The temporal perspective of epistemics in Dutch[1]

Annemarie VAN DOOREN — *University of Maryland, College Park*

Abstract. A series of experiments is conducted on naïve native speakers of Dutch and English to study the scope relation between tense and epistemic modality. The results are consistent with the claim that epistemics scope *over* tense (Stowell 2004, Hacquard 2006, a.o.), and challenge recent research that states that epistemics can, or must, scope *under* tense (von Fintel and Gillies 2007, Rullmann & Matthewson 2018): Dutch and English participants in a Truth Value Judgment Task judge sentences to be *false* when the past tense forms of the modals *have to* and *moeten* 'have to' are used to make an epistemic claim that held at a time before speech time, and *true* when they are used to make an epistemic claim that holds at speech time. Moreover, English participants in an Acceptability Judgment Task judge sentences to be *infelicitous* when the same past tense form of *have to* is used to make an epistemic claim that held at a time before speech time. Besides these general patterns, the results show variation within and across the two languages, which leads to interesting new questions about the interaction between tense and (epistemic) modality.

Keywords: modality, epistemics, tense.

1. Introduction

This paper is concerned with the interaction between tense and epistemic modals in Dutch and English. Epistemic modals such as *must* and *moeten* 'must' in (1) are used to express the notion of a likelihood. A series of experiments is conducted, the results of which are consistent with the relatively old claim that epistemic modals scope *over* tense (Groenendijk & Stokhof 1975, Iatridou 1990, Stowell 2004, Condoravdi 2002, Hacquard 2006, 2010, a.o.), and challenge more recent papers which argue that English and Dutch modals can or must scope *under* tense (von Fintel & Gillies 2007, Rullmann & Matthewson 2018).

(1) a. John ***must*** be home, since his car isn't in the parking lot. *epistemic*
b. Marie ***moet*** wel op vakantie zijn, want ze neemt haar telefoon niet op. *epistemic*
Mary must PRT on trip be because she takes her phone not on
'Mary must be on a trip, as she's not picking up her phone.

The debate revolves around the question where epistemic modals are interpreted relative to tense within their own clause. On the surface, modals seem to appear below tense, since they bear tense morphology. And indeed, *non-epistemic* modals, such as the deontics in (2) are interpreted in the scope of tense: The interpretations of the past tense modals in (2) are of obligations that held at some point in the past. Note that the semi-modal *had to* is used to demonstrate the interaction in English, as English modal auxiliaries such as *must* do not carry tense morphology; all Dutch modals however carry tense morphology.

[1] I would like to thank Valentine Hacquard, Anouk Dieuleveut, Ailís Cournane, Nick Huang, Dan Goodhue, Alexander Williams, and the reviewers and the audience at Sinn und Bedeutung 24. Many thanks also go to the ModSquad @ UMD/NYU, as well as the volunteers who participated in the Dutch experiment. This project is supported in part by NSF grant #BCS-1551628.

 In: M. Franke et al. (eds.)
Proceedings of Sinn und Bedeutung 24, vol. 1, pp.143–160. Osnabrück University.

(2) a. John ***had to*** be home by 10, since his parents told him so. *deontic*
b. Marie ***moest*** om 10 uur thuis zijn van haar ouders *deontic*
Mary must.PST at 10 hour home be from her parents
'Mary's parents obliged her to be home by 10.'

Some researchers have argued that *epistemic* modals in contrast outscope tense (Groenendijk & Stokhof 1975:68-69, Iatridou 1990, Stowell 2004, Hacquard 2006, 2010, 2011 and Hacquard & Cournane 2016, a.o.). The crucial datapoints involve again the interpretation of past tense epistemics, since two possible orderings of the scope-bearing elements are in principle available: The epistemic modal could scope *under* tense, giving a sentence like (3) the interpretation of a *past* likelihood (3i) (a past *temporal perspective*, Condoravdi 2002), or the epistemic modal could scope *over* tense, locating the evaluation time of the modal at the local time of evaluation, which in matrix sentences is speech time. The sentence is thus interpreted as a *current* likelihood (3ii) (a present *temporal perspective*).

(3) John ***had to*** be home, since his car wasn't in the parking lot.
Option i. 'Given what I *knew then*, it *was* necessary that John was home.'
Option ii. 'Given what I *know now*, it *is* necessary that John was home.'

Note that this second interpretation, which is the one argued for in this paper, is unexpected: While the past tense marker is *on* the modal, it does not backshift the evaluation time of the modal. The epistemic modal is evaluated at speech time. The tense marker is instead interpreted *under* the modal, shifting the event under the modal to the past.

What is at stake? The claim that epistemics scope over tense has been used to show that epistemic modals are structurally *high* in the clause: they take a large sized complement, including a TP (4) (Cinque 1999, Hacquard 2006, 2010), while *root* modals scope *under* tense and are thus structurally *low*. So, the interaction between tense and modality is used to argue that modals that can be used to express both a likelihood (an epistemic) and an obligation (a deontic), are in different structural positions depending on its use.

(4) ModEpistemic > Tense > ModRoot > Verb (Cinque 1999, Hacquard 2006, a.o.)

In contrast, Rullmann & Matthewson (2018), who argue that past tense epistemics scope *under* tense, propose that modals are uniformly in a structurally low position (5).

(5) Tense > ModP > Verb (Rullmann & Matthewson 2012, 2018)

In the next section the crucial datapoints are discussed that researchers have brought forward for the interpretation of past tense epistemics in (3). This discussion results in an empirical impasse: English datapoints that according to Hacquard (2006, 2010, 2016) support the view that epistemics scope *over* tense are argued to be 'marginal at best' by Rullmann & Matthewson (2018). Datapoints that according to Rullmann & Matthewson (2018) support the view that epistemics scope *under* tense are claimed to be special by Hacquard (2006, 2010, 2016) in that they involve *context shift*: under special circumstances, a past temporal operator in a higher clause can make it appear like the epistemic is evaluated in the past, but

in fact, the epistemic is still evaluated at the now shifted local time of evaluation. The impasse warrants a quantitative study using naïve speakers of English, and adding Dutch, in which the crucial datapoints can be tested. Two initial studies are described in section 3. The conclusion follows in section 4.

2. The debate

The two claims about the structural height of epistemic modals make different predictions: If epistemics are structurally *below* the tense marker in their clause, the evaluation time is the time expressed by tense. For past tense epistemics, this means that they are evaluated at a time before speech time (a *past* evaluation time, Table 1). If epistemics are structurally *above* tense, the evaluation time is the local time of evaluation. For past tense epistemics in *matrix contexts*, this means that they are evaluated at speech time (a *present* evaluation time); in *embedded contexts*, the local time of evaluation may be in the past. Arguments in favor of each claim are discussed below.

Table 1: Possible interpretations of epistemics

	Evaluation time of the epistemic	Past epistemic in matrix context	
		Past evaluation time	**Present** evaluation time
Tense > Epistemic (Rullmann & Matthewson 2018)	*Time expressed by tense*	+	-
Epistemic > Tense (Cinque 1999, Hacquard 2006, a.o.)	*Local time*	-	+

2.1. Epistemics scope *over* tense

The view that epistemic modals scope *over* tense can be found in Groenendijk & Stokhof (1975:68-69), Iatridou (1990), Stowell (2004), Hacquard (2006, 2010, 2011) and Hacquard & Cournane (2016), a.o. Stowell's example is in (6); a context that helps bring out the intended reading is two speakers discussing how many people were at a party last night. The epistemic modal seems to be evaluated *at speech time*: Given the evidence available at the time of speech, it *is* necessary that there were at least a hundred people at that party last night.

(6) There **had to** be at least a hundred people there.
'There must have been at least a hundred people there.' (Stowell 2004:626)

It could be argued, however, that in this particular example, the evaluation time is in the past and continues to hold: At some time in the past, it *was* necessary given the evidence that was available then that there were at least a hundred people there (Valentine Hacquard, p.c.; for a similar reasoning on different examples, see Rullmann & Matthewson (2018:326)).

In the example in (7) from Hacquard (2010), this analysis is not available, as there is an explicit contrast between a past and a present evaluation time: In the past, Poirot thought that Mary was home at the time of the murder, but more recently, he established that she was home. The question is, can *had to* in this context target the past evaluation time, which would make the sentence true? Hacquard claims the sentence in (7) with the past tense modal is

judged as false, which supports the claim that *had to* cannot be used to make an epistemic claim that held at a past time, i.e., the epistemic modal cannot scope under tense.

(7) *(Context:) Imagine that the evidence gathered at the beginning of a murder investigation (a week ago) pointed to Mary being home at the time of the murder: both Mary and her roommate testified that they were having lunch together there. Yesterday however, Poirot established that Mary's roommate had lied, as she was seen by several eyewitnesses elsewhere at that time, debunking Mary's alibi.*
Mary **had to** be home (at the time of the crime).

Hacquard & Cournane (2016) furthermore claim that there is a contrast between epistemic *modals*, such as *have to*, and epistemic *verbs,* such as *seem*: Only epistemic *modals* scope over tense. Hacquard & Cournane first set up a context in which there is again both a past and a present evaluation time, as in (8). They then contrast an epistemic modal claim using *seemed* (8a) with *had to* (8b) and state that while (8a) is false in this context, (8b) is true. From this they conclude that *had to* can target a present evaluation time.

(8) *(Context:) Al has been a prime suspect for a crime that occurred last night in Montreal. Up to now, all of the evidence pointed to him being in Montreal last night. But just now, the detective receives fresh evidence that proves that Al was in fact in DC last night.*
a. It *seemed* that Al was in DC last night/ Al *seemed* to be in DC last night.
b. Al *had to* be in DC last night. (Hacquard & Cournane 2016: 4)

Together, the examples in (6)-(8) support the claim that epistemic modals *can*, and in fact *must* scope over tense. What is more, the contrast with *seem* shows that this scope relation is specific to epistemic *modals*, as opposed to epistemic *verbs*: In the exact same set-up, epistemic *modals* target a present evaluation time, while epistemic *verbs* do not.

2.2. Epistemics scope *under* tense

Von Fintel & Gillies (2007) and Rullmann & Matthewson (2012, 2018) argue instead that epistemics *can* (Von Fintel & Gillies 2007), or *must* (Rullmann & Matthewson 2012, 2018) scope under tense. Support is provided by further datapoints, in which epistemic modals appear to have a past evaluation time (section 2.2.1.), and by a re-evaluation of the datapoints brought up in the previous section (section 2.2.2.). The disagreement about which datapoints should be used to show the interaction between epistemic modals and tense, and what the judgment is for these datapoints, demonstrate the need for the experiments in section 3.

2.2.1. Further datapoints

Rullmann & Matthewson (2018) respond to the examples in which a past evaluation time does not seem to be available by stating that while a past evaluation time for past epistemics might be dispreferred, it is available in natural speech and in construed examples (9)-(10). In (9a), for instance, the epistemic claim about there being at least a hundred people seems to

hold *before* speech time. Rullmann & Matthewson (2018:284) argue that these examples show that in Dutch and in English,[2] 'typically tense scopes above the modal'.

(9) a. When Susan arrived at Bob's house, she saw that the place was packed. There **had to** be at least a hundred people there. But she found out later that actually, there were only 60. (Rullmann & Matthewson 2018:298)
b. This morning I opened my phone bill and was shocked when I saw that I owed $10,000. This **had to** be a mistake! Unfortunately, it turned out to be correct. My husband had used my phone on his latest trip to Papua New Guinea, forgetting about the roaming charges. (Rullmann & Matthewson 2018:297)

(10) *(Context:) I was looking for Jan last night. I had searched all his usual haunts except his house and hadn't found him yet.*
Jan **moest** wel thuis zijn.
Jan must.PST PRT home be
'John had to be home.' (Rullmann & Matthewson 2018:285)

Boogaart (2002) notes that past epistemics can occur in such contexts but states that the contexts are marked in that they involve *free indirect discourse*: a discourse in which the perspective is shifted to one of the agents in a story, without it being overtly marked. Hacquard (2006, 2010, 2016) follows Boogaart's proposal and analyzes sentences like (9) on a par with cases of embedded modals (11). She claims that in these cases, epistemics still outscope tense and are as such evaluated at the local time of evaluation (Table 1): The local time of evaluation, however, has been shifted to the past. The epistemic modal is evaluated at the time of the past discourse (9)-(10) or the past tense embedding verb (11).[3]

(11) Two days ago, Poirot thought that Mary **had to** be the murderer. (Hacquard 2011: 28)

Evidence for the shifted time in cases of free indirect discourse comes from deictic temporal adverbials like *now*, which can refer to the narrator's now in sentences like (9b), as in (12), which is past relative to utterance time (Hacquard 2016:57). The same argument can be made for the Dutch example in (10), shown in (13).

(12) *(Modification of (9a)):* This morning I opened my phone bill and was shocked when I saw that I owed $10,000. **Now,** this had to be a mistake! […]
(13) *(Same context as (10))* Jan moest **nu** (wel) thuis zijn.
Jan must.PST now (PRT) home be
'Jan had to be home now.'

A second type of examples presented in favor of epistemic modals being able to have a past evaluation time (scoping *under* tense) is in von Fintel and Gillies (2008). The sentence in

[2] Rullmann & Matthewson (2018) also discuss the non-Indo-European languages St'át'imcets and Gitksan, for which they claim that epistemic modals scope under tense as well. These languages will not be discussed here.
[3] Note that Rullmann & Matthewson (2018) argue that in embedded contexts, epistemics can be further backshifted. This paper focuses on epistemics in matrix contexts but these cases will be evaluated in the future.

(14) seems to mean that at a certain point before utterance time, Sophie thought it was a possibility that there was ice cream in the freezer.[4]

(14) *(Context:) Sophie is looking for some ice cream and checks the freezer. There is none in there. Asked why she opened the freezer, she replies:*
There **might have** been ice cream in the freezer.

Hacquard (2006, 2010) notes that this possibility only arises in *why* questions, and analyzes these cases as instances of a covert embedding attitude (15a) (Hacquard 2006), making it similar to the sentences in (11), or a covert *because* (15b) (Hacquard 2010), which is known to be able to shift perspectives (Stephenson 2008).

(15) a. ~~I thought that~~ there **might have** been ice cream in the freezer.
b. ~~Because~~ there **might have** been ice cream in the freezer.

Rullmann & Matthewson (2018:324) support von Fintel and Gillies' analysis and argue against Hacquard's solutions in (15), as neither type of context leads to a past evaluation time of a modal by itself: They observe that in both contexts, only *might* with an embedded perfect (*might have*), but not *might* alone can give rise to a past evaluation time (16a), (17). Note moreover that this is a possibility with an *overt* embedding (16b). Valentine Hacquard (p.c.) responds that perhaps the past evaluation of *might* in (16a) and (17) is dispreferred because of the competition with *might have*.[5]

(16) a. ~~I thought that~~ there **might be** ice cream in the freezer.
#"At a certain time in the past, it *was* possible that there was ice cream." #past
b. I thought that there **might be** ice cream in the freezer.
"At a certain time in the past, it *was* possible that there was ice cream." past

(17) I looked in the freezer because the ice cream **might be** in there.
#"I looked in the freezer because at a certain point in the past, it *was* possible that there was ice cream." (Rullmann & Matthewson 2018:324) #past

So far, the disagreement between researchers who claim epistemics scope *over* or *under* tense has been about past tense epistemics in narrative contexts and in answers to *why* questions. While there is agreement that in these contexts, past tense epistemics can be used to express an epistemic claim that held before speech time, there is disagreement about what this shows: For Rullmann & Matthewson (2018) and von Fintel & Gillies (2008), it shows that past tense can scope over epistemic modals, while for Hacquard (2006, 2016) it shows that in special cases, a higher temporal operator can backshift the local time of evaluation, which is the

[4] The past evaluation time of the epistemic in (14) could either arise from the perfect raising over the modal at LF (see Condoravdi (2002) for this analysis for non-epistemic modals), or it could be lexically encoded in the multi-word item *might have* (Rullmann & Matthewson 2018).

[5] Another possibility is that for the past evaluation time to be available in (14), an *overt* past tense marker is necessary, either in the form of a perfect marker, as for English *might have*, or a past tense. Initial support for this idea is that the Dutch past tense modal *kon* 'can.PST' with (ia) and without an underlying perfect (ib) can have a past evaluation time as a response to a *why* question.

(i) a. Het ijs kon daar toch **zijn geweest**. past
the ice can.PST there PRT be been
'The ice cream might've been there, right?'
b. Het ijs kon daar toch **zijn**. past
the ice can.PST there PRT be
'The ice cream might've been there, right?'

epistemic modals' time of evaluation. The question that naturally follows is whether outside of these contexts, epistemics scope *over* or *under* tense.

2.2.2. Re-evaluation of datapoints

Outside of narrative contexts and *why* questions, Hacquard (2006, 2010, 2016) and Hacquard & Cournane (2016) discuss sentences in which they claim past tense epistemics are evaluated at speech time (section 2.1.). Rullmann & Matthewson (2018) however claim that sentences like the ones in (18) are judged 'marginal at best', which casts doubt on the claim that epistemics scope over tense.

(18) *(Context:) Up until just now, all of the evidence pointed to Mary being home last night. But now, fresh evidence proves that Mary's home was empty last night.*
Mary **had to** be out last night. (Rullmann & Matthewson 2018:326)

Rullmann & Matthewson ran a survey on 8 native speakers of English who judged sentences like (18) on a scale from 1-3. They report that 3/8 judged (18) as 'marginal'. What is more, only 3/8 speakers accept Stowell's original sentence from (6) in context (*There **had to** be at least a hundred people there*). They furthermore constructed their own example (19), which is judged as infelicitous or marginal by 6/7 speakers they consulted.[6]

(19) *(Context:) A mother is wondering what her son got up to at a party last night. He emerges from his room holding his head and looking green. She says:*
You **had to** be drunk. (Rullmann & Matthewson 2018:300)

The contradicting claims about the datapoints are intriguing, and call for further investigation. The experiments described in section 3 complement Rullmann & Matthewson's quantitative results in three ways. First, does the low acceptability for (18) reflect an unavailability of past epistemics to be evaluated at speech time, or a mere dispreference? What could influence the judgment of (18) is a preference for sentences like *Mary must have been out/Mary has to have been out* (or even *Mary was probably out*) in these contexts. Here a Truth Value Judgment Task is conducted, which tests whether sentences like (18), while perhaps dispreferred, have an interpretation available in which the epistemic modal is evaluated at speech time. Secondly, while Rullmann & Matthewson (2018:297-298) ran a second survey showing that past tense epistemics are accepted when they make an epistemic claim that held before speech time *in narrative contexts*, it has not been shown that the same holds *in matrix contexts*. This condition is tested here. Finally, while the quantitative results reported so far all involve English, the claim that epistemic modals scope under tense has also been made for Dutch. This language is tested in exactly the same conditions as English.

[6] See Goodhue, Hacquard & Williams (in progress) for an analysis that the use of *have to* (vs. *must*) requires special contexts, which might be responsible for the infelicity in (19).

2.3. Summary of the debate: An empirical impasse

This summary of the debate on the interaction between epistemic modals and tense justifies conducting a quantified experiment on past tense epistemics: In contexts in which there arguably are no additional context shifters, there is disagreement about the judgments themselves. The experiments described in the next section test whether both Dutch and English epistemic modals *can* or *must* scope over or under tense.

3. Experiments on past tense epistemics

Two experiments were run online on native naïve speakers of English and Dutch. Participants were recruited using Amazon Mechanical Turk (https://www.mturk.com) for English and Facebook groups (www.facebook.com) for Dutch. The first experiment was a Truth Value Judgment Task, in which participants were asked whether a sentence was considered true or false given a preceding context. The second was an Acceptability Judgment Task, in which participants were asked whether a sentence sounded fine or not within the given context.

3.1. Truth-Value Judgment Task – Design and methods

For the Truth-Value Judgment Task (TVJT), each trial consisted of a context and a test sentence, as illustrated in (20). Participants were asked whether the sentence in bold was *true* or *false* given the context. Each context presented evidence at two different points in time, past and present, and was followed by one of four types of test sentences, as shown in (21). Judgments on the critical test sentences (21a, 21c), determine whether the epistemic modal scopes *over* or *under* past tense: If the epistemic modal scopes ***over*** tense, sentences containing *had to* should be judged true when they target a *present* evaluation time (21a), and false when they target a *past* evaluation time (21c). If the epistemic modal scopes ***under*** tense, the judgments for (21a) and (21c) should be reversed: sentences with the modal should be judged false when they target a present evaluation time (21a), and true when they target a past evaluation time (21c). The predictions following from the two claims are in (21) in grey.

(20) *Al has been a prime suspect for a crime that occurred last night in Montreal. Up to now, all of the evidence pointed to him being in Montreal last night. But just now, the detective receives fresh evidence that proves that Al was in fact in DC last night.*
The detective says: **Al had to be in DC last night.**

(21) a. The detective says: Al *had to* be in DC last night. (epi>T: true, T>epi: false)
b. The detective says: Al *seemed* to be in DC last night. (false)
c. The detective says: Al *had to* be in Montreal last night. (epi>T: false, T>epi: true)
d. The detective says: Al *seemed* to be in Montreal last night. (true)

The controls are the same sentences with *seemed* (21b, 21d). Since *seem* uncontroversially scopes under tense (8), sentences with *seemed* are expected to be judged true when they target a past evaluation time (21d), and false when they target a present evaluation time (21b). The two factors tested in the TVJT are summarized below.

(22) Lemma: *had to/moest* 'had to' vs. *seemed/leek* 'seemed'
Target evaluation time: PRES vs. PAST

Following Hacquard & Cournane (2016), this set-up arguably does not contain context shifters (section 2) and as such, there should be agreement about what the results would show. The contexts used for the Dutch version of the experiment were the same; the epistemic modal in the test sentence is *moest* 'had to', and the epistemic verb is *leek* 'seemed'.

Materials Eight contexts similar to (20) were created. For each context, there were four possible test sentences, depending on (1) the lemma tested (*seem/leek* vs. *have to/moeten*) and (2) whether the sentence targeted a past or present epistemic evaluation time. Note that the particle *wel* 'yet' was present in the Dutch *moest* 'had to' sentences but not in the *leek* 'seemed' sentences. It is possible to get an epistemic interpretation for *moest* without *wel*, but it was added to make sure the participants understood the modal epistemically (and not deontically).

In order to avoid a narrative context (see section 2), direct discourse was used in the test sentences, and present tense was used in the context and the question asked. The question asked after each test sentence was 'Is the sentence in bold true or false?' Two further comprehension questions were asked about the contrasting claims in the context after each test question, which measured the attentiveness of the participants. The questions following (20) were 'Where did the detective originally think Al was?', and 'With the new evidence, where does the detective now think Al was?'. The experiment contained all eight contexts, with each of the four conditions in (21) tested twice. The factors *lemma* and *target evaluation time* were counterbalanced across participants.

The experiment additionally contained one training item and eight fillers, which consisted of similar contexts with a past and present evaluation time, and used epistemic adjectives and adverbs to keep participants focused on epistemic claims, in the present or future tense. An example is in (23); all test items and fillers can be found at http://www.annemarievandooren.com/papers-2/.

(23) *Two friends are betting on which sports team is going to win. Fred thinks the red team is going to win, as they have the better defense. But Paul informs him that the red's team best defender is out with an injury and the rest of the team can't win without him.*
Fred says: **The red team is probably not going to win.**
Filler question: 'Is the sentence in bold true or false?'
Comprehension question 1: 'Did Fred originally think the red team was going to win?'
Comprehension question 2: 'With the new evidence, does Fred now think the red team is going to win?'

Participants 40 native speakers of Dutch and 40 native speakers of English based in the United States participated in this study. English participants were recruited through Amazon

Mechanical Turk and paid for their participation. Dutch participants were recruited through e-mail and Facebook Groups targeting Dutch teachers and university students and volunteered.[7]

Procedure Before starting the experiment, there was one training item with feedback. Participants then judged 16 items in total, eight test items plus eight fillers. Since the 40 participants in both the Dutch and the English experiment were given eight test questions, each of the four conditions was judged 80 times.

Data analysis The responses of the participants were analyzed using a two-tailed binomial test (R Studio Core Team, 2008) to investigate whether the proportions of answers with 'true' are higher or lower than expected by chance.

3.2. Truth-Value Judgment Task – Results

In the English experiment, accuracy on comprehension questions for both trials and fillers is high (mean on all contexts: 86.2% correct, with no differences between the 16 contexts). One participant who performed under 75% on the comprehension questions was excluded. Accuracy on fillers is also very high (mean on all fillers: 94.7% correct).

In the Dutch experiment, accuracy on comprehension questions for both trials and fillers is very high (mean on all contexts: 93.8% correct, again with no differences between the 16 contexts). One participant who performed under 75% on the comprehension questions was excluded. Accuracy on fillers is lower than in English (mean on all fillers: 69.7%[8]).

The main results for English are shown in Figure 1. English sentences with *had to* are judged true 89.7% of the time, when the target is a *present* evaluation time, and 6.4%, when the target is a *past* evaluation time. Sentences with *seemed* are judged true with a *present* target 84.6% of the time, and 47.4% of the time with a *past* target. A binomial test (two-sided) indicates that the proportions of the responses in the present/*had to* and present/*seemed* condition are higher than expected by chance, and in the past/*had to* condition lower than expected by chance (0.5) (Table 2).

The main results for Dutch are shown in Figure 2. Sentences with *moest* 'had to' are judged true 64.1% of the time when the target is a *present* evaluation time, and 23.1% of the time when the target is a *past* evaluation time. Sentences with *leek* 'seemed' are judged true 46.2% of the time with a present target and 74.4% with a past target. A binomial test (two-sided) indicates that the proportions of these responses in the present/*moest* and past/*leek* condition are higher than expected by chance, and in the past/*moest* condition lower than expected by chance (0.5) (Table 2).

[7] The Dutch experiment was originally launched on Amazon Mechanical Turk, but there were no responses.

[8] There was also a high amount of variation: While 13/40 participants had a perfect score on the eight fillers, 17/40 performed under 75%. I leave it for further investigation to examine more closely potential differences between the Dutch and English fillers, which lead to the lowered performance of some of the Dutch participants.

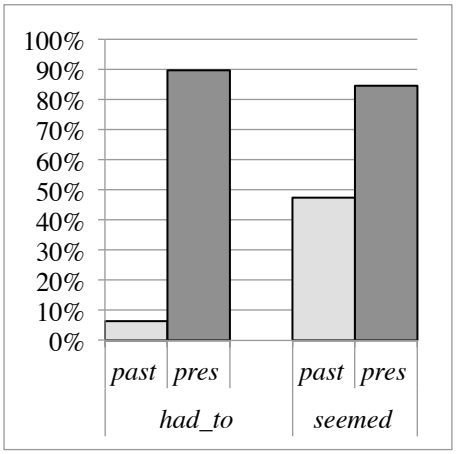

Figure 1: Proportions of 'true' answers by lemma with a pres(ent) or past target in English (n=78 per condition)

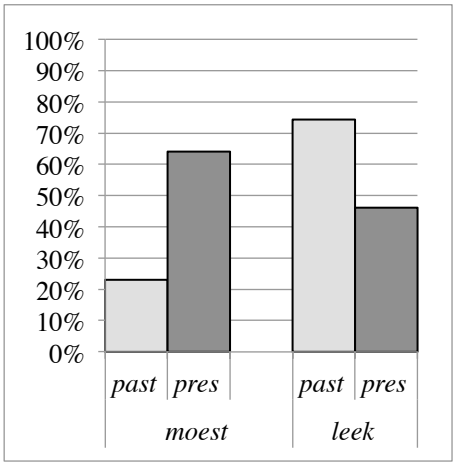

Figure 2: Proportions of 'true' answers by lemma with a pres(ent) or past target in Dutch (n=78 per condition)

Table 2: Results of the 2-tailed binomial tests versus chance results

Lemma \ Target	**Past** evaluation time	**Present** evaluation time
had to	*p<0.0001***	*p<0.0001****
seemed	*p=0.734*	*p<0.0001****
moest 'had to'	*p<0.0001****	*p=0.017**
leek 'seemed'	*p<0.0001****	*p=0.571*

3.3. Truth-Value Judgment Task – Discussion

Dutch results The results are consistent with the hypothesis that epistemic *moeten* 'must' scopes *over* tense in Dutch non-narrative contexts: In contexts with a past and present evaluation time, a sentence with *moest* is judged false when it targets a past evaluation time, while the same sentence with *leek* 'seemed' is judged true. *Moest* moreover tends to be judged as true when it targets a present evaluation time. These results on *moest* are unexpected if the Dutch modal scopes *under* tense.

If Dutch epistemic modals scope *over* tense, why are there still 'true' responses for epistemic *moest* when it targets a past evaluation time, and 'false' responses when it targets a present evaluation time? There are two things to note here. First, there was variation both between[9] and within participants, which I will probe further in future work. Second, some contexts may not have been completely clear about what the evidence showed at the two different time points. One context in particular led to both the highest percentage of 'true' responses with *moest* targeting a past evaluation time (55.6%; average per context 25%), and the highest percentage of 'false' responses targeting a present evaluation time (77.8%; average per context 30.6%). The Dutch version of this context lacked an indicator marking the

[9] One participant consistently gave judgments in line with the interpretation of *moest* as scoping under tense.

incompatibility between the past and the current belief that was present in the English version, which may have lead to the different outcome.[10]
An unexpected result concerning the control *leek* 'seemed' is that participants are at chance with a present target. Why are sentences like (24) not always judged as false?

(24) *Ad was tot op dit moment de hoofdverdachte voor een misdrijf dat gisteravond plaatsvond in Utrecht. Tot nu toe wees al het bewijs erop dat hij gisteravond in Utrecht was. Maar nu ontvangt de detective een nieuw bewijsstuk dat bewijst dat Ad eigenlijk in Breda was gisteravond.* De detective zegt: Ad leek in Breda te zijn.
Ad has been a prime suspect for a crime that occurred last night in Utrecht. Up to now, all of the evidence pointed to him being in Utrecht last night. But just now, the detective receives fresh evidence that proves that Ad was in fact in Breda last night. *The detective says: Ad seemed to be in Breda.*

Note that this result does not change the results for *moest*, as there is still a contrast between *leek* and *moest* in both conditions. What could be the case in (24) is that some participants were focusing on where Ad actually was last night, and since the information provided by the embedded clause answers this question, they may have ignored the past tense morphology on *leek* 'seemed'. In the condition with *leek* and a past target (*Ad seemed to be in Utrecht*), the information provided by the embedded clause does not tell us where Ad was last night, which might have resulted in a more careful parse.

English results Had to behaves as expected under the hypothesis that epistemic modals scope over tense: It is judged 'true' with a present target and 'false' with a past target. What the experiment fails to show, however, is an informative contrast with the control, *seemed*: With a present target, participants judge sentences with *seemed* as true too, most likely for the same reason as Dutch participants. With a past target, as in (25), English participants are at chance.

(25) *(English version of (24), using Montreal as a past target and DC as a present target): The detective says: Al seemed to be* ***in Montreal****.*

Note first that this result does not falsify the hypothesis that *had to* scopes over tense; the results on *had to* are consistent with the claim that *had to* scopes over tense, and inconsistent with the claim that *had to* scopes under tense. What the result fails to show is that in identical contexts, *had to* cannot target a past evaluation time while *seemed* can. There might therefore be a problem with the experiment itself.

What is more, the English results contrast with the Dutch results, as Dutch participants judge sentences with *leek* 'seemed' targeting a past evaluation time as true. The issue thus seems to concern English *seemed* specifically. One possibility is that the sentences with *seemed* do not

[10] Another potential explanation could be that the modal tested in this context does not have an epistemic interpretation for some of the participants (cf. the variation in judgments mentioned in section 2; Boogaart (2002) & Rullmann & Matthewson (2018:326) claim that past tense Dutch epistemic modals cannot be used to express an epistemic claim with a present evaluation time, while Barbiers (1995:202) claims they can). A follow-up study will therefore be an acceptability judgment study to see whether there is variation between datapoints and where this variation comes from.

focus on the epistemic claim as much as *leek*; participants might therefore respond purely to the information provided by the embedded clause and ignore *seemed*. In a follow-up experiment described in section 3.4., the focus was shifted to the epistemic claim by setting up a direct contrast between a past and a present epistemic claim in the test sentence, similar to (26). The prediction is that in this set-up, when *had to* is used to express a past epistemic claim, it is judged as infelicitous while the same sentence with *seemed* are not.[11]

(26) John {**seemed, had to**} to be home, but now I'm not so sure anymore.

3.4. Acceptability Judgment Task – Design and methods

The goal of the follow-up study is to find out whether part of the results from the TVJT were due to a problem with the experimental set-up. Participants did not judge that the English control item *seemed* could be used to target a past evaluation time, in contrast to previous claims (8). The explanation investigated by means of the Acceptability Judgment Task is whether the context did not highlight the past target enough. In a context in which the focus is on the past target, is it possible for *seemed* to target a past tense evaluation? If so, we can compare the results with *had to* to find out whether in the same set-up this epistemic modal can scope under tense. This follow-up was only conducted in English since the TVJT already showed that in the same set-up, Dutch *moest* 'had to' cannot target a past evaluation time while *leek* 'seemed' can.

For the Acceptability Judgment Task (AJT), each trial consisted of a context and a test sentence, as illustrated in (27). Participants were asked for their judgment on the sentence in italics, for which they were given a binary choice: Does the sentence sound *fine* or *(a little) strange* within the given context? Each context was followed by one of the two types of sentences in (28), and the judgment of the test sentence in (28a) determines whether English *had to* can scope *under* tense. Since the sentence contains an explicit contrast between two epistemic statements (*the earth is/isn't stationary*), the epistemic statement expressed by the modal in the first part of the sentence is forced to be evaluated at a time before speech time. If *had to* can scope ***under*** tense, sentences containing *had to* should therefore be judged as *fine*. If *had to* cannot scope under tense, sentences containing *had to* should be judged as *(a little) strange*. The predictions following from the two claims are in (28) in grey.

(27) A professor of ancient Greek culture discusses the ideas of some early philosophers. He says: The ancient Greeks worried much about astronomy, but they had some beliefs that have since been shown to be false. For instance,
*While the earth **had to** be stationary, it actually isn't.*

(28) a. *While the earth **had to** be stationary, it actually isn't.* (epi>T: not fine, T>epi: fine)
b. *While the earth **seemed** to be stationary, it actually isn't.* (fine)

The comparison is with the same sentences with *seemed* (28b), which are expected to be judged as *fine*. The factor tested is summarized below.

(29) Lemma: *had to* vs. *seemed*

[11] Another possible follow-up experiment would be to choose an epistemic verb that inherently focuses more on the epistemic claim, i.e. *appear*.

What the TVJT and the acceptability task have in common is that they do not have a context in the past tense (section 2). In contrast to the TVJT, however, the test sentences in the acceptability task contain an explicit contrast between a past and a present epistemic claim, which should help the participants focus on the available interpretations for *seemed*.

Materials Six contexts similar to (27) were created. For each context, there were two possible test sentences, depending on the lemma tested (*had to* vs. *seemed*). The epistemic claim expressed in the first part of the sentence was always negated in the second part of the same sentence (*it actually isn't*).

Each test sentence was followed by the question "Does the sentence in italics sound fine to you within the given context?", which could be answered by making a binary choice: "Yes, this sentence sounds fine", which is interpreted as being felicitous, or "No, this sentence sounds (a little) strange", which is interpreted as being infelicitous.

The experiment contained either two *seemed* trials or two *had to* trials. The two conditions were tested between participants in order to prevent participants directly contrasting *had to* and *seemed* themselves: While it could be that there is a dispreference for epistemic uses of *had to*, this experiment is conducted to find out whether the interpretation under discussion is *available*. There were six versions of the experiment, each containing two out of six contexts. The contexts were in a fixed order within each version of the experiment.

Two training items and four fillers were created in which outdated and updated beliefs were contrasted by means of epistemic attitude verbs, adverbs and adjectives. All test items, training items and fillers can be found at http://www.annemarievandooren.com/papers-2/.

Participants 34 native naïve speakers of English, currently based in the United States, participated in this study. While this number of participants is small, it makes it possible for every unique sentence to be seen by at least five individuals. This makes it comparable to Mahowald et al's (2016) proposal for small-scale acceptability studies. Given that there were six unique sentences for each of the two conditions, and every participant saw two test sentences, at least 30 participants are necessary. Participants were recruited through Amazon Mechanical Turk and paid for their participation.

Procedure Before starting the experiment, there were two training items with feedback. Participants then judged six items in total, two test items plus four fillers. Since all participants were given two test sentences, the 34 participants gave 68 judgments in total.

Data analysis The responses of the participants were analyzed using a two-tailed binomial test (R Studio Core Team, 2008) to investigate whether the proportions of answers with 'sounding fine' or 'sounding (a little strange)' were higher or lower than expected by chance.

3.5. Acceptability Judgment Task – Results

The average accuracy on fillers was high (mean accuracy: 90.4%). Six participants performed under 75% on fillers and were excluded.

28 participants are included in the results: 13 participants in the *had to* condition, and 15 participants in the *seemed* condition. The sentences in the *had to* condition are judged as 'sounding fine' 38.5% of the time, while sentences in the *seemed* condition in the same set-up are judged as 'sounding fine' 96.7% of the time (Figure 3). A binomial test (two-sided) indicates that the proportion of 'sounding fine' responses for the *seemed* condition is higher than expected by chance ($p<0.0001$***).

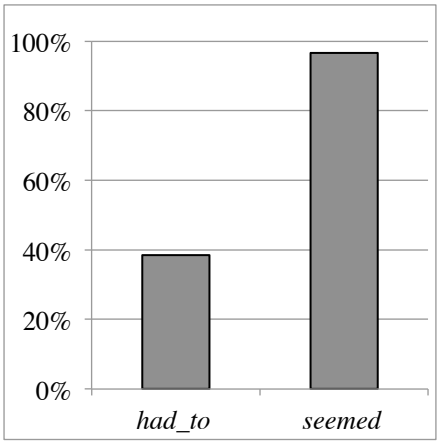

Figure 3: Proportion of 'fine' answers with a past target for had to *(n=26) and* seemed *(n=30)*

3.6. Acceptability Judgment Task – Discussion

Seemed can be used to express a past epistemic claim: In a set-up in which a past epistemic claim is contrasted with a present epistemic claim, native naïve speakers of English judge sentences in which *seemed* is used to express the past epistemic claim as 'sounding fine'. The contrasting results for *seemed* in the Acceptability Judgment Task as compared to the TVJT support the explanation suggested at the end of section 3.3: The set-up of the TVJT in combination with this particular epistemic verb might not have made the participants focus on the past epistemic claim, which led to a rejection of sentences with *seemed* targeting a past evaluation time. In the current set-up, the contrast is explicit in the test sentence, which likely led to the different outcome.

The results support the claim that the epistemic modal *have to* cannot scope *under* the tense marker in its own clause. While sentences in which *seemed* is used to express a past epistemic claim are judged as 'sounding fine', identical sentences with *had to* are not. This result on *had to* is unexpected if the English modal scopes *under* tense.

Why did some participants judge the trials with *had to* as fine? I.e., why is the acceptability for *had to* with a past evaluation time more than 38%, while it is much lower in the TVJT? This question will be left for future research but one interesting thing to note is that participants were quite consistent in their judgments: Each participant was given two trials of the same condition, and 8/10 'fine' judgments for *had to* trials were given by the same

participants. That means that out of the 13 participants in the *had to* condition, four participants judged both sentences as 'sounding fine', seven participants judged both sentences as 'sounding (a little) strange', and only two participants had mixed judgments.

3.7. General discussion

The two experiments together provide support for the claim that the Dutch epistemic modal *moeten* 'have to' and the English modal *have to* do not scope under tense in non-narrative contexts. While in identical set-ups, the past tense form of the epistemic verbs *lijken* 'seem' and *seem* can be used to make a past epistemic claim in matrix contexts, the same is not true for the two epistemic modals. This result is unexpected if the modals scope *under* tense. The results from experiment 1 furthermore provide support for *moeten* and *have to* being able to scope *over* tense: in matrix contexts, past tense *moest* and *had to* can be used to express a present epistemic claim about a past event.

Besides these general tendencies, there is variation within and between speakers of the same language, and across the two languages, which will need to be explored further. For instance, is there systematic variation between datapoints, and if so, where does it come from (fn. 10)? Finally, the results raise a question concerning the results reported in Rullmann & Matthewson (2018): Are the results they report indeed caused by a dispreference for this use of *had to*, as compared to *must have been* or *has to have been*? Why would there be such a dispreference?

4. Conclusion

Can epistemic modals in Dutch and English be interpreted within the scope of the tense in their own clause? The results of two online experiments support the claim that for Dutch *moeten* and English *have to*, the answer is 'no'. In the contexts tested in the two experiments, sentences with the past tense form of the Dutch epistemic verb *lijken* 'seem' are judged true (TVJT) and with the past tense form of English *seem* as felicitous (Acceptability Task) when they are used to express a past epistemic claim. In contrast, sentences with the Dutch past tense epistemic *moest* and the English past tense epistemic *had to* in exactly the same set-up are not. These results are inconsistent with the claim that *moeten* and *have to* scope under tense.

The results instead support the claim that *moeten* and *have to* are able to scope over tense (34), as they can be interpreted at the local time of evaluation, which in the matrix contexts tested here is speech time. Indeed, sentences with past tense *moest* and *had to* are judged as true by native speakers when they are used to express a present epistemic claim about a past event (TVJT). The tense *on* the modal is in these sentences interpreted *under* the modal, shifting the event under the modal to the past.

(34) ModEpistemic > Tense > ModRoot > Verb

In sum, the results seem to support the claim that epistemic modals scope *over* tense (Groenendijk & Stokhof 1975, Iatridou 1990, Stowell 2004, Condoravdi 2002, Hacquard 2006, 2010, a.o.), and challenge the claim that English and Dutch modals can or must scope *under* tense (von Fintel & Gillies 2007, Rullmann & Matthewson 2018).

What makes epistemic modals so special? After all, both epistemic verbs (*seem/lijken*) and root modals (deontics) scope *under* tense: it is the combination of being a modal and being used to express epistemic modality that results in the opposite scope relation. Cinque (1999) proposes that there is a universal hierarchy of functional projections in which epistemics scope over tense. Therefore, functional items used to express epistemic modality are in a structural position above tense. Hacquard (2006, 2010) claims that it follows from the specific syntax and semantics of modals: Modals that are located above tense get an epistemic interpretation *because* they are interpreted relative to a high speech event, while modals that are located below tense are not interpreted relative to the speech event and instead can only get a non-epistemic, root interpretation.

A final empirical contribution of this paper is the development of a new experimental design that can be used to test whether epistemics scope over or under tense in a language. The Acceptability Judgment Task described in section 3.4. directs participants' focus on the epistemic claims themselves, as there is a direct contrast between a past and a present epistemic claim in the same sentence, which seems to help in getting the intended interpretation.

References

Barbiers, S. (1995). The syntax of interpretation. Holland Institute of Generative Linguistics, the Netherlands.

Boogaart, R. (2002). The past and perfect of epistemic modals. In L. de Saussure, J. Moeschler, and G. Puskás (Eds.), *Recent Advances in the Syntax and Semantics of Tense, Aspect and Modality,* pp. 47-69. Berlin: Mouton the Gruyter.

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective.* Oxford: Oxford University Press.

Condoravdi, C. (2002). Temporal interpretations of modals. Modals for the present and for the past. In D. Beaver et al (Eds.), *Stanford Papers on Semantics*, pp. 59-87. Stanford: CSLI Publications.

von Fintel, K. and A. Gillies (2007). An opinionated guide to epistemic modality. In T. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology: Volume 2*, pp. 32-62. Oxford: Oxford University Press.

Groenendijk, J. and M. Stokhof (1975). Modality and conversational information. *Theoretical Linguistics 2,* pp. 61–112.

Hacquard, V. (2006). Aspects of Modality. Ph.D. Thesis, MIT.

Hacquard, V. (2010). On the event relativity of modal auxiliaries. *Natural Language Semantics 18(1),* pp. 79-114.

Hacquard, V. (2016). Modals: Meaning categories? In J. Blaszczak, et al. (Eds.), *Mood, aspect, modality revisited: New answers to old questions,* pp. 45-74. Chicago: University of Chicago Press.

Hacquard, V. and A. Cournane (2016). Themes and variations in the expression of modality. In C. Hammerly & Brandon Prickett (Eds.), *Proceedings of the 46th Annual Meeting of the North East Linguistic Society, Vol. 2,* pp. 21-42. Amherst MA: GSLA.

Iatridou, S. (1990). The past, the possible, and the evident. *Linguistic Inquiry 21(1),* pp. 123-129.

Mahowald, K., P. Graff, J. Hartmann & E. Gibson (2016). SNAP judgments: A small N acceptability paradigm (SNAP) for linguistic acceptability judgments. *Language*, *Volume 92(3), pp. 619-635.*

R Development Core Team (2008). A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Rullmann, H. and L. Matthewson (2018). Towards a theory of modal-temporal interaction. *Language 94(2),* pp. 281-331.

Stowell, T. (2004). Tense and modals. In J. Guéron and J. Lecarme (Eds.), *The Syntax of Time*, pp. 621-636. Cambridge: MIT Press.