On the interaction of tense, aspect and modality in Dutch

Richard van Gerrevink and Helen de Hoop[1]
Radboud University Nijmegen

1. Introduction

Consider the following fragment, taken from the Spoken Dutch Corpus (CGN) (boldface is ours):

(1) *Dus dat **moest betaald worden** want dat was achthonderd piek en die kerel zei*
so that must$_{IMP}$ paid become because that was 800 pop and that guy said
*dus dat zij het niet gedaan hadden dus Ilse **heeft** dat **moeten betalen**.*
so that they it not done had so Ilse has that must pay

"So that had to be paid because it was eight hundred guilders and that guy said that they did not do it so Ilse had to pay it.

In Dutch past tense can be expressed by an imperfective past form, e.g. *betaalde* 'paid', or by a present perfect form such as *heeft betaald* 'has paid'. Applied to the modal verb *moeten* 'must' in the fragment in (1), we thus can get the imperfective past form *moest betaald worden* 'had to be paid' as well as the present perfect form *heeft moeten betalen* 'had to pay'. The fragment above is about a broken window that had to be paid. The guy who lived in the apartment with the broken window does not want to pay the costs for the repair (*that guy said that they did not do it*: the third person plural includes the guy himself; *it* refers to breaking the window). As a consequence, Ilse had to pay it. The question that we will address in this chapter is why first the imperfective past form is used and then the present perfect to refer to the event of *het kapotte raam moeten betalen* 'to have to pay the broken window' in the past.

[1] We would like to thank our colleagues from the Optimal Communication research group for lively discussions on the topic of modality and tense interactions. We would also like to thank two anonymous reviewers for their useful comments. Helen de Hoop gratefully acknowledges the Netherlands Organisation for Scientific Research (NWO) for financial support.

We will argue that the crucial difference between the two grammatical tenses in combination with the modal verb *moeten* 'must' in Dutch is that only the perfective form gives rise to an actuality entailment. Thus, the sentence in (1) entails that Ilse actually did pay the broken window. The imperfective form does not have such an actuality entailment. That is, the phrase *dat raam moest betaald worden* 'that window had to be paid' does not entail that in the actual world somebody did indeed pay the window. We will analyse this difference in actuality entailment between an imperfective tense modal and a present perfect tense modal as the result of bidirectional optimization involving three potentially conflicting faithfulness constraints based on three interacting grammatical factors, *viz.*, modality, aspect, and tense. While a unidirectional optimization analysis cannot account for the relevant paradigm of forms and meanings in Dutch, we will argue that a bidirectional optimization can.

In section 2 we will first introduce and explain the notion of 'actuality entailment'. In section 3 we will turn to the interaction of the three factors that come into play in sentences such as (1) above, i.e., the modal verb *moeten*, perfective aspect, and past tense. With the help of bidirectional Optimality Theory we will demonstrate in section 4 how the interpretations resulting from the interaction between grammatical tense and modality arise. Section 5 will contain the conclusions of this chapter.

2. Actuality entailment

Hacquard (2006) starts her discussion of the relation between aspect and modality with the observation that perfective aspect places the running time of an eventuality within a reference time interval (provided by tense or adverbials of time). Imperfective aspect places the running time of an eventuality around the reference interval. Eventuality is used as a general term to refer to states, processes and events, cf. Bach 1986. In Hacquard's view, what aspect does is quantify over eventualities, it makes a connection between the time of reference and the time that the eventuality takes place. This can be illustrated by the sentence pair and their formal representations given in (2) below (the examples are adapted and translated into Dutch from Hacquard 2006):

(2) a. *Gisterochtend las ik een boek.*
yesterday-morning read_{IMP} I a book
"Yesterday morning, I read a book."
$\exists t \exists e[t \subseteq \tau(e)$ & I read a book(e)]

b. *Gisterochtend* *heb* *ik* *een* *boek* *gelezen.*
yesterday-morning have I a book $\text{read}_{\text{PERF}}$
"Yesterday morning I read a book."
$\exists t \exists e[\tau (e) \subseteq t$ & I read a book(e)]

The sentence in (2a) contains an imperfective past verb, the sentence in (2b) contains a perfective participle with a present tense auxiliary. In (2a) the event of reading a book has not come to an end in the time frame (t) of yesterday morning, whereas in (2b) the event of reading a book has come to an end within this time frame, implying that the subject finished the book yesterday morning.

With respect to modality, Hacquard (2006) takes the view that modal words, which are used to refer to possibilities and necessities, have the ability to go beyond directly observable facts and should therefore be analysed in terms of 'possible worlds' (cf. Kratzer 1981, 1991, Kaufmann et al. 2006). Each possible world represents an alternative of how the world could be. Hacquard posits that modal auxiliaries quantify over different sets of possible worlds. A sentence like *Jane must go to bed at 9 o'clock* states that in all possible worlds in which Jane's parents are obeyed, Jane goes to bed at 9 o'clock. According to Hacquard, the meaning of necessity in this example follows from the fact that the sentence is universally quantified: Jane goes to bed at 9 o'clock in all worlds in which she obeys her parents.

Hacquard stresses that it is important to note that the actual world does not have to belong to the set of possible worlds that the modal quantifies over. It can be the case that Jane is a disobedient girl and in fact never goes to bed at 9. This makes clear that modals enable language users to "talk about non-actual (but possible) situations by invoking other worlds than the actual one" (Hacquard 2006: 2).

The semantics for aspect and modality seem to work fine independently from each other, but Hacquard (2006) finds that they are problematic when we want to model the interaction between the two factors. In order to understand the problem, Hacquard uses the following sentence pair:[2]

[2] The examples in (3)-(4) are taken from Hacquard (2006) and are originally in French. We translated Hacquard's examples into Dutch, as exactly the same effects that she found for the original sentences in French are found for the Dutch translations.

(3) *Om naar de dierentuin te gaan,* *kon ik de trein nemen.*
in-order to the zoo to go can$_{IMP}$ I the train take
"In order to go to the zoo, I could take the train."

(4) *Om naar de dierentuin te gaan,* *heb ik de trein kunnen nemen.*
in-order to the zoo to go have I the train can$_{PERF}$ take
"In order to go to the zoo I could take the train."

A possible world analysis for sentence (3) says that there is a world among all possible worlds in which the subject goes to the zoo and in which the subject takes the train to get there. This does not entail that the subject in fact went there. It is perfectly possible that the subject in (3) was aware of the possibility to take the train, but decided to go by car in the end. The proposition in (3) would be true in both scenarios. Sentence (4), on the other hand, has an entirely different possible world semantic analysis. In order for that sentence to be true, the subject needs to have travelled to the zoo by train in the actual world. Stating that she did not go by train would create a contradiction. Hacquard calls this effect of the proposition having to hold in the actual world 'actuality entailment', following Bhatt (1999).

Hacquard (2006) defines actuality entailment as follows (following Bhatt 1999): a sentence is said to have actuality entailment if "the complement [is] forced to hold in the actual world" (Hacquard 2006: 3). This means that the implication that the complement of the modal expression holds in the actual world cannot be cancelled.

This actuality entailment does not only come up with possibility modals, as in (3)-(4). If we were to replace the forms *kunnen* 'can' by forms of *moeten* 'must', we also get actuality entailment in the perfective sentence (4). As we saw earlier with the Jane's bed time examples, Hacquard found that modal verbs themselves do *not* force their complement to hold in the actual world. The main question for Hacquard (2006) is how this effect of actuality entailment with modal verbs can arise in combination with perfective aspect, but not in combination with imperfective aspect? She concludes that the actuality entailment that arises in these perfective modal constructions is a side effect of the interplay between modality and aspect. At this point, we will perform some tests for actuality entailment of the modal verb *moeten* 'must' in various tenses (varying in aspect) in Dutch.

The tense-modality paradigm that will be examined in the present section is given in Table 1.

Table 1. The tense-modality paradigm of *moeten* 'must'

Tense	Form
Present tense	Hij moet dat betalen. he must$_{PRES}$ that pay "He has to pay that."
Present perfect tense	Hij heeft dat moeten betalen. he has that must$_{PERF}$ pay "He has had to pay that."
Imperfect past tense	Hij moest dat betalen. he must$_{PAST}$ that pay "He had to pay that."
Pluperfect tense	Hij had dat moeten betalen. he had that must$_{PERF}$ pay "He should have paid that."

For each of the sentences in Table 1 we will now determine whether they are forced to hold in the actual world or not. This is tested with the help of positively and negatively formulated descriptions of the state of affairs in the actual world. The results can be found in (5) below:

(5) [Present tense]

a. *Hij moet dat betalen.* [en hij doet het ook / maar hij doet het niet]
he must$_{PRES}$ that pay
"He has to pay that." [and he will do it / but he will not do it]

[Present perfect tense]

b. *Hij heeft dat moeten betalen.* [en hij heeft het ook gedaan /# maar hij heeft
he has that must$_{PART}$ pay
het niet gedaan]
"He has had to pay that." [and he has done it /#but he has not done it]

[Imperfect past tense]

c. *Hij moest dat betalen.* [en hij deed het ook / maar hij deed het niet]
he must$_{PAST}$ that pay
"He had to pay that." [and he did it / but he did not do it]

[Pluperfect tense]

d. *Hij had dat moeten betalen.* [#en hij heeft het ook gedaan /maar hij heeft het
he had that must$_{PART}$ pay
niet gedaan]
"He should have paid that." [#and he has done it /but he has not done it]

The example with present tense in (5a) combines with a positive as well as a negative description of the actual world. Hence, *hij moet dat betalen* 'he has to pay that' does not necessarily imply that he also will pay. Thus present tense deontics do not have actuality entailment in Dutch. Present perfect deontics as the one in (5b) on the other hand, do not allow for a negative description of the actual payment: *hij heeft dat moeten betalen* 'he has had to pay that' implies that he indeed has paid. Thus (5b) has actuality entailment. Imperfect past deontics as in (5c) combine with both types of descriptive sentences; this form does not require its complement to hold in the actual world so it does not have actuality entailment. We do not know whether the subject did or did not pay in the end.

Note that the pluperfect deontic in (5d) gives the most striking result. This form excludes a positive description of the state of affairs. The only fitting description here is a negative one: the complement is forced *not* to hold in the actual world, resulting in a non-actuality entailment. The only possible interpretation for this form is that the subject in fact did *not* pay.

3. Conflicting constraints

In the previous section we have discerned three factors that interact in the interpretation of modals. We will now try and determine for each individual factor what its precise role in the interaction is, starting with modality.

There is in fact little agreement, let alone a clear-cut definition accepted by everyone on the meaning or use of modality, but most scholars agree that modality principally involves the notions of possibility and necessity (cf. Van der Auwera and Plungian 1998). The type of modality in the fragment in (1) above is a clear case of deontic modality. Deontic modality involves external circumstances, which permit or oblige the participant to engage in the state of affairs, but these circumstances originate from some kind of social, ethical or personal norm. Some examples are given in (6):

(6) a. You may enter the plane now.
b. You must enter the plane now.

Sentence (6a) illustrates a case of deontic permission. Some authority figure, presumably a boarding agent at an airport in this example, allows the participant to enter the plane. In (6b) we are dealing with a case of deontic necessity: the participant is obliged to enter the plane, a possible reason may be that he has passed the gate and airport regulations dictate that once passed the gate, one cannot go back.

Narrog (2005) does provide a clear-cut definition of modality in terms of factuality:

> "Modality is a linguistic category referring to the factual status of a state of affairs. The expression of a state of affairs is modalized if it is marked for being undetermined with respect to its factual status, i.e. is neither positively nor negatively factual." (Narrog 2005: 184)

Although Narrog's factuality approach does not really give us any clues as to how we should analyze the exact meaning of 'must', its advantage is that it leaves no doubt whatsoever about what modality is all about. His definition dictates that an expression of a state of affairs is modalized when it is undetermined for its factual status, hence we should look at each instance of 'must' and try and determine what its effect is on the expression of the state of affairs. That is, each instance of the auxiliary *must* automatically should make an expression less factual than the non-modalized variant of that same expression.

Narrog's (2005) notion of factuality seems closely tied with the notion of actuality entailment that we found in the previous section while discussing Hacquard (2006). Recall that a proposition has actuality entailment when the complement is forced to hold in the actual

world. If a proposition is forced to hold in the actual world, it happens or has happened and therefore is a fact. Thus a proposition with actuality entailment displays positive factuality.

Narrog's (2005) view corresponds to the insight from formal approaches to modality in which modals invoke possible (accessible) worlds that enable us to talk about non-actual situations (cf. Kratzer 1981, 1991; Hacquard 2006). In section 2 we discussed the fact that a deontic modal reading of 'must' does not have actuality entailment. A sentence such as *Jane must go to bed at 9 o' clock* does not entail that Jane indeed goes to bed at 9 o' clock (ever) (Hacquard 2006). The same holds for the epistemic modal reading of 'must': If *must* universally quantifies over the set of possible (accessible) worlds in sentence (7a) below, then at first sight this would imply that Jane in fact *is* the murderer in the actual world.

(7) a. Jane must be the murderer.
b. Jane is the murderer.

However, the option that Jane is not the murderer in the actual world, is still open. One could say, for example, "Jane must be the murderer but she isn't". To some language users this sentence may sound infelicitous because of the explicit negation. This negation implies [-fact] and as such seems to be in direct opposition of the modal expression, especially when uttered by one and the same speaker. However, in a dialogic context, with someone other than the speaker replying 'No, Jane is not the murderer' to (7a), this apparent conflict between the modal expression and the negation is solved. Moreover, if the speaker would have been completely certain about Jane being the murderer, the statement in (7b) would have been more appropriate.

The same problem is also discussed in Kaufmann et al. (2006) who provide the following pair of sentences:

(8) a. It rained overnight.
b. It must have rained overnight.

As pointed out by Kaufmann et al. (2006), if the modal in (8b) universally quantifies over a set of worlds which includes the actual world, one might predict (8b) to entail (8a), but this prediction is not borne out. Rather, (8b) carries "an implication of uncertainty" that (8a) does not: "By using an epistemic necessity modal, a speaker is signaling uncertainty as to whether

(…) the actual world is one to which the most plausible explanation for the evidence at hand applies" (Kaufmann et al. 2006: 87-88).

Comparing the non-modal with the modal, the proposition in (8a) has shifted from having positive factuality in its non-modal form to an undetermined status in its modal form in (8b). This is also in accordance with the Dutch imperfect past tense modal we discussed in the previous sections. Modality thus adds uncertainty as to the factual status of the proposition. However, as also pointed out in the previous sections, the use of the present perfect tense seems to cancel this effect of modality. That is, the present perfect modal in sentences such as (1) and (5b) above clearly has a positive factuality status.

So far, the findings for the imperfect present and past tense corroborate Narrog's (2005) hypothesis that modalising the proposition in these tenses results in a shift from positive factuality into undeterminacy with respect to factuality. This is exactly what modality should do according to Narrog. An obvious problem for Narrog's hypothesis of modality's meaning is posed by the perfective tense. If modality adds uncertainty with respect to factuality status, how can it be that perfective aspect displays factuality? If we would follow Narrog (2005), the fact that this grammatical tense does not display undeterminacy with respect to factuality would mean that such a form should not be seen as modal. However, denying the possibility of a modal reading for perfective modal verbs means saying that a modal verb in perfective contexts has no effect on the meaning of the proposition and that this form therefore would mean exactly the same as its non-modal counterpart. This cannot be the case (as also pointed out by Hacquard 2006).

We observed that modalisation of imperfect past tense leads to undeterminacy of factuality for the proposition. This observation does make it very plausible that modality essentially adds an element of uncertainty or undeterminacy to the proposition. We will formulate this as a constraint:

(9) FAITHMODAL: A modal verb leads to undetermined factuality status.

In an Optimality Theoretic framework (Prince and Smolensky1993/2004), such a constraint is expected to be violable and potentially conflicting with other constraints. As Hacquard (2006) points out, perfective aspect seems to be able to 'neutralize' the effect of modality. We interpret this interaction between perfective aspect and modality in terms of a conflict between FAITHMODAL and another constraint involving perfective aspect.

The relation between perfective aspect and determinacy for factuality can be explained pretty easily. Boogaart (2007) argues that there is a difference between imperfective aspect on the one hand and perfective aspect on the other in that only the former is capable of expressing simultaneity with a certain reference point in time. By contrast, perfective aspect can only express precedence or succession of an eventuality with respect to a certain reference point in time. More specifically, perfective aspect denotes that the eventualities described are completed or have come to an end. In our view, when referring to an eventuality that has come to an end, a speaker automatically implies that she knows that the eventuality has actually taken place in reality. If an eventuality has come to an end in reality, it can only be described as a fact. One cannot use perfective aspect to describe an eventuality and leave the factual status of the eventuality undetermined at the same time. This inherent characteristic of perfective aspect explains why the present perfect tense, whether it is a modal form or not, yields a positive factuality score. Perfective aspect implies that the eventuality has come to an end and this comes with actuality entailment. Again, this can be formulated in a constraint, which we assume to be violable:

(10) FAITHPERFECT: Perfective aspect means the eventuality described is completed and thus a fact.

However, when this meaning of a completeness of the eventuality is borne out in the interpretation of present perfect tense, how can it be that this is not the case with a modal pluperfect tense? We have seen that an utterance like *hij had dat moeten betalen* 'he should have paid that' implies that no payment has been made by the subject. Despite the presence of perfective aspect, the event has not taken place. This is where the last factor comes into play: tense.

One of the most basic assumptions about past tense is that it is used to refer to eventualities or situations that have taken place in the past, before the moment of utterance. However, cross-linguistically, there are numerous examples where past tense marking is not used to refer to some eventuality that took place in the past. Janssen (1994), for example, discerns the following set of uses of the imperfect past tense in Dutch: advice; a diffident phrasing of an unlikely potentiality; a wish; politenes and quotative use. The following sentences are examples of these uses (the examples and translations are taken from Janssen (1994), the glosses are added by us):

(11) *Nou, maar ik vertrok morgen*!
well but I left tomorrow
"Well, but I would leave tomorrow!" [advice]

(12) *Wat denk je? Zou hij nog wel komen*?
what think you would he still PRT come
"What do you think? Will he come?" [unlikely potentiality]

(13) *Vertrok hij nu maar*!
left he now PRT
"If only he would leave." [wish]

(14) *Ik wilde graag een borrel*.
I wanted please a drink
"I would like a drink." [politeness]

(15) *Gisteravond vertrok je morgen en nu vertrek je overmorgen!*
yesterday-evening left you tomorrow and now leave you the-day-after-tomorrow
Wat moet ik nu geloven?
what $must_{PRES}$ I now believe
"Last night you were leaving tomorrow and now you say you are leaving the day after tomorrow! What am I to believe?" [quotative]

The past tense forms in examples (11)-(15) are all instances of an irrealis: the eventualities described have not actually taken place. It appears that the past tense can be used not only to refer to past time, but also to mark something as contrary to fact. Iatridou (2000) posits that these two uses follow from one and the same function of the past tense, namely that of exclusion. Past tense marking can range over time as well as over worlds. What it does is exclude the actual time or world from the topic time (i.e. the past tense use) or world (i.e. the contrary to fact use) respectively. Whenever this exclusion function is used, it is often implicated that the event or state under its scope does not hold at the moment of utterance, as shown by the example in (16) (taken from Iatridou 2000):

(16) John was in the classroom.

In (16) the simple past form *was* implicates that John is not in the classroom anymore. The state of John's being in the classroom is excluded from the actual time of reference. This implicature can, of course, be cancelled by a follow-up like *and he still is*. Hogeweg (to

appear) takes this analysis a step further. She argues that the use of past tense in a counterfactual construction is a grammaticalization of the observed pragmatic implicature that when John *was* in the classroom in the past, he *is not* in the classroom anymore in the present. Past tense marking implicates that the eventualities described are the case in the past, but are not the case presently, at the moment of utterance. This explains why counterfactual interpretations (or interpretations with negative factuality) arise for sentences with past tense such as pluperfect modal forms. In a sentence like *hij had dat moeten betalen* 'he should have paid that', the past tense form implicates that the event described is not true in the actual, present world. The modal pluperfect form shows negative factuality or, in other words, a non-actuality entailment. In order to explain this phenomenon we need the constraint in (17):

(17) FAITHPTI: Faith past tense implicature: the past tense implicature that the eventuality described is not true at the moment of utterance holds.

It should be noted that the feature [+fact] is not intended as the full semantics of perfective aspect and [-fact] is not intended as the full semantics of past tense. We only concentrate on a small aspect of the total interpretation of the sentence and merely make use of the effects of perfective aspect and past tense on the factuality status of the sentence. Furthermore, whether the eventualities described in a sentence take place in the past, present or future is an entirely different part of the overall interpretation process which we do not attempt nor desire to explain in the current chapter.

The three constraints proposed so far allow us to account for the factuality of modal forms in various tenses. For convenience, the findings with respect to modality, tense and factuality status, are summarized in Table 2:

Table 2. Factuality of the tense-modality paradigm of *moeten* 'must'

Tense	Form	Factuality status
Present tense	Hij moet dat betalen. he must$_{PRES}$ that pay "He has to pay that."	+/-
Present perfect tense	Hij heeft dat moeten betalen. he has that must$_{PERF}$ pay "He has had to pay that."	+
Imperfect past tense	Hij moest dat betalen. he must$_{PAST}$ that pay "He had to pay that."	+/-
Pluperfect tense	Hij had dat moeten betalen. he had that must$_{PERF}$ pay "He should have paid that."	-

The pattern of actuality entailment for modal expressions in different tenses that we found in the previous section leads to the pattern of factuality in Table 2 above. In Table 2, present tense and imperfect past tense modal forms did not have actuality entailment and thus display undeterminacy with respect to their factuality status. They get a score of +/-. Present perfect tense has actuality entailment and therefore has positive factuality. Pluperfect tense has non-actuality entailment and thus negative factuality: the proposition necessarily does not hold in the actual world. Please note that these factuality scores are the optimal interpretations in a neutral context where there is no other conflicting information present.

In this section we have examined the role of three separate factors in the interpretation of modal verbs in various tenses in Dutch. We found that tense, aspect and modality each have different, conflicting effects on the proposition. When comparing modal forms with their non-modal counterparts, we concluded that modality itself has an effect of creating undeterminacy with respect to the factuality status. With aspect on the other hand, we argued that the property of denoting completeness of the eventuality results in a positive factuality status. Finally, we showed that past tense marking induces negatively factual interpretations. In the next section we will model how speakers and hearers resolve the conflict between these different forces in order to produce and correctly interpret the various forms.

4. A bidirectional optimization analysis

As indicated in the previous section, there are three forces at work in the tense-modality paradigm in Dutch: tense, aspect and modality. The interpretations of the forms in the paradigm are the result of a conflict between the forces that are exerted by these different factors. A theoretical framework that has often been applied for resolving conflicts between different forces is Optimality Theory (OT) (Prince and Smolensky 1993/2004). OT explains language phenomena in terms of violable constraints. These constraints express general statements with respect to language and they can be in conflict with each other. The constraints are ordered in a constraint hierarchy on the basis of their strength. Constraints that are higher in the hierarchy should be satisfied more than constraints that are lower in the hierarchy. OT specifies the relation between the input and output. For each input, several possible output candidates are evaluated against the constraints. The output that satisfies the ranked constraints best emerges as the optimal output for the given input. Optimality Theory has been used first in the fields of phonology and syntax and has later been applied to semantics/pragmatics as well (Hendriks and de Hoop 2001, Blutner 2000).

In order to model the production and interpretation of the tense-modality paradigm in Dutch simultaneously, we will make use of bidirectional Optimality Theory. Bidirectional OT is an extension of unidirectional OT and it is used to integrate conversational implicatures (i.e. knowledge of the speaker about the hearer and vice versa) into the process of production and interpretation (cf. Blutner 2000, Blutner et al. 2006). In bidirectional OT a set of relevant, violable constraints is used to determine which form-meaning pair is optimal in the set of possible form-meaning pairs. In our data, there are three possible forms for which we wish to determine the optimal interpretations. We have an imperfective modal form (we will group present and past imperfective together in order to limit the complexity of the bidirectional OT tableau to follow), we have a present perfect modal form and we have a perfect past (pluperfect) modal form.[3] Possible features of the interpretations for each of these forms consist of a positively factual, a negatively factual and an undetermined factual reading (of course, the full interpretation of each form consists of much more than just its factuality status, but for our analysis only the factuality status is relevant). Each form pairs up with only

[3] Please note that we are able to group these imperfective forms together because we can reduce the difference between 'imperfective past' and 'imperfective present' simply to the semantic difference between past and present since we barely present any tense semantics (but mainly modal and aspectual semantics) in our model. By no means do we intend to reduce the full semantics of 'imperfective' merely to the feature [+/-fact].

one of the possible meanings and each possible meaning can only pair up once with a certain form. The entire paradigm of possible form-meaning pairs is given in Table 3:

Table 3. Possible modal form-meaning pairs

Modal form	(Relevant part of the) Meaning
Imperfective (present and past)	+ fact
	+/- fact
	- fact
Present perfect	+ fact
	+/- fact
	- fact
Past perfect	+ fact
	+/- fact
	- fact

Having determined the set of possible form-meaning pairs, we now need constraints which determine the optimal pairs of meaning and form. As we saw in the previous section, modality, perfective aspect and past tense each have an effect on the factuality status of the proposition. We saw that modalizing an utterance leads to an undetermined factuality status. This principle was called FAITHMODAL. The constraint FAITHPERFECT states that perfective aspect means that the eventuality has taken place, i.e. has positive factuality. Finally, we pointed out that this is not the case with the modal pluperfect form, which shows negative factuality. For this, we introduced the constraint FAITHPTI, which refers to the implicature of past tense that the eventuality is not true at the moment of utterance.

If we now combine the set of form-meaning pairs and the constraints in a bidirectional OT tableau, we can see how the optimal form-meaning pairs are derived from the constraints:

Tableau 1. Production and interpretation of the tense-modality paradigm in Dutch

Modal form	Meaning	FAITHMODAL	FAITHPERFECT	FAITHPTI
Imperfective	+ fact	*	(✓)	(✓)
✌ Imperfective	+/- fact	✓	(✓)	(✓)
Imperfective	- fact	*	(✓)	(✓)
✌ Present perfect	+ fact	*	✓	(✓)
Present perfect	+/- fact	✓	*	(✓)
Present perfect	- fact	*	*	(✓)
Past perfect	+ fact	*	✓	*
Past perfect	+/- fact	✓	*	*
✌ Past perfect	- fact	*	*	✓

In Tableau 1 we can see how our three constraints successfully determine which form-meaning pairs are bidirectionally optimal (indicated by the ✌ symbol). For the imperfective modal form, pairings with the positive and the negative factual reading are ruled out, because they violate the most important constraint FAITHMODAL. Since the imperfective modal form does not contain perfective aspect, the second constraint FAITHPERFECT is vacuously satisfied by all possible form-meaning pairs. The same goes for our third constraint: because we grouped present and past together here, we cannot determine whether FAITHPTI is satisfied or violated: present tense vacuously satisfies the constraint, while the combination of past tense with positive or undetermined factuality violates it and only past tense with negative factuality satisfies it. This does not really matter here, however, since the first constraint, FAITHMODAL, already picked out [imperfective, +/- fact] as the first superoptimal form-meaning pair.

The next superoptimal pair is [present perfect, + fact]. The +/- fact reading is the only reading that satisfies the highest constraint FAITHMODAL, and hence would become optimal in a merely unidirectional OT semantic approach (Hendriks and de Hoop 2001). However, in a bidirectional analysis, we see that the optimal +/- fact reading already paired up with the imperfective modal form, and hence it is not available anymore for the present perfect form. The next best option for this form is to pair up with the positive factual reading which satisfies the other high ranked constraint FAITHPERFECT, a constraint that would be violated by a

possible pairing with negative factuality. The FAITHPTI constraint is vacuously satisfied, because there is no past tense present in this modal form.

This leaves us with a third superoptimal pair remaining: [past perfect, - fact]. This pair, even though it violates the two highest ranked constraints, still satsifies FAITHPTI. The bidirectional optimization analysis can be performed on a concrete example, the tense-modality paradigm discussed above.

Tableau 2. Production and interpretation of the tense-modality paradigm in Dutch

Form	Meaning	FAITHMODAL	FAITHPERFECT	FAITHPTI
Hij moe(s)t dat betalen.	+ fact	*	(✓)	(✓)
✌ he must$_{PRES(PAST)}$ that pay	+/- fact	✓	(✓)	(✓)
"He ha(d)(s) to pay that."	- fact	*	(✓)	(✓)
✌ Hij heeft dat moeten betalen.	+ fact	*	✓	(✓)
he has that must$_{PERF}$ pay	+/- fact	✓	*	(✓)
"He has had to pay that	- fact	*	*	(✓)
Hij had dat moeten betalen.	+ fact	*	✓	*
he had that must$_{PERF}$ pay	+/- fact	✓	*	*
✌ "He should have paid that."	- fact	*	*	✓

In Tableau 2 we can see how our set of constraints determines the superoptimal form-meaning pairs. For the deontic imperfective forms the pairing with an interpretation that is undetermined for factuality comes out as superoptimal. This pairing satisfies the most important constraint FAITHMODAL, which dictates that the presence of a modal verb in the form should lead to an interpretation that is undetermined for factuality. Pairing up with a positively or negatively factual interpretation leads to a violation of the highest ranked constraint FAITHMODAL, thus these options are ruled out. This result is reflected by reality, because as we saw earlier in this chapter, imperfective forms combine with positive and negative factuality and thus have an undetermined factual status. This is also the correct reading for the past imperfective modal form that we encountered in the fragment in (1) above. The broken window *moest betaald worden* 'had to be paid', but this does not entail that it indeed was paid in the actual world.

Turning to the present perfect modal form *hij heeft dat moeten betalen* 'he has had to pay that' in Tableau 2, we see that combining with a positively factual reading results in a superoptimal form-meaning pair. Even though an undetermined factual status would violate the most important constraint FAITHMODAL the least, this option is ruled out because the undetermined factual interpretation already paired up with the imperfective forms. The positively factual interpretation results in the next best form-meaning combination. The event that is described by the modal form has been completed, satisfying the second most important constraint FAITHPERFECT, leading to a pairing up with a positively factual interpretation. A negatively factual interpretation would violate this constraint, since it would entail that the event described has not taken place and therefore has not been completed either. Again, this is in accordance with the intuition that we had with respect to the reading of the perfective modal form in the fragment in (1) above. The sentence *Ilse heeft dat moeten betalen* 'Ilse had to pay that' clearly entails that Ilse indeed paid it in the actual world.

Finally, the form-meaning pair of *hij had dat moeten betalen* 'he should have paid that' and a negatively factual interpretation is derived as super-optimal, satisfying the constraint FAITHPTI, while violating the two highest ranked constraints.

In this section we have constructed a bidirectional OT model which explains how the undetermined, positively and negatively factual interpretations pair up with the different deontic modal forms in Dutch.

5. Conclusion

We recognized three different factors that influence the interpretation of the modal verb *moeten* 'must' in Dutch: tense, aspect and modality. In order to explain how the various interpretations come up for the various modal forms, we first set out to determine for each of these factors separately the effect on the meaning of the proposition. We found that modality has the effect of leading to an undetermined factuality status. It is this characteristic of modality that ensures an undetermined factual interpretation for the imperfective modal forms. A conflicting force on the meaning of the proposition is exerted by perfective aspect. We argued that when the proposition has perfective aspect, this entails that the eventuality described by the proposition has come to an end or has been completed. When an eventuality is completed, this entails that it has taken place in the actual world and has a positively factual interpretation. We argued that it is this characteristic of expressing completeness of the eventuality that leads to a positively factual interpretation for present perfect modals. With

respect to the negatively factual interpretation for the pluperfect modals, we posited that it is caused by the past tense in this composite construction. We argued that speakers and hearers make use of the implicature that a proposition described in the past tense does not hold anymore in the actual world, leading to a counterfactual interpretation. We translated the three conflicting forces of tense, aspect and modality into three constraints: FAITHMODAL, FAITHPERFECT and FAITHPTI, and modelled the interaction between these three constraints with the help of bidirectional Optimality Theory in order to show how speakers and hearers deduce what form fits the meaning best and vice versa. The bidirectional optimization model correctly explained and predicted how three form-meaning pairs have become superoptimal in Dutch.

References

Van der Auwera, J. and V. Plungian (1998). Modality's semantic map. *Linguistic Typology 2*, 79-124.

Bach, E. (1986). The Algebra of Events, *Linguistics and Philosophy 9*, 5-16.

Bhatt, R. (1999). Obligation and Possession. In H. Harley (ed.), *Papers from the Upenn/MIT Roundtable on Argument Structure and Aspect, MITWPL 32*. Cambridge, MA: MITWPL.

Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics 17*, 189-216.

Blutner, R., H. de Hoop and P. Hendriks (2006). *Optimal Communication*. Stanford, CA: CSLI Publications.

Boogaart, R. (2007). The past and perfect of epistemic modals. In L. de Saussure, J. Moeschler & G. Puskas (eds.), *Recent advances in the syntax and semantics of tense, mood and aspect*. Berlin: Mouton de Gruyter.

Hacquard, V. (2006). On the interaction of aspect and modal auxiliaries. Paper presented at the Michigan Linguistics and Philosophy workshop.

Hendriks, P. & H. de Hoop (2001), Optimality Theoretic Semantics. *Linguistics and Philosophy* 24, 1-32.

Hogeweg, L. (to appear). What's so unreal about the past? – Past tense and counterfactuals.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry 31*, 231-270.

Janssen, T. (1994). Preterit and perfect in Dutch. In C. Vet and C. Vetters (eds.), *Tense and Aspect in Discourse*, 115-146. Berlin: Mouton de Gruyter.

Kaufmann, S., C. Condoravdi and V. Haizanov (2006). Formal approaches to modality. In W. Frawley (ed.), *The expression of modality*, 71-105. Berlin/New York: Mouton de Gruyter.

Kratzer, A. (1981), The Notional Category of Modality. In H.J. Eikmeyer and H. Rieser, *Words, Worlds, and Contexts*. De Gruyter, Berlin/New York, 38-74.

Kratzer, A. (1991). Modality. In A. von Stechow and D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der Zeitgenössischen Forschung*, 639-650. Berlin/New York: Mouton de Gruyter.

Narrog, H. (2005). On defining modality again. *Language Sciences 27*, 165-192.

Prince, A. & P. Smolensky (1993/ 2004). *Optimality Theory: Constraint interaction in generative grammar.* Oxford: Blackwell. Appeared also as Technical Report CU-CS-696-93, Department of Computer Science, University of Colorado at Boulder, and Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, April 1993.