X</td>
</tr>
<tr>
<td><em>be –ed</em></td>
<td>X</td>
</tr>
<tr>
<td>Low (Periphrasis)</td>
<td></td>
</tr>
<tr>
<td><em>be supposed to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>keep –ing</em></td>
<td>X</td>
</tr>
<tr>
<td><em>be about to</em></td>
<td>X</td>
</tr>
</tbody>
</table>

more grammaticalized (auxiliaries). However, there are two counterexamples to this picture: the auxiliary *shall* is restricted to narrow scope (desirability /
obligation, \( \pi_1 \) or to future tense (\( \pi_2 \)), whereas some of the less grammaticalized periphrastic constructions (have to, ought to, have got to) can have wide scope. However, shall is very infrequent and probably on its way to become extinct. With the exception of shall, all modal expressions with wide scope are formally grammaticalized to an equal or higher degree than modal expressions with medial or narrow scope. So, if only semantically related expressions are compared, there is indeed a strong correlation between scope and level of formal grammaticalization, although even then, there is one counterexample.

Although the domain restricted analysis supports the hypothesis, an alternative explanation might be possible, namely, the age of the grams. As was discussed in 5.4.4 there is probably a relation between age of grams and level of formal grammaticalization. Since it was also hypothesized that operators increase in scope, hypothesis 4 was stated in terms of scope and not in terms of age. However, the relation between scope and age of a gram is not straightforward. If grams change their scope, they will increase in scope, but it is not necessarily the case that grams keep changing their scope. Although the hypothesis abstracted away from this feature, for English, it may be worthwhile to investigate whether it is in fact age rather than scope, that is the prime predictor for formal grammaticalization. Scope may have only indirect influence, as it correlates to a certain extent with age.

What is the English situation? The oldest grams are the past and present tense. However, the meanings of these inflections have not drastically changed for centuries. The scope is the predication (\( \pi_2 \)) from the earliest records on English. The next oldest expressions are the central modals, which have grammaticalized in Old English, with the exception of can/could that grammaticalized in Middle English. Aspectual expressions are younger, especially the prospective and immediate prospective be about to. In Table 6-14 the relation between age and expression form is presented. The first period in which a particular construction clearly came to be used grammatically is taken as the age of the gram.

Table 6-14 shows that there is a very strong relation between age and expression form. For a few expressions, however, there is not enough information available on their diachronic development: from which period should they be considered grammatical? Be about to is said to occur from late Middle English, but it is unknown whether it was a grammatical construction at that time. The earliest lexical uses of the active verb suppose are from the 15th century, but it is unknown when the passive construction with this verb became grammatical. According to Ziegeler (2003), the passive construction occurs late, but she gives no exact dates. Furthermore, there is not enough information
Table 6-14. Relation between age and expression form in English

<table>
<thead>
<tr>
<th>First Important Period of Grammaticalization</th>
<th>Degree of Formal Grammaticalization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
</tr>
<tr>
<td>Before Old English</td>
<td></td>
</tr>
<tr>
<td>Past</td>
<td>X</td>
</tr>
<tr>
<td>Present</td>
<td>X</td>
</tr>
<tr>
<td>Old English</td>
<td></td>
</tr>
<tr>
<td><em>must</em></td>
<td>X</td>
</tr>
<tr>
<td><em>may / might</em></td>
<td>X</td>
</tr>
<tr>
<td><em>will / would</em></td>
<td>X</td>
</tr>
<tr>
<td><em>shall / should</em></td>
<td>X</td>
</tr>
<tr>
<td>Middle English</td>
<td></td>
</tr>
<tr>
<td><em>could</em></td>
<td>X</td>
</tr>
<tr>
<td><em>can</em></td>
<td>X</td>
</tr>
<tr>
<td>Early Modern English</td>
<td></td>
</tr>
<tr>
<td><em>have—ed</em></td>
<td>X</td>
</tr>
<tr>
<td><em>be—ed</em></td>
<td>X</td>
</tr>
<tr>
<td><em>be—ing</em></td>
<td>X</td>
</tr>
<tr>
<td><em>ought to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>need to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>used to</em></td>
<td>X</td>
</tr>
<tr>
<td>Present Day English</td>
<td></td>
</tr>
<tr>
<td><em>be going to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>have to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>have got to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>wanna</em></td>
<td>X</td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td><em>be supposed to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>be about to</em></td>
<td>X</td>
</tr>
<tr>
<td><em>keep—ing</em></td>
<td>X</td>
</tr>
</tbody>
</table>

available on *keep—ing*. So for the formally least grammaticalized constructions the information on age is lacking. This makes it impossible to test the relation between age and expression form thoroughly. Interestingly, the fact that the history of these constructions is not studied in much detail and the fact that
they are not highly grammaticalized may have the same cause: they are not frequently used. Probably, the level of formal grammaticalization is not only dependent on age, but also on the frequency of use during its history (cf. Bybee & Hopper 2001b).

In conclusion, there is a correlation between scope and expression form, but it does not account for all the facts. A different analysis, whereby the age of the expressions is related to the expression form, reveals a very strong relation for the expressions of which the age is known with certainty. The apparent relation between scope and expression form may therefore be a tertiary relation resulting from a relationship between age and scope and between age and expression form. As was already discussed in 5.4.4 there is no straightforward relation between scope and age of a gram: younger grams are earlier at the grammaticalization paths but not every step on a path leads to widening of scope.

6.7 Expression order

The final question to be examined in this chapter is: what is the expression order of TMA expressions? H5 predicts that:

H5: The expression order of TMA expressions iconically reflects the scope relations as follows: \(\pi_3 \pi_2 \pi_1\) Predicate \(\pi_1 \pi_2 \pi_3\)

As it is impossible to consider all the possible combinations of TMA markers in English, the discussion will be restricted to a few core examples.

In aspectual periphrases, the auxiliaries that carry tense, be and have, are farther away from the predicate than the part that signals aspect, -ing, -ed, going to/gonna and about to, consider (189)-(191):

(189) He was about to read.   > past \(\pi_2\) [immediate prospective \(\pi_1\)] [Pred]
(190) He’s gonna read.       > present \(\pi_2\) [prospective \(\pi_1\)] [Pred]
(191) He will be reading.    > future \(\pi_2\) [[Pred]-progressive \(\pi_1\)]

Foley & Van Valin (1984: 213) present some examples that combine different modal expressions. One is presented in (192):

(192) John may have to leave.

In (192), may can only be interpreted as a marker of objective (\(\pi_2\)) or subjective (\(\pi_3\)) epistemic modality and have to as a marker of participant-oriented necessity (\(\pi_1\)). It could be paraphrased as it is possible that John is forced to leave, or I doubt that
John is forced to leave. This meaning cannot be expressed by the opposite order: *John has to may leave. In the dialect of the southern United States, central modals can be combined, such as in the following examples taken from Foley & Van Valin (1984: 231):

(193) John might could do it.
(194) Don't get so far ahead - I may not could make it.

The only possible meaning of (193) is: 'it is possible that John is able to do it.' The only possible meaning of (194) is 'it is possible that I will not be able to make it.' Might and may express epistemic modality (π² or π³) and could expresses ability (π₁). This reflects the predicted order π₃ π₁ Pred or π₂ π₁ Pred. When modal expressions are combined, then, the expressions closer to the predicate have narrower scope than the expressions farther away from the predicate.

A further example is provided by Hengeveld (fc.: ex.13), in which a tense marker is combined with a π₃ modal and a lexical aspect marker:

(195) The tree must have begun to grow.

The order in (195) is the only possible order of 'certainty' must (π₃), past tense have (π₂) and ingressive aspect begin (lexical) and it iconically reflects the scope: π₃ π₂ Pred. All other combinations are impossible, such as must begin to have grown, has begun to must grow, begins to have must grow, etcetera.

Finally, when the habitual (π₂) and progressive (π₁) are combined, again the predicted order arises. See (196)-(197):

(196) Drug using prostitutes now complain about so-called low-life prostitutes who used to be bumming cigarettes who are now bumming condoms. (CWO)
(197) He used to be talking about whether he should run for president. (CWO)

The opposite order, be is being used to talk about … raises a completely different reading, in which be is literally used for the purpose of talking about ….

The only cases in which the expression order is not a clear reflection of scope relations seem to be when tense is expressed on modal markers or on the verb keep in keep –ing, because then, tense occurs in the middle of the expression. Consider (198):
(198) [about a dog] In his younger days, he could out bluff and move a herd of cattle very well.\(^{23}\)

Both the ability and the past tense meaning are expressed by the one full form could, therefore, no order can be established between the two. In forms like needed to, had to or kept -ing, the tense marker belongs to the inner part of the periphrastic expression and there is no reflection of the Scope Hierarchy. In a strict interpretation of the hypothesis, they might be considered counterexamples. However, a real counterexample would be the opposite order, where π1-operators would be farther away from the predicate than π2-operators. In sum, the general picture is that H5 holds for English. The expression order of TMA operators indeed reflects the scope relations.

6.8 CONCLUSION

Scope is clearly reflected in the English TMA system. The Scope Hierarchy makes correct predictions for the diachrony of grams, for the synchronic system, and for expression order. First, the developmental paths of TMA expressions only show an increase in scope. Second, polysemous expressions currently reflect the semantic relatedness of the steps in their developmental path and have only adjacent scopes. Third, the synchronic set of TMA expressions adheres to the implicational hierarchy between the operators. Finally, the expression order of operators reflects scope relations: operators with scope over the predicate are expressed closest to the predicate, operators with medial scope are expressed farther away and operators with scope over the proposition are expressed farthest away.

With respect to frequency and expression form, the hypotheses need some adjustments. Firstly, wider scope correlates with lower token frequency, but this only holds for TMA operators with related semantics and grammatical status. Obligatoriness of expression has more influence on frequency than scope. Secondly, the class of π2-operators is slightly larger than of π1-operators, whereas the opposite pattern was predicted. This suggests that operators that modify the description of the set of possible events and operators that modify the situating of the event are equally relevant categories in English. In contrast, π3-operators seem to be less relevant, as was predicted, since there are far less expression forms and semantic functions with wide scope. Thirdly, with respect to the expression form of operators, there seems to be only an indirect relation between expression form and scope, in that scope correlates to a certain extent with age of grams and the age of a gram, probably in combination with

\(^{23}\)ultimatehorsessites.com/temp/brokenbackranch/puppypage.html
frequency, seems to be a primary determining factor in the degree of formal grammaticalization.

This chapter has presented an overview of the English TMA system. The TMA system was continuously illuminated from the viewpoint of scope. Although some studies on English TMA have mentioned the role of scope, especially the work of Traugott (1982; 1989; 1995; 1996), it has not been taken into account systematically. Furthermore, the scope distinctions have only been broadly specified, for example a distinction between sentence or proposition scope versus scope over the verb or predicate. The definition of scope in FG is exceptional in that it makes a more fine-grained analysis of possible scopes than is done in other frameworks and specifies a clear link to cognitive and communicative factors. This approach has been shown to be fruitful in that it accounts for different characteristics of the TMA system. Needless to say that scope is not the only factor involved; it is only part of the TMA story.

One of the requirements of a functional theoretical model of language is that it can provide accurate descriptions of a specific language. For English, as shown in this chapter, the Scope Hierarchy has proved to be a rather reliable predictor. However, a further requirement is that a theoretical model is typologically adequate. It has to define the boundaries on possible language systems. Implicational hierarchies, such as the Scope Hierarchy, in general describe statistical tendencies, not necessarily absolute universals. The predictional power of the Scope Hierarchy therefore needs to be tested in a large sample of languages. The next chapter examines whether the assumed correlations between scope and properties of TMA systems have crosslinguistic validity.