Evidential restrictions on epistemic modals
Lisa Matthewson
University of British Columbia

1 Introduction

Epistemic modals and evidentials have conceptually distinct definitions: evidentials encode information about the speaker’s source of evidence for the proposition advanced, while epistemic modals introduce quantification over epistemically accessible possible worlds. However, these functions are often both performed by the same morphemes, and for several years researchers have been debating the empirical relation between the two classes. Almost all possibilities have been argued for in this debate, from complete disjointness (e.g., de Haan 1999, Aikhenvald 2004) to complete identity (e.g., Matthewson 2011a, in press).

In recent work on the relation between epistemic modals and evidentials, von Fintel and Gillies (2010, henceforth vF&G) argue that the English epistemic modals must and might encode an evidential restriction, specifically a requirement of indirect evidence. They further suggest that the indirect evidence restriction is not specific to these English modals, but is ‘persistent and cross-linguistically stable’ (vF&G:368), holding ‘in language after language’ (vF&G:367). In this paper I will support, extend and tweak vF&G’s proposals. I will argue that must and might make exactly the evidential contributions vF&G propose, but contrary to what vF&G claim, we should not characterize that contribution as a requirement for indirect inference. I will show that despite first appearances to the contrary, must and might do pattern like other elements cross-linguistically whose primary function is evidential. And following ideas found in Kratzer (2012), I will argue that all epistemic modals encode evidential information, as a matter of definition, since an ‘epistemic modal’ is a modal whose modal base relies on evidence (not on knowledge). Once the connection between epistemic modality and evidentiality is tightened in this way, the fact that elements like must encode restrictions on evidence is no longer a mystery, as vF&G reluctantly suggest (vF&G:368).

The paper is structured as follows. In section 2 I show that that vF&G’s claims about the evidential contributions of must and might are supported by a wider range of data than they originally gave. I then argue that it is not correct to characterize the evidential contributions of the English modals as requiring ‘indirect inference’, at least not in the way this term is typically

---

1 Many thanks to St’át’imcets consultants Gertrude Ned, Laura Thevarge, the late Beverley Frank and the late Rose Whitley. Many thanks also to Luis Alonso-Ovalle, Henry Davis, Kai von Fintel, Paula Menéndez-Benito, Hotze Rullmann, Ori Simchen, the audience at the Workshop on Epistemic Indefinites at the University of Göttingen, the participants in the Modality Seminar at the University of British Columbia, and two anonymous reviewers. This work was supported by SSHRC grants #410-2007-1046 and #410-2011-0431.

used in the evidentials literature. Rather than requiring the absence of a particular type of evidence (say, perceptual evidence), *must* and *might* rule out evidence of a certain level of trustworthiness or reliability. Although evidentials are not usually viewed as encoding trustworthiness distinctions, I show that there are some elements whose primary function is evidential, which do target the trustworthiness dimension. This brings *must* and *might* nicely into line with the broader class of evidentials, exactly as vF&G wanted.

In section 3, I address vF&G’s worry about having to lexically hardwire an ‘indirectness’ (i.e., untrustworthiness) requirement on all epistemic modals. Their intuition is that if all epistemic modals require such a restriction, it should somehow fall out for free, perhaps as a conversational implicature. However, they show that it is difficult to make an implicature story work. I argue that while all epistemic modals encode some restriction on evidence type – as expected, given that their function is to encode information about evidence – it does not have to be an indirectness / untrustworthiness requirement. Since some epistemic modals allow direct evidence and some rule it out, we have no choice but to stipulate indirectness when the facts support it.

Finally, in section 4 I relate my findings to the issue of epistemic indefinites, the topic of the current volume. There are some intriguing indications that epistemic indefinites pay attention to evidential-like notions (Alonso-Ovalle and Menéndez-Benito 2003, Aloni and Port 2010). However, there are also principled differences between epistemic indefinites and evidentials, and there is a clear need for further research on the extent of the connection.

The reader may notice a disconnect between the empirical focus of this paper (two English lexical items, plus brief looks at Quechua and St’át’imcets) and the strength of the conclusion (all epistemic modals in natural language encode evidential information). Obviously, the proposals made here need to be tested in many other languages. The idea that epistemic modals and evidentials are very closely connected rests not just on the material presented here, but on a range of other arguments from a number of languages; see the references in footnote 2 for discussion and debate.

2 Evidential restrictions on *must* and *might*

One of the most worked-out analyses of an epistemic modal as encoding evidential semantics is vF&G’s analysis of *must* (and secondarily *might*). I begin with an overview of their proposal.

2.1 *Must* as an evidential: von Fintel and Gillies (2010)

vF&G argue that all epistemic modals require that the speaker’s evidence for the prejacent proposition is indirect. Their discussion focuses primarily on English *must*. *Must* is predicted to be infelicitous whenever the speaker has direct evidence (as in (1)), and felicitous when the evidence is only indirect (as in (2-3)).

(1) Seeing the pouring rain.

---

2 Westmoreland (1998) proposes that *must* signals that the information was obtained via ‘deduction’.
??It must be raining. (vF&G:353)

(2) Seeing wet raingear and knowing rain is the only cause.
    It must be raining. (vF&G:353)

(3) Chris has lost her ball, but she knows with full certainty that it is in either Box A or B or C.
    She says:
    The ball is in A or B or C. It is not in A ... It is not in B. So, it must be in C. (vF&G:362)

vF&G argue that the indirect evidential signal is what is responsible for the apparent weakness of a must-assertion vis-à-vis its plain counterpart. They thus argue, contrary to the analysis which has been fairly standard since Kratzer (1981), that must $\varphi$ is not semantically weaker than $\varphi$, but on the contrary entails it. Their claims about the non-weakness of must are largely independent of the evidential question: must could have its standard weak semantics and still contribute an indirect evidence requirement. I will therefore set the weakness issue aside here, concentrating instead on what exactly the evidential restriction of must is.

According to vF&G (2010:368), must presupposes that neither the prejacent nor its negation ‘is known through direct evidence or trustworthy reports.’ In terms of formal implementation, vF&G assume that the modal base (as usual) represents the information compatible with what is known in the world at the context. They then propose a special set of propositions representing ‘the privileged information,’ i.e. the ‘direct information’ that the speaker has in the context. The set of propositions representing the direct information is called the kernel (K), and it determines its own special modal base $B_K$, the set of worlds given by intersecting all the propositions in K (vF&G:371). All the propositions in K are true, and the modal bases determined by kernels are therefore reflexive; vF&G claim that one ‘can’t have direct information that P unless it is the case that P. So for a modal uttered at $w$, with respect to a kernel K, we know that $w \in \bigcap K$.4

The lexical entry for must is given in (4). Must $\varphi$ presupposes that the kernel does not directly settle $\varphi$, and asserts that the kernel entails $\varphi$. The definition of ‘directly settle’ is as in (5). A kernel K directly settles whether a proposition P just in case P is either entailed or contradicted by one of the propositions in K.5

(4) Fix a c-relevant kernel K:
   i. $[[\text{must } \varphi]]^{c,w}$ is defined only if K does not directly settle $[[\varphi]]^c$
   ii. $[[\text{must } \varphi]]^{c,w} = 1$ iff $B_K \subseteq [[\varphi]]^c$ (vF&G:372)

(5) K directly settles whether P iff either $X \subseteq P$ or $X \cap P = \emptyset$ for some $X \in K$. (vF&G:374)

---

4 This, combined with the fact that vF&G place certain propositions in the kernel which under standard analyses would appear in the ordering source (and therefore would not have to be true in the actual world), derives the proposed entailment from must $\varphi$ to $\varphi$. But as I said above, the strength issue is not my focus here.

5 vF&G actually offer two distinct implementations of their idea, and (5) is where they start to diverge. The difference between the two implementations is not relevant here.
Any felicitous and true use of *must* is a case where the kernel fails to directly settle whether \( \phi \) (i.e., there is no single proposition in K which entails or contradicts \( \phi \)), but the kernel entails \( \phi \) (i.e., \( \phi \) follows from the modal base determined by K). Now since any proposition which is not itself in K, but is entailed by it, is a proposition we have indirect evidence for, vF&G derive the result they want, namely that *must* \( \phi \) entails \( \phi \), but encodes that the speaker has only indirect evidence for \( \phi \).

Let’s see how it works with a simple example.

(6)  *Billy is seeing the pouring rain.*
    
    K: {it’s raining, ...}
    
    # It must be raining.

Here, Billy has direct evidence that it’s raining, so the proposition that it’s raining is in the kernel. The kernel directly settles the issue of whether it’s raining, so (6) violates the presupposition of *must*. What about the slightly different situation in (7)? Here, ‘Billy’s direct information is that the people coming inside have wet umbrellas, slickers, and galoshes and that rain is the only cause’ (vF&G:372).

(7)  *Billy sees wet raingear and knows rain is the only cause.*
    
    K: {there is wet raingear, if the raingear is wet then it’s raining, ...}
    
    It must be raining.

The kernel in (7) doesn’t directly settle whether it’s raining, as there is no single proposition in K which either entails or contradicts the proposition that it’s raining. So (7) is felicitous. K entails that it’s raining, so (7) is true.\(^6\)

The analysis works well for the simple cases. However, vF&G do not go into more detail than this about the nature of *must*’s evidential restriction. We do not yet know exactly what kinds of evidence count as ‘direct’ enough to go into the kernel, thereby affecting the felicity of utterances containing *must*.

2.2 The evidence restriction of *must*: a bit more detail

According to vF&G, *must* rules out ‘direct evidence’; they sometimes augment this to ‘direct observation or trustworthy reports.’ However, their data do not extend beyond the types of example already given, namely the contrast between a speaker who sees the rain, and a speaker who sees wet raingear. In this section I present a range of data which establish more fully the correctness of vF&G’s generalizations, and which enable us to characterize precisely the evidential contribution made by *must*.

It will be useful to bear in mind the categories of evidence which are typically assumed in the evidentials literature. (8) shows Willett’s (1988) influential cross-linguistic taxonomy of

\(^6\) Notice that part of the ‘direct information’ in (7) is the general-knowledge proposition that if the raingear is wet then it’s raining. This will be relevant below.
evidence types encoded in evidential systems.

(8) Types of Evidence (Willett 1988:57)

<table>
<thead>
<tr>
<th>Direct</th>
<th>Indirect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attested</td>
<td>Reported</td>
</tr>
<tr>
<td>Visual</td>
<td>Second-hand Reasoning</td>
</tr>
<tr>
<td>Auditory</td>
<td>Third-hand Results</td>
</tr>
<tr>
<td>Other sensory</td>
<td>Folklore</td>
</tr>
</tbody>
</table>

According to Willett’s categorization, direct evidence is evidence personally witnessed by the speaker; languages may make finer distinctions within that, with individual evidentials requiring for example specifically visual witness. Within the category of indirect evidence, there are two main types: evidence reported by third persons, and evidence resulting from inference based either on observable results of an event, or on reasoning. vF&G propose that epistemic modals encode indirect inference, corresponding to the rightmost branch in (8).

The first observation about must is that – as vF&G predict – all types of sensory evidence send propositions to the kernel. That is, no distinctions are made between evidence obtained via sight as opposed to the other senses. This is shown in (9-13). In each case, sensory observation causes the relevant proposition to appear in the kernel, which means that the kernel directly settles the relevant question, so must is infelicitous.

(9) *The speaker sees the rain.*
    K: {it is raining, ...}                     VISUAL WITNESS
    # It must be raining.

(10) *The speaker hears people playing Tchaikovsky.*
        K: {they are playing Tchaikovsky, ...}   AUDITORY WITNESS
            # They must be playing Tchaikovsky.

(11) *The speaker smells something good.*
        K: {something smells good, ...}    OLFACTORY WITNESS
            # Something must smell good.

(12) *The speaker tastes something good.*
        K: {something tastes good, ...}    GUSTATORY WITNESS
            # Something must taste good.

(13) *The speaker feels that his/her coat is wet.*
        K: {my coat is wet, ...}       TACTILE WITNESS
            # My coat must be wet.

So far, so simple. Now we turn to cases where the speaker has indirect sensory evidence, i.e.,
evidence which is obtained via the senses, but which neither entails nor contradicts the prejacent proposition. We see in (14-16) that must is felicitous in such cases. Here the context has also been set up so that the entire kernel (including additional propositions known by the speaker) entails φ. The analysis correctly predicts that the must-statements in (14-16) are true in the contexts given.

(14) *The speaker sees people behind a window wincing and holding their ears while listening to music.*
K: {they’re wincing and holding their ears, they’re listening to music, people wince and hold their ears when things are too loud, …}
The music must be too loud.

(15) *The speaker hears someone crying inside the next room.*
K: {someone is crying in the next room, I just saw Susie go into that room, I know the room was empty before, …} Susie must be crying.

(16) *The speaker smells a smell like burning meat.*
K: {I smell something like burning meat, I was cooking meat earlier, I haven’t turned the oven off, …} I must have burnt the meat.

What about reports from third persons? vF&G suggest that trustworthy reports that φ send φ to the kernel. The claim is supported by (17-19). In (17), a trustworthy report that φ renders must φ infelicitous. Unless Belinda’s trustworthiness sent the proposition that Bob is home to the kernel, we would incorrectly predict (17) to be acceptable.⁷

(17) *Belinda tells the speaker that Bob is home.*
TRUSTWORTHY REPORT
K: {Belinda said Bob is home, Belinda is a reliable source about Bob’s whereabouts, Bob is home, …} # Bob must be home.

If the direct evidence contains only trustworthy reports, then untrustworthy reports that φ should not send φ to the kernel. This predicts that in (18), the indirectness presupposition of must is satisfied, since K does not directly settle whether Bob is home. However, the sentence is false in this context (and therefore infelicitous for a different reason, so still marked with a #): given that Belinda’s report is untrustworthy, K does not entail that Bob is home.

---

⁷ vF&G do discuss cases involving reports. They observe, for example, that after reading a newspaper report which tells us that the Knicks lost last night’s game, one cannot say ‘The Knicks must have lost again.’ This is similar to (17). However, vF&G do not contrast this with cases of untrustworthy reports. Consequently, they over-generalize the effect to all reportatives, suggesting (2010:354) that epistemic modals will in all cases treat reports as direct evidence. However, in most evidential systems, reportatives pattern with indirect evidentials. See section 2.3 for further discussion.
The claim that only trustworthy reports count as direct evidence is not yet proven, since both (17) and (18) are predicted to be unacceptable, albeit for different reasons (presupposition failure vs. falsity respectively). Judgments distinguishing presupposition failure from falsity are notoriously unreliable; see von Fintel (2004). However, it is possible to show that the analysis is on the right track. Consider a case where we have a trustworthy report about something other than the prejacent. In (19), there is no report which directly settles the proposition that Bob is home, so \textit{Bob must be home} is felicitous. However we do have a trustworthy report that Bob’s lights are on, and the presence of this report in $K$ ensures that $K$ entails the prejacent. It is crucial for deriving the truth of (19) that the trustworthy report is in the kernel.

(19) \textit{Belinda tells the speaker that Bob’s lights are on.} \\
$K$: \{Belinda said Bob’s lights are on, Belinda is a reliable source about Bob’s lights, Bob’s lights are on, Bob’s lights are only on when he’s home, ...\} \\
Bob must be home.

As above (see footnote 6), it is important that general knowledge propositions appear in the kernel, not just propositions for which the speaker has direct evidence in the utterance situation. Unless the proposition that Bob’s lights are only on when he’s home is in the kernel, $K$ will not entail that Bob is home, and (19) will be incorrectly predicted to be false.

Summarizing so far, vF&G’s analysis predicts the correct range of data for \textit{must}, under the assumption that what counts as direct evidence includes sensory perception, trustworthy reports, and general knowledge. vF&G extend their indirect evidence analysis to the possibility modal \textit{might}. For reasons of space, I am unable to give any \textit{might} data, but the reader can verify that \textit{might} patterns in the same way as \textit{must} for all the data given above. (To see this, as vF&G point out, one must use negated sentences, since otherwise the indirectness signal is swamped by the weak modal force of \textit{might}. For example, we predict that one cannot say \textit{It can’t be raining} when looking at sunshine, etc.)

In the next sub-section I discuss the relation between the evidence restriction encoded by the English epistemic modals and those encoded by elements traditionally analysed as evidentials. For convenience I continue to talk only about \textit{must}, and assume that the results extend to \textit{might}.

2.3 \textbf{The evidence restriction on \textit{must}: relation to evidentials}

We have seen that the kernel for \textit{must} contains the following three types of information:

(20) i. information obtained by sensory observation in the utterance situation  
ii. trustworthy reports  
iii. general knowledge

The inclusion of general knowledge – which we saw above was crucial for the analysis to work –
does not immediately seem to fit with vF&G’s characterization of the kernel propositions as ‘the pieces of direct information explicitly given by the context’ (2010:374). Kai von Fintel (p.c.) suggests that ‘Maybe, the story should be that while often we arrive at such conditionals via an indirect inference (induction/abduction), they over time become part of the trusted kernel.’ This points to the unifying property behind the three apparently disparate sources of evidence listed in (20): trustworthiness.8

Trustworthiness is not a property which is typically mentioned as relevant by the evidentials literature. To see the puzzle posed by must, consider again Willett’s taxonomy of evidence types in (8). We see that the types of evidence which are, or can be, classified as ‘direct’ by must include not only those which Willett counts as direct, but most of what he counts as indirect, including some reported information, and even inferred information from results (e.g., (14-16)) or reasoning (e.g., (3)).9 It is thus not correct to claim, as vF&G do, that epistemic modals like must require evidence from the right-most branch of Willett’s taxonomy, indirect inference.

It is not just Willett’s divisions which show a mismatch with those of must. Aikhenvald’s (2004:43) definition of a ‘direct’ evidential is that it involves ‘sensory experience’. This rules out many general knowledge propositions, and is clearly a subset of what is counted as ‘direct’ by must. The distinction made by must between trustworthy and untrustworthy reports also runs counter to the empirical situation in languages which have reportative evidentials. Usually, reportative evidentials are used whenever the information was obtained by a report, regardless of whether the report is judged to be trustworthy or untrustworthy. And all reported information is almost always classified as indirect evidence by evidential systems.

However, must is not unique in its singling out of trustworthy evidence. I argue in Matthewson (2011b) that trustworthiness is actually one of three dimensions of meaning which evidentials encode:

(21) Dimensions of meaning encoded in evidential restrictions:
1. Evidence type: whether the evidence is visual, sensory, reported, etc.
2. Evidence location: whether the speaker witnessed the event itself or merely some of its results
3. Evidence strength: the trustworthiness/reliability of the evidence

In Matthewson (2011b) I discuss evidentials from a range of languages, and argue that several of them encode (sometimes along with other information) whether or not the speaker considers their evidence to be trustworthy. One good example is the Cuzco Quechua direct/best possible grounds evidential =mi (Faller 2002, 2011).10 Faller argues (2002:18) that =mi requires that ‘the

8 It also points to a significant role played by the speaker, which is side-stepped in the denotation for must given in (4) above. The speaker judges what counts as trustworthy evidence in the context. I set this aside here, but see for example Stephenson (2007) for relevant discussion.
9 Due to space limitations I cannot discuss thirdhand reports or folklore, but here too, it appears that trustworthiness is the relevant criterion for determining whether must is felicitous.
10 Also its Wanka Quechua counterpart. Aikhenvald (2004:161) writes that in Wanka Quechua, ‘The direct evidential expresses the speaker’s firm belief that what they are talking about is true.’
speaker has the best possible source of information required for the type of event described.’ The restriction imposed by \( =mi \) thus appears to be at least partly a restriction on the strength of the evidence the speaker has, rather on the type. This is supported by the fact that \( =mi \) is licensed by any of the following three types of evidence (Faller 2010):

(22) i. Direct evidence, in cases where the described event is directly observable or otherwise directly accessible
   ii. The ‘next best thing’, in cases where the event is not observable (including reliable reports)
   iii. Undisputed common and learnt knowledge

Cases of ‘the next best thing’ and common or learnt knowledge are illustrated in (23) and (24) respectively. (‘Ev’ indicates the source of the speaker’s information. \( =mi \) has an allomorph \( =n \).)

(23) paqarin Inés Quṣuq-ta=\( n \) ri-nqa
tomorrow Inés Cuzco-ACC=\( BPG \) go-3FUT

‘Inés will go to Cuzco tomorrow.’

Ev: Inés told the speaker that she will go to Cuzco tomorrow (Faller 2011)

(24) a. 1945 wata-\( n \) segunda guerra mundial=qa tuku-rqa-n
1945 year-LOC=\( BPG \) second war world=TOP end-PST-3

‘World War II ended in 1945.’

Ev: learnt in school (Faller 2011; see also Aikhenvald 2004:162)

b. yunka-\( n \) k’usillu-kuna-qa ka-n
rainforest-LOC=\( BPG \) monkey-PL-TOP be-3

‘In the rainforest, there are monkeys.’

Ev: speaker knows it as part of Quechua culture (Faller 2002:133)

The similarity between the evidential requirements of must and those of \( =mi \) is striking. In both cases, the relevant kinds of evidence – those which are either disallowed (for must) or allowed (for \( =mi \)) – are the most trustworthy kinds of evidence.

It is important that the trustworthiness distinction is still an evidential notion, and does not reduce to speaker certainty about the prejacent proposition. This is true for must under vF&G’s analysis, since for them, the speaker of must \( \phi \) can easily be certain that \( \phi \) is true, but must lack a single trustworthy-evidence proposition that \( \phi \). For Quechua, Faller observes that pure certainty does not license \( =mi \). For example, suppose we are having a meeting and Mary’s partner calls to say that she is sick and cannot attend. Suppose further that I completely trust Mary’s partner. I still could not use \( =mi \) to report to the others that Mary is sick; I would have to use the reportative evidential instead (Martina Faller, p.c.). The difference between this case and the above is that in the latter case, the speaker reports on the state of another person, while in the former case, the speaker reports on the speaker’s own state.

She gives examples of the use of this evidential where the speaker did not directly witness the events described, but has a reliable secondhand source. See also Izvorski (1997), Lecarme (2008) and Waldie (in prep.) on the idea that evidentials can encode the reliability or strength of the evidence.
cases in (23-24), where =mi is licensed, is that it is in principle possible to directly witness Mary’s sickness. Therefore, direct witness counts as the ‘best possible’ evidence for this event, and =mi is disallowed if the speaker lacks that best evidence. In (23), in contrast, Inés’s future plans are not directly observable; therefore the best possible evidence about her plans is her own report. This shows that =mi depends on the very highest level of evidence-trustworthiness, and that what counts as the highest level of trustworthiness is dependent on the type of event being reported.

Summarizing so far, we have seen that kernel-membership for must is defined in terms of trustworthiness of the evidence, and that there is at least one other element which has been analyzed as an evidential which also encodes trustworthiness. The comparison with Quechua =mi provides a cross-linguistic argument that the restriction on must really is an evidential distinction. However, the argument would not be very strong if the comparison were only between must and =mi. The claim that both must and =mi rely on trustworthiness (rather than on more traditional evidential notions such as those discussed by Willett) could just as easily show that =mi is not an evidential.

For reasons of space I cannot address this question fully here. However in Matthewson (2011b), I discuss a range of other elements which are clearly evidentials (as they encode traditional evidential notions such as sensory witness), but which also care about the trustworthiness of the evidence. For example, the evidential determiners in Nivacle (Matacoan-Mataguayan) encode whether or not the speaker has had, at some point in the relevant individual’s lifespan, the ‘best possible’ type of sensory evidence for the existence of that individual (Gutiérrez and Matthewson 2012). It thus seems that cross-linguistically, trustworthiness of the evidence is something which evidentials can choose to pay attention to. This parallel between the semantics of epistemic modals and evidentials supports the argument for the evidential nature of epistemic modals.

### 2.4 Comparison with a reportative evidential

vF&G note (2010:381) that an obvious extension of their work ‘would explore whether the tool of structured modal bases can help to give a principled account of the variety of evidential systems found cross-linguistically.’ In this section I show that vF&G’s kernel-based analysis can be extended for use with at least one evidential which is more ‘traditional’ in that its primary function is to encode evidence type: the St’át’imcets (Lillooet Salish) reportative ku7. A simple example of the use of this reportative is given in (25).

(25) wá7=ku7 ku=sts’ëts’qwaz’ l=ta=stswáw’cw=a be=REPORT DET=trout in=DET=creek=EXIS

‘[reportedly] There are trout in the creek.’ (Matthewson et al. 2007)

Matthewson et al. (2007) show that ku7 passes a range of tests for being an epistemic modal, and patterns in all relevant respects like other modals in the language. Their analysis of ku7 is given in (26).

(26) [[ ku7 ]]cw is only defined if c provides a modal base B such that for all worlds w’, w’ ∈
B(w) iff the reported evidence in w holds in w'.
If defined, $[[ku7]]^w = \lambda f_{<st,st>}. \lambda p_{<s,t>}. \forall w'[w' \in f(B(w)) \rightarrow p(w')]$

The choice function f in (26) is a means of allowing weaker-than-universal force for the modal; it narrows the domain of worlds quantified over. According to this analysis, $ku7$ presupposes that the modal base is given by reported evidence, and asserts that in all worlds in the subset of modal base worlds selected by the choice function f, the prejacent is true.

I will now give a kernel-based analysis of $ku7$. Recall that in vF&G’s analysis, the kernel contains the trustworthy information (the speaker’s direct sensory evidence, trustworthy reports, or general knowledge). The definedness condition on reportative $ku7$ can be modeled as a restriction on the kernel; it must contain the information that there was a report that $\phi$, $\phi$ itself need not be in the kernel, but it can be. That is, unlike $must$ (but like $=mi$), $ku7$ does not disallow that the kernel contains a single proposition which directly settles the prejacent. These ideas are captured in (27).

$$(27) \quad [[ku7 \phi]]^w \text{ is defined only if K contains a proposition of the form 'someone said $\phi$.'}$$
$$\quad \text{If defined, } [[ku7 \phi]]^w = 1 \text{ iff } f(K) \subseteq [[\phi]]^c$$

The choice function f selects a non-empty subset of the kernel worlds (which are presupposed to be worlds in which someone said $\phi$). A sentence $ku7 \phi$ is true iff in all worlds in the subset of kernel worlds selected by the choice function f, $\phi$ is true. This captures the same facts as the earlier, non-kernel-based analysis in (26), but is a bit more precise about the nature of the report which has to exist.11

As stated above, $ku7$ does not require that the report itself be trustworthy. That is, while the fact that there was a report must appear in the kernel, the content of the report may or may not also appear in the kernel. This is illustrated in (28), where K1 and K2 are two different kernels corresponding to two different discourse contexts. With K1, the speaker considers Belinda a reliable source, and with K2, she doesn’t. With K1, the speaker believes the prejacent proposition, and with K2, she doesn’t necessarily believe it.12

$$(28) \quad Belinda \ tells \ the \ speaker \ that \ Bob \ is \ home.$$
\begin{align*}
\text{wá7=ku7} & \quad \text{=ta=tsítcw-s=a} & \quad \text{kw=s=Bob} \\
\text{be=REPORT} & \quad \text{in=DET=house=3POSS=EXIS} & \quad \text{DET=NMLZ=Bob} \\
\text{‘Bob is [reportedly] home.’} & \\
\text{K1: } \{\text{Belinda said Bob is home, Belinda is a reliable source, Bob is home, ...} \}
\end{align*}$$

11 (27) (like (26)) raises the issue of the limits on possible evidential restrictions in natural language. For example, an anonymous reviewer asks whether an evidential could require K to contain a proposition of the form ‘someone said my uncle said $\phi$,’ or ‘someone believed $\phi$ for a while.’ This important and well-known problem goes beyond the bounds of the current paper; see Speas (2004), McCready (2010) for discussion.

12 As shown by Matthewson et al. (2007) and as predicted by (27), even with K2 the speaker must believe that it is at least possible that Bob is home. This is predicted because a non-empty subset of the kernel-worlds (those selected by the choice function) must be $\phi$-worlds.
K2: {Belinda said Bob is home, Belinda is not a reliable source, ...}

Ku7 differs from must in the ability of the prejacent to appear in the kernel; recall that must disallows trustworthy reports from directly settling $\phi$. My analysis of ku7 correctly predicts that unlike must, ku7 will be good even when the reports are so trustworthy and/or frequent that the information has become common knowledge. This is shown in (29-30).

(29) *It’s the day after the election in the U.S. You tell your husband:*

\[
\begin{align*}
t’\text{cum} &= \text{ku7} & k &= \text{Obama} \\
\text{win} &= \text{REPORT} & \text{DET} &= \text{Obama} \\
& \quad \text{‘Obama [reportedly] won.’} \\
& \quad \# \text{ Obama must have won.}
\end{align*}
\]

(30) *You’re talking to your grandchild about geography.>*

\[
\begin{align*}
k\text{éla7} &= \text{ku7} & \text{xelh} &= \text{káku7} & \text{South Pole} &= \text{a} \\
\text{first} &= \text{REPORT} & \text{DEIC} &= \text{South Pole} &= \text{EXIS} \\
& \quad \text{‘It’s [reportedly] cold at the South Pole.’} \\
& \quad \# \text{ It must be cold at the South Pole.}
\end{align*}
\]

However, ku7 may not be used when the speaker has direct sensory evidence for $\phi$. This is shown in (31).

(31) *You were invited to Ted’s wedding and you went there and watched him get married. Marilyn (Ted’s sister) didn’t see you at the wedding and didn’t know you had been invited. She told you “Ted got married.” Later, you see me and you tell me:*

\[
\begin{align*}
\# \text{ melyih} &= \text{ku7} & \text{kw} &= \text{s} &= \text{Ted} \\
\text{marry} &= \text{REPORT} & \text{DEIT} &= \text{NMLZ} &= \text{Ted} \\
& \quad \text{‘[reportedly] Ted got married.’} \\
& \quad \text{ (Matthewson et al. 2007)}
\end{align*}
\]

Negative restrictions of this type are not usually written in lexical entries in the evidentials literature; it is assumed that they follow as conversational implicatures (e.g., Faller 2002). If we did want to hardwire the restriction, we could give the kernel internal structure, distinguishing the speaker’s entire information state (K) from a subset of propositions representing those obtained via direct sensory evidence (K$_\text{SENS}$). Ku7 could then require that K$_\text{SENS}$ does not directly settle $\phi$.

We have seen that vF&G’s kernel-based analysis can easily be applied to elements, like the St’át’í’mcets reportative, whose primary function is to indicate evidence source, and which therefore fall under a traditional definition of evidentials. Is there any advantage to the kernel-based analysis in (27) over the original one in (26)? More broadly, does vF&G’s approach provide any insight into what types of evidentials we will find cross-linguistically? The answer to this is not clear to me at this stage. The kernel analysis is most obviously advantageous for elements which rely on trustworthiness (the dimension of ‘evidence strength’ in (21)), and less obviously relevant for elements relying on the traditional notion of evidence source. However,

\[\text{ Thanks to an anonymous reviewer for raising this issue.}\]
the tool of structured modal bases might be useful for elements which encode complex evidential restrictions, such as the St’át’imcets reportative if the restriction against direct sensory evidence is hardwired, as outlined immediately above. In Matthewson (2010), I provide an analysis of another St’át’imcets evidential, lákw7a, utilizing internal structure within the kernel. The restrictions on lákw7a are presented in (32):\textsuperscript{14}

\begin{equation}
[[ \text{lákw7a } \varphi ]]_{cw} \text{ is defined only if:}
\begin{align*}
\text{i. } & K_{SENS} \text{ contains information which bears on } \varphi \\
\text{ii. } & K_{VIS} \text{ [the set of propositions obtained through vision] does not directly settle } \varphi 
\end{align*}
\end{equation}

Analyses of this type have the advantage that they force us to confront the question of which types of sub-sets of K are targetted by natural language. On the other hand, very similar results are achieved by e.g., Portner’s (2006) theory of conversational update, according to which propositions presented by interlocutors are placed into different sets, including the common ground, the propositions relying on reports, the propositions relying on direct sensory evidence, etc. Further research is required on the relative merits of the various formal analyses.

3 All epistemic modals are evidentials, but they are not all indirect

So far I have argued that epistemic modals carry evidential information, and that the evidential contributions of modals parallel those of some elements whose primary function is evidential. In this section I address vF&G’s worry about whether the evidential restrictions of epistemic modals should follow from conversational implicature, rather than being lexically stipulated.

vF&G note that they ‘have not found a language whose expression of epistemic necessity fails to carry an evidential signal of indirect inference’ (2010:367), and they regret that they ‘have to leave as unsolved the mystery of why this seems to be happening with every epistemic necessity modal that we have come across’ (2010:368). They later extend the generalization – and hence the mystery – to include possibility modals (2010:369). Here I will argue that there are some epistemic modals which do not carry an ‘indirectness’ restriction in vF&G’s sense (which, following the discussion above, I will call an ‘untrustworthiness’ restriction). I will then argue that while an untrustworthiness restriction is not universally present on epistemic modals, some specification about evidence is universally present. This will however no longer be a stipulation, but will follow from the fact that an epistemic modal is simply a modal which cares about evidence (as argued by Kratzer 2012).

Recall what is meant by the untrustworthiness requirement on English epistemic modals: the kernel may not directly settle the prejacent. We have actually already seen an example of an epistemic modal which lacks this requirement: the St’át’imcets reportative ku7. I showed in (28-30) that ku7 is felicitous even when the prejacent proposition appears in the kernel.

The reader might however not be fully convinced by the ku7 example, given that I admitted that ku7 is infelicitous if the speaker directly witnessed the relevant event (see (31)). If we refine the definition of ‘(un)trustworthiness’ so that it relies not on whether the kernel directly settles φ, but

\textsuperscript{14} See Matthewson (2011a) for a non-kernel-based analysis of lákw7a.
on whether the set of propositions for which the speaker has direct sensory evidence (KSENS) directly settles \( \varphi \), then \( ku7 \) still counts as requiring untrustworthiness. Note, however, that this would be a weaker untrustworthiness requirement than that of \textit{must}, as \textit{must} does not allow general knowledge propositions to directly settle \( \varphi \), while \( ku7 \) does (see the contrast between \textit{must} and \( ku7 \) in (29-30)).

However, there is an even more striking potential case of an epistemic modal which lacks an untrustworthiness requirement: Cuzco Quechua \( =mi \).

Recall that \( =mi \) has the opposite requirement about evidence from English \textit{must}: \( =mi \) requires that the speaker have the most trustworthy kind of evidence for the proposition being advanced (including direct sensory evidence). It is therefore very relevant that Faller (2011) advances a modal analysis of \( =mi \).

Faller proposes that \( =mi \) is a necessity modal, and requires a non-empty modal base that contains propositions describing the speaker’s perceptions. \( =mi \) makes no reference to an ordering source, and a proposition of the form \( =mi \ \varphi \) asserts that \( \varphi \) is true (under the assumption that the speaker’s perceptions have not deceived them).

In short, there seems no reason to conclude that all epistemic modals encode an ‘untrustworthy evidence’ requirement. This in turn means that we do not need to worry about stipulating the untrustworthiness requirement of \textit{must}. \textit{Must}’s evidential contribution turns out to be just one of the multitude of different evidential contributions encoded by evidentials cross-linguistically.

What is universal to all epistemic modals, I suggest, is that they all make \textit{some} evidential contribution. In that case, the term ‘epistemic’ should be actually be retired, as argued by Kratzer (2010), who writes that ‘the term “epistemic modal” is a misnomer,’ because knowledge is not involved. Kratzer argues that knowledge is not required for an epistemic modal partly on the basis of cases like those in (33). These can be felicitously asserted even if the speaker knows that the newspaper reports or the testimony are false, and therefore knows that the prejacent propositions are false.

(33) a. According to the newspaper reports, the thief \textbf{must} have entered through the kitchen window.
   b. According to the testimony we heard, the defendant \textbf{must} be suffering from severe emotional distress.

If epistemic modals do not rely on knowledge, suppose instead that their modal bases depend on

---

\[ ^{15} \text{Interestingly, I am not aware of a reportative which hardwires indirectness in vF&G’s sense (requiring untrustworthy reports). Thanks to Henry Davis (p.c.) for pointing this out.} \]

\[ ^{16} \text{Lecarme (2008) also proposes a modal analysis of direct evidentials in Somali.} \]

\[ ^{17} \text{An anonymous reviewer asks what explains the pervasiveness of the indirect type of epistemic modal. This may have a functional explanation, since a modal which lacks indirect evidential semantics is very similar in truth conditions to a plain assertion. Note, however, that we do not have a large sample of languages on which to base the claim that indirect epistemic modals are more numerous. Even for the languages we know about, it is possible that elements with a similar semantics to \( =mi \) have mistakenly been analyzed as not part of the modal system.} \]

\[ ^{18} \text{Not all epistemic modals allow the speaker to believe that the prejacent is false; Kratzer observes that replacing \textit{according to} with \textit{given} in (33) removes this possibility.} \]
evidence. An epistemic necessity modal $M_{\phi}$ applied to $\phi$ then does not mean ‘all (stereotypical) worlds compatible with the speaker’s knowledge are $\phi$-worlds’, but ‘all (stereotypical) worlds compatible with the speaker’s evidence regarding $\phi$ are $\phi$-worlds.’ If this idea is right, then the reason epistemic modals have appeared to be about knowledge is perhaps simply that it is very normal to infer something about the speaker’s knowledge from their assertions about their evidence. And if the idea is right, then it stands to reason that it will be extremely common – if not universal – for ‘epistemic’ modals to encode restrictions on the type of evidence the speaker has.\(^1\)

4 Epistemic indefinites and evidential restrictions

So far I have argued that epistemic modals in at least English, and possibly all languages, always encode evidential information. In this final section I address the potential implications of my proposals for epistemic indefinites, the topic of the current volume.

Epistemic indefinites, as the reader is no doubt well aware, are indefinites which convey the speaker’s lack of knowledge about who or what might satisfy the existential claim. A Spanish example is given in (34), where algún signals that the speaker doesn’t know who María married.

(34) María se casó con algún estudiante del departamento de lingüística
‘María married a linguistics student.’ (Alonso-Ovalle and Menéndez-Benito 2010:2)

The modal inference for algún, according to Alonso-Ovalle and Menéndez-Benito (henceforth AO\&MB), can be represented as in (35). There must be at least two different individuals who are possibilities for the speaker.

\begin{equation}
\exists w',w'' \in D_{w}\{x : P(w')(x) & Q(w')(x)\} \neq \{x : P(w'')(x) & Q(w'')(x)\}
\end{equation}

(35) (where $D_w$ is the set of worlds compatible with what the speaker believes in $w$, and $P$ and $Q$ are two properties) (AO\&MB 2010:7)

The ignorance effect of elements like algún is clearly epistemic in nature: the speaker’s epistemic state is limited in some respect. However, it is not clear whether my proposals about

\(^{1}\) I should clarify the lack of contradiction between my proposal that epistemic modals depend on evidence rather than knowledge, and McCready’s (2008, 2010) claim that ‘evidence is knowledge.’ McCready’s proposal is that speakers use evidentials only when they know – rather than merely believe – that they have evidence of the relevant type. When their confidence in the evidence is reduced from knowledge to belief via skeptical arguments (e.g., if they are reminded that their senses might be deceiving them), they will decline to use an evidential. This proposal is compatible with my claim that the modal base is not determined by the speaker’s knowledge. In (33a), for example, the speaker has evidence (the newspaper report), which she knows – rather than merely believes – to exist (she read the newspaper, and has not been subjected to skeptical arguments). She asserts that in all worlds compatible with the evidence in the newspaper report, the thief entered through the kitchen window. However, she is still free to disbelieve the content of the report.
the evidential contributions of epistemic modals should transfer directly to epistemic indefinites. The reason is that epistemic indefinites, according to current analyses, do not themselves introduce epistemic modal semantics, or even directly restrict a higher epistemic modal. For AO&MB, the semantic contribution of algún is an anti-singleton requirement on the domain of quantification, and the modal variation inference in (35) arises as a conversational implicature. For Aloni and Port (2010), the contribution of German irgendein and Italian un qualche is to introduce a shift in the method of identification of the individual (modeled as a shift in conceptual covers; see their paper for details). For Aloni and Port, the modal variation effect derives not from Gricean implicatures, but from the conceptual cover shift plus a felicity condition stating that the domain shift has to happen for a reason. Given that the epistemic effects are derived indirectly in both these analyses, it is not immediately obvious how to transfer the proposal that all epistemic modals place a restriction on evidence.

In spite of this, we can see some tentative connections between evidential epistemic modals and epistemic indefinites. For example, AO&MB (2003) discuss a difference between English some and Spanish algún which appears to be evidential in nature. They point out that some can be used when making a claim about an individual for whom the speaker has direct perceptual evidence, but algún cannot. Thus, (36) is infelicitous in the context described, while its English translation with some is fine. The minimally different context in (37), where there is no direct perception of the professor, is fine with algún.

(36) L and P are talking in the lounge of the Math department, on one of those long, boring, predictable winter evenings when everybody is doing what they are supposed to. Neither L nor P knows anybody there. All of a sudden, a burst of Brazilian music starts to play in an office. Believe it or not, there is a guy dancing lambada on his desk. Unbeknownst to L and P, the guy is Rino Cusper, the famous statistician.

¡Mira! algún profesor está bailando la lambada encima de la mesa.

‘Look! Some professor is dancing lambada on his table.’ (AO&MB 2003:4)

(37) Strange things happen! L and P are outside the office where a faculty meeting is being held. Suddenly, they hear lambada music coming out of the office and sounds that unequivocally indicate that a person is dancing to the music.

Strange things happen! L and P are outside the office where a faculty meeting is being held. Suddenly, they hear lambada music coming out of the office and sounds that unequivocally indicate that a person is dancing to the music. (AO&MB 2003:5)

The question is whether the ban on direct perception with algún is something that needs to be stated separately (as it could be on an evidential/epistemic modal), or whether the contrast between (36) and (37) already falls out from the analysis of algún as requiring there to be at least two individuals who the speaker considers possibilities. AO&MB point out that (37) is infelicitous because it is not the case that any professor could be dancing lambada; the speaker knows that it is that single professor she is looking at who is dancing.

We shouldn’t conclude too hastily, though, that nothing evidential is going on in (36-37). AO&MB and Aloni and Port (2010) use such data to illustrate the issue of methods of identification, something which may be related to evidentiality. The idea is that it is not sufficient to talk about whether an individual counts as a possibility for the speaker; we need to consider the way(s) in which the speaker can identify an individual. In (36), for example, P can
identify the professor by ostension, but she can’t identify him by name. Aloni and Port argue that epistemic indefinites require that the method of identification be different from the one required for knowledge in the relevant context. They also propose that epistemic indefinites across languages can encode restrictions on methods of identification. Thus, German irgendein is felicitous if the speaker can identify the individual by ostension but not by name, but Italian un qualche (like Spanish algún) is bad in similar cases.

The distinction between different methods of identification does not have an obvious analogue in the propositional epistemic modal domain, or in the evidentials literature. While ‘ostension’ could be assimilated to visual evidentiality, description and naming can arise through various evidence sources, including third-party reports and inference as well as visual evidence. It seems that there might be a principled difference between evidence about the truth of a proposition, which we dealt with in the first three sections of the paper, and evidence about the identity of an individual, which epistemic indefinites are typically concerned with.

Perhaps a more fruitful comparison for epistemic indefinites would be evidentials within the nominal domain. These have been discussed for Somali by LeCarme (2003), for various languages by Imai (2003), and for Nivacle by Gutierrez and Matthewson (2012). Gutierrez and Matthewson show that in Nivacle, the determiners encode whether or not the speaker has had, at some point in the relevant individual’s lifespan, the best type of sensory evidence (usually visual evidence) for the existence of that individual. Imai (2003) shows that in Malagasy, various evidential notions are encoded by demonstratives and pronouns, including for example a requirement for non-visual sensory evidence.20 Here again, however, there are differences with epistemic indefinites. To my knowledge, elements which have been identified as nominal evidentials do not encode information about methods of identification such as naming vs. description. Instead, they encode things like which sense the speaker used to perceive the individual (sight, hearing, etc.). Nor do elements which have been identified as nominal evidentials display any of the characteristic ‘free choice’ effects of epistemic indefinites like algún. It is clear that there is room for much future research on the connections between propositional epistemic modals, epistemic indefinites, and evidentials.

5 Conclusion

In this paper I have argued in support of vF&G’s empirical claim that English must and might require that the speaker lacks a certain kind of evidence for the prejacent proposition. Specifically, the set of propositions representing the speaker’s trustworthy evidence must not contain any single proposition which entails the prejacent. I have shown that this evidential

---

20 Within nominals, there is often a close connection between evidentiality and deixis. Many of the nominal evidential elements reported on by e.g., Lecarme (2003) and Imai (2003) encode information about the speaker’s (non-)perception of the individual at the utterance time, and hence also give deictic information. The Nivacle determiners are not indexical in this way. In Nivacle, an individual who is absent at the utterance time will still require a ‘best sensory evidence’ determiner if the speaker has visually witnessed the individual at any point in that individual’s lifespan. These are interesting differences to be investigated in future research on nominal evidentiality.
restriction can – in spite of initial appearances – be brought into line with the restrictions encoded by other evidential elements cross-linguistically, and I have broadened the claim to propose that all ‘epistemic’ modals, in all languages, are quantifiers over worlds compatible with bodies of evidence (rather than with knowledge states). As such, it is no longer a mystery that such elements encode restrictions on evidence. However, which precise restrictions on evidence each element encodes is not predictable, and must be lexically stipulated.

These proposals raise several questions. First, there is the empirical question of whether it is right to generalize the results from English, St’át’imcets and Quechua to other languages. This will obviously have to await further research, although it is significant that there is already a substantial body of research arguing for a close connection between epistemic modality and evidentiality (see footnote 2).

With regard to the broader relation between epistemic modals and evidentials, I have argued here that a correlation holds in one direction: epistemic modals all encode evidential information. I have not addressed whether the implication holds in the other direction – i.e., whether all evidentials are modals; see Faller (2002), Murray (2010), Peterson 2010, among others, for the claim that they are not. Even with a one-way correlation, the question arises of why out of a set of elements which can receive such similar analyses, some are typically called ‘epistemic modals’, and some are called ‘evidentials.’ A first answer to this might run as follows. Elements which are called epistemic modals are modals whose paradigms are impoverished with respect to the fine-grainedness of their evidential distinctions. English modals, for example, require a lack of trustworthiness, and that’s about it, and consequently evidentiality has been less salient in the study of English modals. Moreover, English modals do other interesting things, like lexically distinguishing modal force, and allowing non-epistemic interpretations. On the other hand, elements which are called evidentials are modals which make richer evidential distinctions, tend not to make distinctions in modal force, and lack non-epistemic interpretations. In spite of these differences, however, all these elements can receive essentially similar analyses, as modals which quantify over worlds compatible with certain kinds of evidence.

References

Tübingen.
Kratzer, Angelika 2010. Collected papers on modals and conditionals. Ms., University of Massachusetts, Amherst (published in revised form as Kratzer 2012).
Matthewson, Lisa 2010. Evidential restrictions on epistemic modals. Paper presented at the


Waldie, Ryan, Tyler Peterson, Hotze Rullmann, and Scott Mackie 2009. Evidentials as epistemic modals or speech act operators: Testing the tests. Paper presented at the Workshop on the Structure and Constituency of Languages of the Americas 14, Purdue University.
