Indexicality: *de se* anaphora in discourse

Lecture 6: Perspectival expressions

1. Tense and point of view

There’s an old debate in metaphysics about subjective vs. objective characterizations of time. In modern work, McTaggart (1908:458) started with the observation that:

Positions in time, as time appears to us *prima facie*, are distinguished in two ways. Each position is Earlier than some, and Later than some, of the other positions. And each position is either Past, Present, or Future. The distinctions of the former class are permanent, while those of the latter are not. If M is ever earlier than N, it is always earlier. But an event, which is now present, was future and will be past.

He took the Earlier/Later relations to be more objective but *not* more “fundamental to the nature of time” than Past/Present/Future; in fact

the distinction of past, present and future is as *essential* to time as the distinction of earlier and later, while in a certain sense, as we shall see, it may be regarded as more *fundamental* than the distinction of earlier and later.

Accordingly, times can be ordered in two different ways: the first ordering, the “A-series”, assumes that times are ordered relative to a changing standpoint, and is reflected in expressions like *today* and *tomorrow*. The second, or “B-series”, is just a strict linear order of moments, reflected in *before* and *after*.

The usual approach to modeling time in formal semantics is decontextualized, so that tenses pick out arbitrary times in a strict order as the “locations” of the truth of propositions; i.e., this is a B-series. Here’s a puzzle that arises out of the way of modeling time, pointed out by the tense logician Arthur Prior (1959:17):

One says, e.g. "Thank goodness that's over!", and not only is this, when said, quite clear without any date appended, but it says something which it is impossible that any use of a tenseless copula with a date should convey. It certainly doesn't mean the same as, e.g. "Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954", even if it be said then. (Nor, for that matter, does it mean "Thank goodness the conclusion of that thing is contemporaneous with this utterance". Why should anyone thank goodness for that?)

Thomason (2015) offers this nice discussion of Prior’s example:
Prior...considers the case of someone who has forgotten the date, but who, “in a dateless haze” says on June 15, 1954, “Thank goodness that’s over!” To add detail, suppose this person thinks, about yesterday’s visit to the dentist, that she’s glad that the visit is over. This is not at all the same as thinking you’re glad that the visit ended before June 15, 1954. And the confused person who believes on June 15, that today is June 14, does not believe that June 15 is June 14. But the standard theories of tense found in formal semantics predict this incorrect pattern of beliefs.

Thomason speculates about how we might deal with this puzzle:

In Montague-style semantics, a proposition is a set of worlds—if time is taken into consideration, this means that a proposition is a set of histories. Then the proposition expressed by Today is June 14 on June 15 would be the set of histories in which June 14 is June 15, i.e. the empty set. This is not a worthy object of belief. However, suppose we represent propositions in a more general way, as a set of centered times—these are simply worlds associated with a time location, of ordered pairs consisting of a world and a time. This is, of course, a generalization of the simpler idea that propositions are sets of worlds, because we can identify an observer-independent proposition—a set $W$ of worlds—with the set of all pairs $<w, t>$ such that $w \in W$.

Then Prior’s problem doesn’t arise. For instance, the person who thinks on June 15 that today is June 14 believes a centered proposition containing pairs $<w, t>$ such that $t$ occurs on June 14. This proposition is not empty. And similar problems have solutions that seem equally natural.

...In the literature on centered worlds, time is only one of the many aspects of self-location that are incorporated in these worlds. But temporal self-location is an important case, and illustrates the general ideas well.

So Thomason is suggesting that we need a perspectival account of tense, one in which tenses like the English present are anchored in the point of view of some salient ‘now’ in the world of interpretation. This temporal perspective then is modeled in terms of centered worlds where the center is a time. One might think of such a center as the origin in McTaggart’s A-series. This discussion suggests that the attitudes expressed by tensed clauses—even main clauses, as in Prior’s example—may be centered on such an origin.

There is further evidence that tense has some properties related to those we’ve observed for indexicals: while some are fixed relative to the utterance time (UT), others shift to be anchored to the event time (ET) of an embedding clause. English tense is UT-anchored (except for the “historical present”, a sort of FID use of the present)—i.e. always anchored to the time of utterance: the UT serves as the origin, and the past and future functions yield a perspective from that origin: back along the time line or forward into the (possibly branching, irrealis) future. But there are a number of languages like Korean where in embedded contexts the anchoring time is the event time of the embedding clause (see, e.g., Yoon 1996). So:
The English (140a) can only be true if what John said implied that Mary is pregnant at the utterance time; it may also have the so-called “double accessibility” reading (Abusch 1988, Enç 1987) wherein Mary was pregnant at the time of John’s saying as well as at the utterance time. Compare this to (141), which doesn’t commit John to Mary still being pregnant at UT:

(141) John said that Mary was pregnant.

But the Korean (140b) has no double accessibility reading. Moreover, while there is no past tense in the complement clause (which makes it ‘non-past’ according to Yoon), Mary’s pregnancy is interpreted as obtaining at the ET of the matrix, John’s saying time (in the past). If the past tense –ess is added to that complement clause, the resulting interpretation would be that John said Mary had been pregnant—i.e. it would yield a past-from-the-past interpretation.

Korean tense in embedded clauses—complements, adverbal clauses (both temporal and non-), and relative clauses—is almost always interpreted relative to the ET of the embedding clause, as we see in this paradigm from Yoon (1996; Chapter 5), with the past tense –ess highlighted:

   John-TOP Mary-NOM leave-PAST-COMP feel-PAST-DEC
   'John felt that Mary had left.'

   John-TOP Mary-NOM leave-PAST-COMP feel-NONPAST-DEC
   'John feels that Mary left/has left.'

(144) John-un nayil Mary-ka ttena-ess-tako nukki-keyss-ta. [psbl Fut COMP]
   John-TOP tomorrow Mary-NOM leave-PAST-COMP feel-MOD-DEC
   'John will feel tomorrow that Mary left.'

If we were to swap the non-past for –ess in (142) – (144), this would systematically alter their interpretations, so that in each case the leaving of the complement clause would temporally overlap the matrix ET at which John holds his attitude—in the past in (142), the present in (143), the future in (144).

Embedded under matrix past in (142), the embedded past has a perfective English translation (past relative to the matrix past ET); under non-past in (143), it yields a simple past (past relative to the matrix present ET); under the futurate modal –keyss in (144), it yields a past-in-the-future interpretation (past relative to some future event, so that the leaving itself is potentially future).

We see parallel behavior in temporal adverbial clauses, here with ttay, ‘time’ or ‘when’:

   Mary-NOM arrive-NONPAST-REL time John-NOM leave-PAST-DEC
   'John left when Mary was arriving.'
In (145) under matrix past -ess, Mary’s arrival is understood to overlap John’s past leaving; under futurate –keyss in (146), the arriving is still simultaneous, now with a future leaving.

The emerging pattern is that Korean tense is always interpreted relative to the ET of an embedding clause; only in matrix clauses does it take the UT as anchor. A *prima facie* exception occurs with relative clauses. Consider Yoon’s (147):

    John-NOM  Seoul-to  go-NONPAST-REL person-Ace  seek-PAST-DEC

a. 'John sought a person who was going to Seoul (at the time of his seeking).'</n
b. 'John sought a person who is going to Seoul [at UT].'

(147) has two meanings, as noted: one (a) in which the non-past relative clause is understood to be simultaneous with the past ET of the matrix seeking, but another (b) in which the non-past relative clause tense has a present interpretation, as if interpreted relative to UT.

But notice that this is not the case if we replace the intensional matrix predicate *chach-ess-ta* ‘seek’ in (147) with an extensional predicate like *po-ess-ta* ‘saw’ (Heedon Ahn, p.c.). The result would have only the relative-to-matrix (a) reading, where the seeing temporally overlaps with the going to Seoul. This is consistent with Yoon’s (1996) explanation of the (b) reading for (147): it is the consequence of a *de re* reading of the entire object *a person who was going to Seoul*, in which it is understood to take wide scope over the intensional *chach-ess-ta*, so that it behaves as if anchored to the UT. If it proves generally correct that we only get the (b) reading under intensional predicates, this explains the apparent exception, and we get a very simple resulting pattern of interpretation for Korean tenses: they are interpreted relative to the ET of an embedding clause whose scope they fall under at LF.

### 2. What’s a perspectival expression?

We just saw that a shifting tense like those in Korean is perspectival in the temporal sense: depending on a temporal anchor given by a matrix clause. Then we might conceive of natural language tense as an perspectival expression, with a profile that includes the components of indexical meaning that we mapped out in Lecture 4. We can think of using the anchoring ET to establish an origin on a B-series order over times: the ET serves as the time from which we can set up relations of past, present and future, effectively establishing an A-series. This is not a doxastic perspective, unlike that in *now*. Hence, we find that the tense shift does not depend on a shift in doxastic perspective: Korean tenses shift under any embedding ET, while shifting the

---

1 I am told (Kate Ritchie, p.c.) that some philosophers grant a special metaphysical status to the origin on the A-series. That is not my intention here. Rather, the “origin” here is merely a reference point locally established in the interpretation of the utterance, relative to which the ‘past’, ‘non-past’ and ‘future’ of Korean tense can be determined.
anchor of *now* requires either FID or an FID-like interpretation of the complement of an attitude, as we have seen. Similarly, in order to shift *come* some doxastic perspective other than that of the speaker or addressee must be salient (as in attitude complements or FID), while locative *to the left* is not so constrained.

Table 3: Some more perspectival expressions, compared

<table>
<thead>
<tr>
<th>Components: Indexicals:</th>
<th>Familiarity(^2)</th>
<th>Classificatory</th>
<th>Deictic</th>
<th>Relational</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>this,</em> / <em>that,</em> (locative, +/- demonstration)</td>
<td>(d_i)</td>
<td>(if any) Descriptive Content Cond’n on (d_i)</td>
<td>location in (w^<em>) of (&lt;a,t&gt; = \ominus^k) for (\ominus^k \in {\ominus^</em>, \ominus^\text{FID}})</td>
<td>(&lt;\text{loc}(&lt;a,t&gt;,w^<em>),v&gt;) (v) the demonstrated vector with origin in (a) at (t) in (w^</em>)</td>
</tr>
<tr>
<td>(i_i)</td>
<td>(d_i)</td>
<td>(\emptyset)</td>
<td>(&lt;a,t&gt; = \ominus^*)</td>
<td>(&lt;&lt;&lt;a,t&gt;,w^*&gt;,\text{Dox}&gt;)</td>
</tr>
<tr>
<td>Amharic 1(^{st}) person -ñññ</td>
<td>(d_i)</td>
<td>(\emptyset)</td>
<td>(&lt;a,t&gt; \in {\ominus^*, \ominus^\text{say}, \ominus^\text{tell}})</td>
<td>(&lt;&lt;&lt;b,t^*,w&gt;,\text{ATT}&gt;) ATT the CG or embedding attitude, (w \in \text{ATT}(&lt;&lt;a,t&gt;,w^{\text{interp}}&gt;))</td>
</tr>
<tr>
<td>English past tense</td>
<td>(d_i)</td>
<td>(d_i \subseteq RT)</td>
<td>(t = \text{UT})</td>
<td>(&lt;t,\text{Past}&gt;) (cf. A-series with (t) as origin)</td>
</tr>
<tr>
<td>Korean past tense -ess</td>
<td>(d_i)</td>
<td>(d_i \subseteq RT)</td>
<td>(t = \text{ET} ) of embedding clause, if any, else (\text{UT})</td>
<td>(&lt;t,\text{Past}&gt;) (cf. A-series with (t) as origin)</td>
</tr>
<tr>
<td><em>to the left</em></td>
<td>(d_i)</td>
<td>Descriptive Content of complement holds of (d)</td>
<td>(\text{loc}(a,\text{ET},w), \alpha) the familiar anchor</td>
<td>(&lt;v,\text{Origin}&gt;, v) a vector from (\text{Origin}) to (\text{loc}(d,\text{ET},w))</td>
</tr>
<tr>
<td><em>come</em></td>
<td>(d_i)</td>
<td>locative adverbials, if any</td>
<td>(&lt;a,t&gt; = \ominus^k) for (\ominus^k \in {\ominus^*, \ominus^\text{say}, \ominus^\text{FID}, \ominus^\text{att}})</td>
<td>(&lt;&lt;&lt;a,t&gt;,w^{\text{interp}},\text{Dox}&gt;, w^k \in \text{Dox}(&lt;&lt;a,t&gt;,w^{\text{interp}}&gt;))</td>
</tr>
</tbody>
</table>

\(^2\) For the indexicals, this is familiarity of the denotatum. For *come*, it is familiarity of the presupposed end-point of a path.
Here the English and Korean past tenses differ only in the anchoring origin, which shifts under embedding, to yield a different temporal perspective—as given by that embedding ET: the perspective is the set of times accessible via the Past function—which takes the B series order over times plus the origin—effectively an A series centered on that origin, and yields the set of times which precede the origin in the order.

I’ve also added the perspectival locative to the left. Such locative PPs pick out a location in space (the denotatum) relative to two other locations: a point of view and a reference point (the “landmark” or “ground” in traditional locative terminology). All three locations are weakly familiar to the interlocutors (at the time of utterance of the PP, if not prior to utterance), including the denotatum $d$. The classificatory component is the descriptive content of the complement of the expression—either explicit (to the left of the tree) or implicit/elliptical (the middle of the battlefield in (148)):

(148) John watched the parley group in the middle of the battlefield. To the left, the British were arrayed, to the right the French.

This descriptive content constrains the reference point $d$, serving as the index component: $d$ is the location of the entity satisfying that descriptive content. Then the location $d$, picked out by the whole PP is a sub-space to the left of $d$. However, since a battlefield—like a tree—has no inherent left side, the to-the-left relation must be determined relative to a point of view, that of the origin. The origin is the location in the 3D space (at the event time (ET) and in the world of evaluation of the clause in which the locative PP occurs) of a familiar anchor $a$. Again, the anchor may be the location of a person (e.g., John or the speaker) or an arbitrary location (from the top of Knob hill), either explicit (From John’s point of view) or implicit in context: In (148), John is the pragmatically given anchor, his location the origin. The elided index $d$ is the location of the group in the middle of the battlefield. The origin plus a vector extending from the origin (John’s location) to the index $d$ constitutes the POV of the origin, yielding John’s perspective, the visual field accessible from that POV. Then the Relational component tells us that the denotatum $d_i$ is a subspace of the left part of that visual $3^D$-field (a subspace of the space to the left of the vector and near the index from John’s point of view).

As we discussed in Lecture 5, shifting indexicals like Amharic –ññ differ from their English counterparts in the possibility of being anchored not only to the (purported doxastic) point of view of the actual speaker, but to that of the agent of embedding ‘say’ or ‘tell’. Hence, the index in a shifted use is not that of the speaker at the time of utterance, as we see with $I$, but that of the reported agent as they purport to be, yielding the de se.

The major difference between these perspectival expressions is in the character of the “space” in which the relevant perspective is mapped out: three-dimensional actual space for to the left, the two-dimensional temporal space for the tenses, and the space of possibility for the doxastic.

---

3 In work on the formal semantics of locatives, it has been argued that some locatives require another vector. E.g., behind requires us to establish a vector extending away from the reference location from the agent’s point of view. I ignore that here for simplicity. See Bohnemeyer (2012).
perspectivals, including the classical indexicals. These different “spaces” are reflected in the
different accessibility relations which serve as the second elements of the POV component of
these expressions’ meanings:

So now we have expanded our purview to consider the perspectival character not only of
indexicals and demonstratives, but of certain locatives, deictic motion verbs, and now tenses.
What do all the expressions in Tables 1 and 3 have in common, that would lead us to
characterize them all as perspectival?

Characteristics of perspectival expressions:
1. perspectival content is presuppositional, not part of proffered content, has no local effect,
   and tends to project
In particular, all perspectival expressions carry two kinds of presupposition:
2. display general properties of anaphoric expressions, including:
   • Partee’s (1984) paradigm, extended as in Lecture 1; note that Partee’s original
discussion was intended to argue the anaphoric character of tense
   • the usual scope constraints on antecedence (Kamp 1981, Heim 1982, Chierchia &
Rooth 1983, Roberts 1996)
   • the display of pseudo-scope behavior as defined above: via projection as a function of
the satisfaction of the anaphoric presupposition
3. presuppose deictic components of their meaning—an origin and POV, relative to which one
can identify an index—which are crucial to the determination of the expression’s proffered
content.
In addition:
4. insofar as different POVs may become relevant, we find variable anchoring—not only to
the default speaker/UT/utterance location, but to other entities of the appropriate type which
are sufficiently salient.
Doxastic perspectivals display additional features deriving from the space of doxastic
 possibilities over which their perspective is defined:
5. share a pattern of restricted distribution of possible perspectival anchors, with
   corresponding constraints on interpretation. Contexts which license shifting must make
salient not just a doxastic agent, but the doxastic perspective per se of such an agent,
thereby limiting the possible anchors to:
   • the speaker and addressee are always available as perspectival anchors
   • the speaker is the default perspectival anchor
   • the addressee is the pragmatic default in questions (“interrogative flip”)
   • the agent of an attitude predicate may serve as perspectival anchor
   • in Free Indirect Discourse (FID), the agent whose point of view is being represented
   • in modal subordination (Roberts 1989) involving epistemic modality, across both
   modal auxiliaries and attitude predicates (Heim 1992, Roberts 1996), an agent whose
views are serving as accommodated Modal Base may serve as perspectival anchor
   under the scope of the modal
6. yield a de se interpretation of any anaphoric element whose antecedent is the perspectival
anchor.
3. More types of perspectival expressions

Here we will extend the inventory of perspectival expressions just a bit more, to give a better sense of the wide range of relevant expressions. We’ll consider two more types in some detail: Potts’s conventional implicatures triggers, and epistemic modal auxiliaries. Others will have to be left for a future occasion.

3.1 Potts’ Conventional Implicature Triggers

Many natural language expressions have three distinct types of semantic content. One type of expression where we see all three is epithets, like the bastard in (149):

(149) [Clara:] I’m so mad at George. That bastard took my headset!

In uttering that bastard as an epithet, the speaker makes the following anaphoric assumption, clear to any competent addressee:

a) there is an entity who is both recently salient in the context of utterance and, of all such salient entities, most likely to be the speaker’s intended antecedent of that bastard

Then since there is only one way to resolve the anaphora in the context given in (149), to the salient George, the sentence in which the epithet occurs carries the following implications:

b) ‘Clara thinks George is a bad character’

c) ‘George took Clara’s headset’

Clearly, each of these has different status in the utterance. No one would maintain that (a) is part of what Clara meant when she said That bastard took my headset!. Instead it’s part of the machinery that she uses to convey what she means—important to doing so, but not the point. And though, unlike (a), Clara does clearly mean to convey (b) by using the epithet that bastard intending it to be understood as anaphoric to George, (b) doesn’t seem to be the main point of (149) either. For example, one cannot reply to Clara’s utterance with (150):

(150) That’s true! George is a bastard. However, he didn’t take your headset—I did.

The original concession of truth would be understood instead to pertain to (c), which is the main point of the utterance.

How do these implications arise? Notice that one would find comparable implications on any occasion of use of that bastard, suggesting that they arise from the epithet’s regular semantic content. Epithets are anaphoric; like pronouns, they occur in clauses which predicate something

---

4 One can of course respond to (1) with (i):

(i) It’s true that George is a bastard, but he didn’t take your headset.

But unlike that’s true in (150), (i) does not involve anaphora to the preceding utterance. On the classic analysis, this construction involves extraposition of the complement, with pleonastic it subject. And on any reasonable analysis, the complement is effectively the subject in this construction. Hence (i) would be an indirect response to (1), not a judgment of its truth, as is (150).
of the antecedent’s denotation; and in addition, they carry some additional implication about that individual’s character. If that is so, what does the differing status of these implications tell us about how they come about in the course of interpretation? I want to argue that the special status of implication (b) arises because the epithets like that bastard are perspectival.

Potts (2005) argued that epithets are members of the class of expressions he called Conventional Implicatures, including not only epithets but non-restrictive relative clauses (NRRCs), nominal appositives, speaker-oriented and topical sentential adverbials, and a variety of other types of expressions. According to him, they all display the following characteristics:

i. CIs contribute to the conventional meaning of the utterance in which they occur, and hence are non-cancellable;
ii. CIs are not at-issue in the utterance in which they occur; but
iii. CIs are not presupposed (and in fact are typically new information); nonetheless
iv. CIs always project: they never occur in the semantic scope of at-issue operators in whose syntactic scope they occur;
v. CIs do not interact semantically with the at-issue content of the utterance in which they occur; and
vi. CIs are speaker-oriented, except in direct quotations (when they are attributed to the reported speaker). Roughly, this is to say that they attribute to the speaker some attitude represented by the content of the CI.

Accordingly, Potts’ system generated CIs compositionally in a process parallel to but completely independent of the calculation of at-issue content.

In their review of Potts, Amaral Roberts & Smith (2007) agreed about properties (i) - (iv), but offered empirical evidence that he was wrong about (v) and (vi). Instead:

v’. CIs interact anaphorically with at-issue content in both directions, though it doesn’t seem possible to directly bind across the two kinds of content.

vi’. CIs are always understood to represent a particular point of view which is being considered by the interlocutors, which is by default that of the speaker, but may also be that of the addressee (especially in interrogative mood), the agent of an embedding attitude, or the agent whose POV is being reported in a literary text.

(v’) is illustrated in Amaral et al.’s (151) - (154):

Anaphora across types of CI and at-issue content:
(151) Stan Bronowski, who took an exam, passed it with flying colors.
(152) Several students, most of them linguists, missed the bus.
(153) Every professional man I polled said that while his wife, who had earned a bachelor’s degree, nevertheless had no work experience, he thought she could use it to get a good job if she needed one.
(154) In each class, several students failed the midterm exam, which they had to retake later.

(153) and (154) are especially striking because the anaphoric antecedent takes narrow scope under an at-issue operator in the same sentence: In (153), a bachelor’s degree (in the CI) is understood to be that of the wife of some arbitrary instantiation of every professional man, and in (154), several students (in the at-issue content) takes narrow scope relative to the quantificational
adverbial in each class. Notice that in none of these interactions is binding involved, so that the anaphora is arguably *discourse anaphora*, of the same general type as is observed in donkey anaphora.

With respect to (vi’), in (155) we see the usual speaker-orientation that arises when the utterance is considered out of the blue (when the speaker’s point of view is arguably the default adopted). But in (156) – (159) the point of view has shifted to some salient agent other than the speaker (or author):

**Anchoring of CIs to a shifted point of view:**

- **appositive:**
  (155) Sheila believes that Chuck, a psychopath, is a very nice guy and fit to watch the kids.
  (156) [speaker and addressee are forensic psychologists who agrees that Chuck is a psychopath:]
    Sheila believes that Chuck, a sweetheart if she ever met one, is fit to watch the kids.

- **NRRC:**
  (157) Joan is crazy. She’s hallucinating that some geniuses in Silicon Valley have invented a new brain chip that’s been installed in her left temporal lobe and permits her to speak any of a number of languages she’s never studied. She believes that her chip, which she had installed last month, has a twelve year guarantee.
    Again, saying, *That’s wrong* to the last claim seems to take issue with Joan’s belief that the chip has a twelve year guarantee, and not with her belief that she had it installed the previous month.

- **expressive:**
  (158) [Context: We know that Bob loves to do yard work and is very proud of his lawn, but also that he has a son Monty who hates to do yard chores. So Bob could say (perhaps in response to his partner’s suggestion that Monty be asked to mow the lawn while he is away on business):]
    Well, in fact Monty said to me this very morning that he hates to mow the *friggin’* lawn.

- **epithet:**
  (159) [In an historical novel set in 19th century England:] Johanna was dismayed by the fact that she suddenly found herself endebted to Lord Collins. Now what would she do when the rent came due tomorrow?! *That bastard* couldn’t care less if she was homeless! No, worse!—it suited his purposes perfectly.

Although (156) attributes to Sheila the belief that Chuck is a sweetheart, it doesn’t seem that this is the main point of the assertion. The latter is that she finds him fit to watch the kids. We can see this if we embed under a downward entailing attitude like *doubt*:

(160) Sheila doubts that Chuck, who’s a real sweetheart, is available to babysit tonight.

In the context in (156), we understand the NRRC *who’s a real sweetheart* in (160) to reflect Sheila’s judgment about Chuck, not the speaker’s. But that isn’t part of the content of Sheila’s reported doubt: She doubts his availability, not his sweetheartness. The sweetheart attribution plays a role in the embedded content something like an auxiliary explanation for Sheila’s belief.
In (156) and (157) the appositive anchor is pragmatically driven (on pain of contradiction with the Common Ground) to shift to the agent of the embedding attitude predicate. In (158), for the same reason, it’s understood to be anchored to the agent whose speech is being indirectly reported. And in the FID in (159), the descriptive content of the epithet is most naturally attributed to the character whose state is being described.

Harris & Potts (2009) conducted experimental studies that supported hypothesis (iv′): CIs are always anchored to some salient point of view, but not necessarily that of the speaker.

Note, as well, that anchoring to a salient point of view other than that of the speaker leaves open the possibility that the not-at-issue implications triggered by CIs do on some occasions take narrow scope with respect to an at-issue operator, specifically in case said operator itself takes wide scope with respect to the agent whose point of view is taken to anchor the CI:

(161)  [two psychologists talking about the diagnosed psychopath Chuck, whom they agree is a terrible person, especially a menace to children:] Every housewife in the neighborhood, totally charmed by our Chuck, believed that that sweetheart was fit to watch her kids.

Here, the point of view adopted for the interpretation of the epithet is that of the arbitrary housewife (and clearly not that of the speaker). Hence the implication is that x thinks Chuck is a sweetheart, with a free variable x which is bound by the at-issue subject every housewife in the neighborhood.

So in terms of our analysis, CIs presuppose are understood to be true from the salient doxastic perspective of some anchoring agent, with the usual limitations on how that perspective can be made salient. It is the stringent limitations that led Potts to original claim that CIs were always speaker-anchored.

The perspectival character of epithets like that bastard in (149) now help to explain properties iv and vi′ of CIs, repeated here:

iv. CIs always project: they never occur in the semantic scope of at-issue operators in whose syntactic scope they occur;

vi′. CIs are always understood to represent a particular point of view which is being considered by the interlocutors, which is by default that of the speaker, but may also be that of the addressee (especially in interrogative mood), the agent of an embedding attitude, or the agent whose POV is being reported in a literary text.

The projective properties of CIs arise from the fact that they are directly attributed to their doxastic anchor—typically the speaker, but sometimes some other agent whose doxastic perspective is sufficiently salient (vi′). The projection (iv) is then pseudo-scope, resulting from anaphoric anchoring to the doxastic anchor.
3.2 Epistemic modals

When we use modal auxiliaries like must and might with an epistemic sense, there is general agreement that the resulting interpretation is generally relative to some understood body of information available to an epistemic agent, some of which may be made explicit with an if-clause. Consider:

(162) It must be raining.
(163) If your umbrellas are wet, it must be raining.

Roughly, the speaker who asserts (162) seems to implicate that she has access to a body of evidence whose truth would entail that it’s raining. In (163), she relativizes the claim to the hypothetical assumption that the addressees’ umbrellas are wet; if everything else she knows (about the weather and umbrellas and their uses, etc.) is true and it’s true that the umbrellas are wet, then we can conclude that it’s raining.

Note that a body of information which some agent (purports to) believe (true or not) is a doxastic perspective in the sense we’ve defined in earlier lectures. Here I will (very briefly) offer two reasons to think that epistemic modal auxiliaries (EMAs) and the assertions we make with them are perspectival: they display evidence of the restricted availability of perspectival anchors which we saw in other doxastic perspectivals, and they display evidence of wide scope.

There has been a good deal of discussion and debate about whose epistemic state might be appealed to in the semantics of epistemic modals, e.g. in Hacking (1967), DeRose (1991), Egan et al. (2005), MacFarlane (2005), Stephenson (2007), and von Fintel & Gillies (2007a). As Yalcin (2007) says, “It is a striking fact that these questions do not have obvious answers.” Everyone notes that there is at least a default tendency to understand the speaker to be the relevant agent whose epistemic state is at-issue. DeRose (1991) proposes a speaker inclusion constraint to capture this in the strongest way, requiring the speaker to be at least part of the group with the relevant state:

**speaker inclusion constraint:** the relevant community [must] include the speaker. Hence “whenever S truly utters a might be F, S does not know that a is not F.” (p.5)

But as discussed by von Fintel & Gillies (2007a), this requirement is much too strong. Here is a sample of the types of anchors attested for epistemic modal auxiliaries:

- speaker at Utterance Time (©*):
  (164) John might be the thief.
  (165) This suggests that Angela must be in Austin right now.

Especially out of context, with no explicit suggestion to shift to a different point of view, this is the only reasonable way to anchor an EMA: the speaker’s point of view is always the default, for all perspectival expressions. And the following is also consistent with DeRose’s proposal:
• Group containing speaker at some actual time ≠ Utterance Time:

(166) Given what we knew at the time, John might have been the thief.  [von Fintel & Gillies]

Here the adverbial explicitly shifts the doxastic point of view to that of the speaker and others in the anaphorically retrieved denotation of we at some time prior to the utterance time.

But we also see cases where the grounds appeal to appear to be solely those of the addressee:

• Addresser:

(167) Where might you have put the keys?  [von Fintel & Gillies]

As with evidential particles in many languages, this amounts to Interrogative Flip (Faller 2002)—wherein the anchoring perspective naturally becomes that of the addressee when seeking information they might offer.

• Arbitrary group containing addressee at some hypothetical time:

(168) [Military trainer:] Before you walk into an area where there are lots of high trees, if there might be snipers hiding in the branches use your flamethrowers to clear away the foliage.  [von Fintel & Iatridou 2003]

In (168) note the imperative mood and the dependent plural your flamethrowers. This constitutes generic instructions issued to the addressees. Then the if-clause constitutes a precondition for carrying out the instructions—they are for circumstances in which the relevant information leaves open the possibility that there are snipers. But then, of course, the only reasonable anchor for might would be the trainee who has to decide whether the correct conditions obtain.

• Explicit third person orientation:

(169) As far as Bill knows, John might be the thief.  [von Fintel & Gillies]
(170) From John’s point of view, it must be raining.
(171) This suggested to George that the polar ice cap might be melting.
(172) John thinks it must be raining.

(169), (170) and (171) illustrate several ways to explicitly shift the intended point of view. Note that the center’s time may be the present (as is the only possibility with temporally present must in (170)), or may be some past time (as in (171), the past suggestion-time). With attitude predicates, like think in (172), shifting to the agent is the default (Stephenson 2007). Note, however, that this may not be a necessary shift. For example, in (173) and (174) the anchor probably includes the speaker (with or without the addressee), though perhaps a group that includes both the speaker and John, but the latter hasn’t yet drawn the relevant conclusion from the information to which he’s privy:

(173) John won’t acknowledge/hasn’t yet conceded/hasn’t realized that it must be raining.
(174) Has John realized that it must be raining?
Jefferson Barlew (p.c.) offers a case where the anchor would not include the agent of the attitude, here the speaker:

(175) Bill claims to have discovered evidence allowing him to deduce who killed Frank. If that’s true, I hope it must be George.

The only reasonable, RELEVANT interpretation to give must in (175) is as anchored to Bill and his new evidence. If that seems a bit forced, it just illustrates how strong the tendency is to anchor the EMA to the agent of the attitude hope, here the speaker. Dowell (2011) notes that it’s certainly possible to shift to an agent other than the agent of the embedding attitude when we use explicit shifting adverbials, as in this variation on one of her examples:

(176) Leiter believes that, for all Blofeld and No.2 know, Bond might be in Zurich.

and she gives other convincing cases without explicit adverbial shifting (see her discussion of extensions of the Mastermind example p.23ff).

- Third person in extended modal subordination: (Roberts 2015)

(177) Suppose you were John. Where would you go now to find Clarissa? You might find her with Sidney in New Orleans, or maybe you would find her with her aunt in Chicago.

The supposition proposed to the addressee is to shift his point of view to that of John. The sequel contains a series of modals: would, might, maybe, each of them understood relative to the point of view of John at that point. The disjunction suggests alternative possible answers to the QUD introduced by the preceding interrogative, might and maybe enumerating the possibilities epistemically accessible to him.

- Third person in Free Indirect Discourse (FID): (Roberts 2015)

(178) John pondered his situation. Where was Clarissa now? She might be in New Orleans with Sidney. But she might be in Chicago.

Here, as typical in FID and contrasting with (177), there is no explicit suggestion to the addressee to shift point of view from that of the speaker to that of the reported agent, John (Eckardt 2014). Assume that the novel takes place in the 19th century, but we are reading it in the 21st. Then the use of now in querying Clarissa’s location (e.g., if) suggests the shift.

- Multiple bodies of evidence, one more “objective” than the other:

(179) Given the results of the DNA tests, John might be the thief. But if we take the eyewitness seriously, John can’t have been the thief. [von Fintel & Gillies]

The speaker in (179) explicitly suggests first one body of evidence, the DNA tests, and then another, the eyewitness’ account, which make contradictory predictions about whether John might have been the thief. This is a case where the speaker (and probably the addressee, as well) knows the DNA evidence, and that it can be interpreted to so that John might have done it. The if-clause in the last sentence suggests that the speaker and addressee haven’t yet decided whether to take the eyewitness’ evidence seriously; then can’t in the main clause
reflects the consequences of the doxastic state they would be in if they were to suppose that that evidence is correct. Another possibility is that can’t is anchored to the eyewitness himself, as would be explicit in But according to the eyewitness. It may not make a truth conditional difference which anchor we chose for can’t, since the same body of evidence would be brought to bear on the interpretation in both cases. But the evidential semantics of EMA can’t would argue for the speaker(+addressee) anchor: If the eyewitness saw the theft, that person would have the best available evidence and the speaker could simply report, But according to the eyewitness, John was the thief. So here the EMA is used because from the interlocutors’ point of view the evidence is hearsay. This example is especially nice because it shows how context not only suggests the identity of the intended discourse center anchor, but in some cases may suggest the particular body of evidence which is supposed by that agent, going beyond his firm beliefs.

von Fintel & Gillies (2007a) also suggest that objective epistemic modals might be anchored to non-human knowledge sources (logs, charts, etc.), and that in such cases this leads to a more “objective” interpretation where the speaker’s opinion doesn’t count:

(180) The hulk might be in these waters. [von Fintel & Gillies 2007a, after Hacking 1967]

But whatever the knowledge sources for (180), if the speaker knows of them, and the propositions they make true, then s/he’s privy to that evidence. If there’s an implication that the body of evidence is “objective” or “consensual” or generic, then the modal statement may be taken to have more force—after all, evidentiality is about the speaker’s judgment of the quality of her evidence for certain beliefs, and we are more likely to believe that for which we have better evidence. Such examples seem particularly prone to the ‘we’ interpretation we saw earlier, which arises in a group where the Common Ground contains such “objective” evidence; and I suspect the common consensus about the value of the information strengthens the sense of objectivity, as well. But (180) is still evidential, and the non-human provenance of the evidence doesn’t mean that the anchor itself—the epistemic agent who has that information—is non-agentive.

The pattern of occurrence observed above is that typical of anchors for indexical expressions generally, wherein the presupposed center who anchors the indexical’s interpretation varies systematically with the context. With EMAs, not only does the anchor vary, but it then gives clues about the body of evidence which is intended to restrict the modal’s domain.

Summarizing, we see that for EMA domain restriction:

• in root declaratives, evidence to which the speaker or inclusive ‘we’ is privy is the default
• in root interrogatives, evidence to which the addressee is privy is the default, i.e. interrogative flip (167)

5Note (Palmer 1990) that can’t has an epistemic interpretation, while can typically does not. And that epistemic can’t always has scope order ¬◊, whereas epistemic mightn’t (as in Mightn’t George be the thief?) has only ◊¬. Thus, epistemic can arguably is a Negative Polarity Item (probably of a particularly restrictive sort, requiring negation), in its epistemic interpretation (but not in the deontic or dynamic) requiring a downward entailing context for felicity.
• in imperative conditional protases, as in (168), evidence to which the addressee is hypothetically privy is the default, as this information is useful for determining the applicability of conditional advice or directions
• in attitude complements, evidence to which the attitude agent is privy is the default
• with explicit adverbials like those in (169), (170) and (171), we see a shift similar to that in attitude complements.
• in modal subordination the evidentiality is anchored to the relevant agent in the epistemic subordinating context
• in FID the default is the evidence of the doxastic agent whose perspective is adopted by the author.

In almost all of these cases, the default can be pragmatically over-ridden. The resolution of the anchoring presupposition of an EMA is anaphoric, hence subject to the usual constraints on anaphora resolution: familiarity and salience of the intended discourse referent antecedent, pragmatic plausibility, and coherence of the resulting resolution. A useful contrast is between the anchoring of EMAs in attitude complements and those governed by explicit shifting adverbials. With the adverbials, the shift appears to be obligatory, conventional, whereas it is only the pragmatically governed default in the attitude complements. In the latter it is just that the agent of the attitude is the most salient center in that context, so that anaphora resolution naturally points to that center, all other things being equal.

This, then, is evidence of just the sort of variable anchoring—limited and most often to the agent of ©*, the speaker at utterance time—that we found with indexicals and other doxastic perspectivals like come. Certainly, the interpretation of an EMA crucially relies on the doxastic perspective of the relevant anchor. And it seems that this anchoring is presuppositional and not part of the proposition expressed—the truth conditions are only relativized to the relevant belief state, not the agent whose state it is.

One other reflex of presuppositional anchoring to the doxastic perspective of a discourse interlocutor, also observed with the indexicals, is the strong tendency for an apparently wide scope reading. We find that in EMAs as well.

Though it is often claimed that epistemic modals cannot take narrow scope relative to other operators (Palmer 1990, Brennan (1993), Hacquard 2013), von Fintel & Gillies (2007b) offer the following examples to argue that this is not always the case:

(181) Bill thinks that there might have been a mistake. [Attitude predicate over modal]
(182) Where might you have put the keys? [Question over modal]
(183) The keys might have been in the drawer. [Past over modal]
   present perfect: shifts the time of the center to one earlier than ST
(184) There can’t have been a mistake. [Negation over modal]
   NPI epistemic can’t
The semantics proposed above offer a different perspective on this matter. It’s true that non-intensional operators tend not to scope over epistemic modals, leading to von Fintel & Iatridou’s (2003) Epistemic Containment Principle:

The Epistemic Containment Principle: Epistemic modals tend to take obligatory wide scope with respect to a wide class of quantifiers.

In support of this, they offer (185):

(185) Every candidate might win. [von Fintel & Iatridou 2003]

They claim that on the epistemic understanding of *might*, (185) “has no true reading if there is at most one winner of the election, even if there is no candidate that we know is going to lose.” There is a possible reading of (185) with the universal wide, but the modal in this reading is dynamic, not epistemic: ‘each candidate is capable of winning’. I agree that there’s a lot of interference in this example from the dynamic interpretation of the modal, but perhaps that’s partly because the dynamic reading is so very closely related in this case to a speaker-anchored epistemic.

But to control for interference from that possibility we can select a predicate that doesn’t lend itself to being something someone is ‘capable of’, i.e. isn’t a naturally dynamic complement. I also use a non-universal quantifier:

(186) [spoken by a teacher whose students are having trouble with standardized tests:] Given what I know about these kids, many of them might be amenable to working with a tutor.

I readily get an interpretation for (186) which can be paraphrased as ‘there are many $x$ s.t. $x$ might be amenable to working with a tutor’, and in fact find that more natural than the ‘it might be that many are amenable’ interpretation. So the epistemic containment principle is not particularly robust.

Also, as we might expect on the present account, it’s very easy to get the wide scope reading of a universally quantified NP if that NP ranges over agents of an embedding attitude predicate that takes an EMA in its complement:

(187) Every candidate believed, on the basis of his own polls, that he might win. Some of them were right, and others wrong.

But here, of course, the agents in question act as discourse centers that satisfy the doxastic anchoring presupposition of *might*. And, as with presupposition satisfaction generally, if some operator takes wider scope than the anchor for a modal, that operator will take wide scope over the modal as well:

(188) Every year most candidates believed, on the basis of their own polls, that they might win.

with scope order: *every - most - might*
In other words, one key to the tendency to wide scope of EMAs is actually the pseudoscope typical of constituents whose lexical content involves an anaphoric presupposition. Recall that we considered cases like the following, involving anaphoric Bridging:

(189) If you park a car on a steep hill, engage the emergency brake.
(190) The gears on this car tend to slip. If you park on a hill, engage the emergency brake.

In (189), the antecedent of the anaphoric description *the emergency brake* is understood to be the (weakly familiar, Roberts 2004) emergency brake of the car introduced in the *if*-clause. Since the antecedent is understood to be non-specific—the arbitrary car parked on a steep hill, and hence falls under the scope of the operator associated with the bare conditional (per Kratzer, a universal modal operator), the definite description seems to take narrow scope relative to that operator, as well. But in (190), the weakly familiar antecedent is the emergency brake of the car referred to with the demonstrative *this car*. Since that antecedent is introduced outside of the conditional—we’re talking about a particular emergency brake, a singular entity—the particularity extends to the denotation of the emergency brake itself. Thereby, though the definite description doesn’t get displaced in Logical Form, it *seems* to have wide scope over the conditional operator just because its antecedent does. This is pseudoscope.

Similarly, modal anchoring for EMAs as described here is most often to one of the interlocutors, as reflected in the initial plausibility of DeRose’s (1991) Speaker Inclusion Constraint. This is like the anchoring of so-called “pure” indexicals like English *I, we or you*, as reflected in the use of global-only anchoring for indexicals in Kaplan’s (1979) account. It is global anchoring which gives rise to pseudo-wide-scope, and the default tendency for global anchoring which gives rise to the initial plausibility of the claim that EMAs always take wide scope. The examples considered in this section argue that this initially plausible claim is false.

Summarizing: Epistemic Containment is more a tendency than a “principle”, and it reflects the characteristics of default anchoring and projection: pseudo-scope.

4. **Perspective and Localism in Linguistics**

I have made a case that the interpretations of indexicals co-vary systematically with a certain kind of contextual factor, the doxastic perspective adopted by the speaker at that point in the discourse. But the proposed notion of perspective has much broader ramifications for natural language understanding and, hence, for semantics and pragmatics. A case can be made that a generalization of the notion of perspective plays a role in a large and superficially quite diverse class of expressions of which the indexicals are just a special subset: the perspectival *expressions*.

If time permitted, we might investigate a much longer list of candidates for perspectival expressions. These include:
• so-called “predicates of personal taste” like *tasty, fun*, where some presupposed agent’s subjectively justified judgment plays a crucial role in the truth conditions (e.g., see Mitchell 1986, Lasersohn 2005 and much subsequent work)
• logophoric pronouns (Sells 1987, Pearson 2013)
• conjunct-disjunct marking in Tibeto-Burman languages, including Newari (Delancey 1992; Hargreaves 2005; Zu 2015, to appear; Coppock & Wechsler 2016)

The basic idea I’ve been working with here—that one can find a formal parallel between a long list of expressions which depend for their interpretation on some implicit perspective—grows out of a long tradition in linguistics, with roots in the ancient grammarians, known in contemporary literature as *localism*:6

[Localism is based upon] the hypothesis that spatial expressions are more basic, grammatically and semantically, than various kinds of non spatial expressions (...). Spatial expressions are linguistically more basic, according to the localists, in that they serve as structural templates, as it were, for other expressions; and the reason why this should be so, it is plausibly suggested by psychologists, is that spatial organization is of central importance in human cognition” (Lyons 1977:718).

Lyons’ localism starts with the application of this concept to the semantics of deixis, locatives and tense, but extends it as well to what he calls the use of demonstratives in “empathetic deixis”:

The conditions which determine this empathetic use of the marked member of these deictically opposed demonstratives and adverbs are difficult to specify with any degree of precision. But there is no doubt that the speaker’s subjective involvement and his appeal to shared experience are relevant factors in the selection of those demonstratives and adverbs which, in their normal deictic use, indicate proximity. At this point deixis merges with modality... (Lyons 1977:677)

A similar impulse can be observed in otherwise quite wide divergent modern approaches to semantics, from the functionalist work of Talmy, who extends localism to temporal and causative contexts: “Situations that involve state and change of state seem to be organized by the human mind in such a way that they can be specified by structures homologous with motion structures” (1975:234), through Fillmore’s (1975) *Lectures on Deixis*, Langacker’s *Space Grammar* (1978), Mitchell’s (1986) ambitious attempt to provide a unified framework for the semantics of point of view, Sells (1987) on logophoric pronouns (and the more recent work on these due to Pearson 2013), Doron (1991) on point of view, Speas & Tenny’s (2003) introduction into the Logical Form of sentences in Universal Grammar of a syntactic projection for Point of View, and the most recent work on perspective and the *de se* cited throughout, much of which hints at the general notion of perspective (especially that of Moltmann 2012 and Pearson 2013), and much more besides. But, instructive and insightful though all this work is, none of it quite captures the central notion of perspective in such a way as to found a satisfying general account. Why has this been so difficult to pin down?

---

6See Fortis (2012) for a nice history and overview of localism in linguistics, from which several of the references cited here were drawn.
I would argue that, as the range of expressions alluded to suggests, the semantics and pragmatics of perspective in human discourse is reflected throughout language, any language, but in so deep a fashion as to only rarely be truth conditionally evident: It is not generally part of the content that we proffer each other in our usual intercourse: what we talk about. Rather, it is a central supposition that grounds much of what we say and understand. That is, perspectival anchoring of interpretation is fundamentally presuppositional, and hence pragmatic, and only surfaces here and there in explicit truth conditional semantic content. But once we identify it for what it is, we see evidence of it throughout that content. All the authors cited just above have had a genuine intuition of this underlying unity.

The indexical expressions as characterized above, are dependent on a specifically doxastic perspective, and thus constitute a proper sub-class of the perspectival expressions:

\[
\text{Indexicals} \subseteq \text{Perspectival Expressions} \subseteq \text{Anaphoric Expressions}
\]

Of course, the evidence for these claims goes well beyond the purview of the current work. But though focused on the interpretation of indexical expressions, a deeper goal of this work has been to offer a preliminary characterization of the essential features of the perspectival anchoring of interpretation and to offer some simple tools for modeling it and tracking how perspective changes in systematic ways across discourse. This brief and partial inventory of perspectival expressions is intended to sketch the cross-categorial, cross-linguistic vista which this work opens up. I hope this will help to stimulate further cross-linguistic investigation.

References not on the course site:


https://www.academia.edu/1125714/A_Flexible_Contextualist_Account_of_Epistemic_Modals


