Future Displacement and Modality

Fabrizio Cariani
University of Maryland, College Park

[Forthcoming in Oxford Handbook of Philosophy of Language 2nd Edition, Ernie Lepore and Una Stojnic (eds.)]

Among the most remarkable features of human language is that it offers speakers the ability to represent situations that go beyond the here and now of any one conversational situation. Linguists refer to this feature as displacement (Hockett et al., 1960; von Fintel and Heim, 2011) and go on to observe that it comes along a variety of distinct dimensions. For example, we are evidently able to talk about events from which we are spatially displaced (events happening elsewhere), temporally displaced (events happening at different times), or modally displaced (events that are not actual, but may be reconstructed as happenings in other possible worlds).

This chapter focuses on temporal displacement. Like its parent category, temporal displacement is also realized in many different ways, often even within the same language. Many, but not all, human languages incorporate a tense system—a grammatical system consisting of inflectional morphemes whose job it is to locate the subject matter of an utterance with its time (Comrie, 1976, p.9). Another option is for temporal displacement to occur by means of lexical items. An example of this sort of device, and one which will be prominent in this discussion, is the English auxiliary will. A final option for temporal displacement is complex phrases: English examples include on January 1st, 6120, or next Monday. These complex phrases are peripheral to the present aims and will not be discussed again, leaving the focus on grammatical and lexical devices of temporal displacement.

Theories of future displacement will be surveyed here in a rather opinionated fashion. I will rely heavily on a guiding hypothesis already articulated within an

---

Thanks to Justin Khoo, Angelika Kratzer, Paolo Santorio, and Alexander Williams, for conversations, exchanges and feedback.
existing body of research. This is the claim that the linguistic and cognitive implementation of future displacement is fundamentally different from the linguistic and cognitive implementation of past displacement. In particular, future-directed discourse is interwoven with modality in a way that is simply not true of past-directed discourse.

The structure of the chapter is as follows. §1 establishes some history and background. §2 aims to settle down on operative definitions of some core terms such as ‘tense’, ‘aspect’ and ‘modality’. §3 investigates the varieties of ways in which one might claim that there is an important connection between future reference and modality. §4 surveys the main diagnostics that lead linguists and philosophers to accept modal treatments of future discourse. §5 contrasts two styles of semantic implementation of a modal future thesis and illustrates how the temporal and modal elements might live together in a single semantic analysis of temporal modals. §6 considers and rejects an objection to the main working hypothesis.

1 Greatest hits

1.1 The tense logic paradigm

The standard semantic picture of tenses is built on the foundation of tense logic in the style of Prior [Prior, 1957, 1967, 1969]. This picture envisages tense operators, respectively for past and future. Past and future tenses are understood as operating existential quantification over, respectively, past and future instants. Thus, the truth conditions of Jodie Foster won an Oscar are roughly equivalent to At some point in the past, Jodie Foster won an Oscar. By contrast, the truth conditions of Jodie Foster will win an Oscar are roughly equivalent to At some point in the future, Jodie Foster will win an Oscar. We could introduce sentential operators, FUT and PST to express these concepts, and endow them with the semantics in (5-a)-(5-b)

(1) a. $[\text{FUT } A] = \lambda p \lambda w \lambda t. \exists u > t, p(w)(u)$
b. $[\text{PST } A] = \lambda p \lambda w \lambda t. \exists u < t, p(w)(u)$

A piece of notational convention: I use the variable $p$ to range over functions from world/time pairs to truth-values.

---

1 Inevitably, the chapter leverages claims and arguments drawn from my recent book on this topic [Cariani, 2021]. However, it is no mere repetition, or summary, of those arguments. Though the book is still hot off the press at the time of the present writing, my theoretical stance has shifted somewhat, and unsatisfactory developments in the book’s architecture can be improved if we settle on definitions of some fundamental terms.
Although the tense logic analysis was not introduced by Prior as a piece of natural language semantics, it has made its way into natural language semantics thanks to contributions such as Montague (1973) and Dowty (1979). Once it is turned into a claim about natural language, it appears immediately desirable to add some bells and whistles to the tense-logic analysis. Thus, in some contexts, the quantification associated with the tenses might be restricted to smaller intervals—smaller, that is, than the entire future or past. A famous example in Partee (1973) can be used to illustrate this point. Suppose that, as I am driving on the freeway, I utter:

(2) Val didn’t turn the stove off.

The tense-logical analysis provides two interpretations for (2) depending on whether the tense is supposed to outscope the negation (3-a) or viceversa (3-b).

(3) a. past [not [Val turn the stove off]]
   b. not [past [Val turn the stove off]]

If given the tense-over-negation structure, (3-a), the sentence means that at some past time I didn’t turn off the stove. That reading is evidently too weak. If given the negation-over-tense structure, (3-b), the sentence means that at no past point in time I turned off the stove, which is evidently too strong. Partee’s gloss on this example is that “[t]he sentence clearly refers to a particular time [...] whose identity is generally clear from the extra-linguistic context” (Partee, 1973, p. 602-603).

It is important to pause here and notice that there are two ways of developing Partee’s suggestion. Narrowly understood, it points to an analysis in which context directly provides the temporal coordinate, obviating the need for existential quantification altogether. So, (2) would mean that the contextually set past time does not feature an event consisting of Val turning off the stove. This narrow construal is what ordinarily goes by the “referential analysis of tense”. Letting $g$ be an assignment function and giving the tenses an index that the assignment function can apply to, we might provide the following formal statement of this analysis:

(4) a. $[\text{FUT}_{\text{t}}A]^{\delta} = \lambda p \lambda w \lambda t : g(i) > t. p(w)(g(i))$
   b. $[\text{PST}_{\text{t}}A]^{\delta} = \lambda p \lambda w \lambda t : g(i) < t. p(w)(g(i))$

---

2 This matter is aptly discussed in Ogihara (2007). Ogihara questions whether Montague can really be correctly interpreted as aiming to contribute to an understanding of tense in natural language.
A broader interpretation is also compatible with Partee’s remark: could reasonably be analyzed as involving restricted existential quantification (Ogihara, 1996). The quantification would come from the lexical entry of past tense, the restriction would be set from context, once again via the assignment function. Here it might be assumed that \( g(i) \) (the value of the domain variable for temporal quantifiers) is an interval (i.e. a non-empty convex set of instants) to be used in restricting the quantification over times.

\[
(5) \begin{align*}
a. & \quad \lbrack \text{FUT}, A \rbrack^\delta = \lambda p \lambda w \lambda t. \exists u \in g(i), u > t \land p(w)(u) \\
b. & \quad \lbrack \text{PST}, A \rbrack^\delta = \lambda p \lambda w \lambda t. \exists u \in g(i), u < t \land p(w)(u)
\end{align*}
\]

Partee (1973), and then Kratzer (1998), also famously noted other important similarities between tenses and pronouns. Like pronouns, tenses can be bound by prior elements of the discourse. Thus:

\[
(6) \quad \text{Karen played drums and Richard played keys}
\]

can be understood either as describing two independent events happening at different points in time, or more plausibly as describing two events happening at roughly the same time—the past tense on the second verb picking out anaphorically the past tense on the reference time introduced by the past tense sang. Theories like the ones in (4) and (5) both seem well positioned to captured this anaphoric behavior—or at least to reduce theoretical treatment of such anaphoric behavior to more general theoretical treatments of anaphora.

1.2 The Reichenbach paradigm

One problem for the tense logic analysis — one that was already clear to its original developers — is its inability to model relative tense, such as pluperfect and future perfect.

\[
(7) \quad \text{By the time Sonya arrived, I had prepared the sauce.}
\]

\[
(8) \quad \text{By the time Sonya arrives, I will have prepared the sauce.}
\]

Consider the phrase *I had prepared the sauce* in (7). It is inadequate to analyze this as \( \text{PST PST}(I \text{ prepare the sauce}) \). It is similarly inadequate to analyze the phrase *I will have prepared the sauce* in (8) as \( \text{FUT PST}(I \text{ prepare the sauce}) \). In both cases, the analysis of the complex tenses requires some way of anchoring some past-shifting device to Sonya’s arrival. This sort of anchoring does not seem to be within the expressive capacity of linear tense logic.
In a short but extremely influential discussion, Reichenbach (1947, pp.287-288) proposed understanding tenses, including complex tenses in terms of a complex system of references to three points in time. These are: ‘E’ for the event, ‘R’ for a time of reference, and ‘S’ for speech time. Thus in the speech time is whatever time the sentence is uttered at, the reference time is the time of Sonya’s arrival, and the event time is the time of the sauce preparation. So, pluperfect corresponds to an $E \rightarrow R \rightarrow S$ structure, while future perfect corresponds to an $S \rightarrow E \rightarrow R$ structure.

Despite the enormous influence of this paradigm — especially through refinements such as Comrie (1976), I think it is worth setting it aside for our purposes for two reasons. First, there are complexities and unclarities when it comes to how these associations between complex tenses and Reichenbach's temporal anchors are to figure within a compositional semantics (Von Stechow, 1995). Second, we can go a relatively significant theoretical distance by just reflecting on simple tenses, which do not require us to articulate the complexities of the Reichenbachian paradigm.

2 How do we distinguish between, tense, aspect and modality?

Part of the present task is to understand the relative role of tense and modality in powering future reference. This task requires us to have at least an operative understanding of the distinction between tense and modality, and furthermore of how both differ from aspect.

My criteria for reaching such an understanding are fairly loose. I aim to reconstruct these distinctions by taking the practice of linguistics at face value. Because the practice of linguistics — the whole lot of competing views, incompatible ambitions and diverse frameworks — is not entirely coherent, it will occasionally be necessary to do some clean up work to guarantee consistency. Moreover, in some cases, we may have to land on characterizations that adjudicate controversial distinctions, especially when it comes to the relative boundaries of tense and aspect. In sum: we want something like a definition to guide our thinking; we want the definition to be primarily informed by linguistic practice; but we also

---

3Here is Reichenbach’s passage:

Let us call the time point of the token the point of speech. Then the three indications, ‘before the point of speech’, ‘simultaneous with the point of speech’, and ‘after the point of speech’, furnish only three tenses; since the number of verb tenses is obviously greater, we need a more complex interpretation. From a sentence like ‘Peter had gone’ we see that the time order expressed in the tense does not concern one event, but two events, whose positions are determined with respect to the point of speech. We shall call these time points the point of the event and the point of reference.
want to be wary of putting too much stock on some marginal classifications it might entail.

In the background, I assume a broadly truth-conditional picture of meaning. In turn, it will be best to view truth-conditions as having some structure. The truth-conditions of simple declarative sentences typically involve locating events, states, and processes across temporal and modal space, however conceived. Because the disjunctive category of “events, states and processes” occurs frequently, it is helpful to have a technical term for it. I will follow Comrie (1976) here and use “situations” as a catch-all term to cover these.

Enough with the throat-clearing. Against this background we can introduce a rough characterization of tense. As we will use the label, tense is a grammatical system consisting of inflectional morphology whose job it is to anchor situations to certain times, defined by relation to the utterance time, in the process of fixing their truth conditions. The upshot of claiming that the sentence I played soccer involves past tense is that it involves the application of the bound morpheme -ed to the verb (a bound morpheme is one that only occurs as a proper part of a word); the semantic role of this morpheme is to locate the situation emerging from the whole verb phrase in the past. The requirement that tense be part of a grammatical system of bound morphemes directly implies that the devices that achieve future reference in English—auxiliaries like will and phrases like going to—are not tenses. Here is Huddleston and Pullum (2002) summarizing their rationale for this classification in their Cambridge Grammar of the English Language.

We distinguish sharply between the grammatical category of tense and the semantic category of time. [...]. Once this distinction is clearly drawn, it is easy to see that English has no future tense: will and shall belong grammatically with must, may, and can, and are modal auxiliaries, not tense auxiliaries.

More generally, tense—and future tense specifically—is not present in all languages (Tonhauser, 2015; Bochnak, 2019).

Evidently, not all the bound morphemes in verbs’ inflectional systems convey temporal information. Aspect, mood, and gender features are all examples of

---

4What I assume does not entail that every part of language is well understood only by reflecting on truth-conditions. The assumption is that an account of linguistic meaning can helpfully feature a module that captures how the truth-conditions of various simple declarative sentences depend on the semantic values of their components. Even this much might be incompatible with certain “internalist” conceptions of semantics, on which semantics is not in any interesting sense about truth-conditions but about relations between concepts (Pietroski, 2018). The problem is not that what I have to say cannot be said in that idiom. It is instead that what I have to say must be said in some idiom or other and I am not very good at speaking that other idiom consistently and precisely.
other features that might be targeted by inflectional morphology across languages. Roughly speaking, *aspect* is a system whose job it is to specify the internal temporal structural properties of eventualities. A characterization on these lines is Comrie’s (1976) claim that the category of aspect involves “[t]he different ways of viewing the internal temporal constituency of a situation”. Thus the difference between *I read a book* and *I was reading a book* is largely one of aspect. Unlike tense markers, there is no general demand that aspect markers be grammatical devices, as aspectsual constraints might emerge directly from the lexicon (Hamm and Bott 2021). The line between aspect and tense can sometimes be blurry. For example, prospective aspect can easily get mixed up with future tense. The difference here can be helpfully illustrated by the Reichenbach framework, since the job of prospective aspect is to locate the relevant situation in the future of the *reference* time as opposed to the future of the *utterance* situation.

It was a little painful to settle on (provisional) characterizations for tense and aspect. But it is even more painful to settle on one for modality. One point that conjures substantive agreement is that modality is not a grammatical category, as modals across language appear in a disparate variety of syntactic categories (Kratzer 1981, Kratzer 2012, §2.2). At the same time, while it seems clear that modality is a semantic category, it is not be identified with a type of lexical entry. After all, modals seem fairly heterogeneous when it comes to the kinds of lexical entries they allow. In Cariani (2021, §3.1) I note that most of these arguments really only require an interesting sufficient condition. I offer the following one: consider a semantic theory $T$ that includes a type for possible worlds: then the modals in $T$ are exactly those primitive expressions of the language whose semantic value belong to a type that is constructed in any way by using the basic type of worlds. For semantic theories that are not based on a type-theoretic universe, or theories that are not based on a universe involving worlds this will need revision. But for our purposes it will be enough.

3 The varieties of modal involvement in future displacement

In the introduction, I outlined a basic hypothesis — that the linguistic devices that power future displacement have modal features. The present section aims to get clearer about what that hypothesis comes down to, how broadly it ranges over languages, and whether claiming that a certain expression has a modal component rules it out from being a tense. It is helpful to approach this as a “provincial”, language-specific matter and then broaden our gaze towards a more universal perspective.

Before doing that, it is important to emphasize that it is a mistake to confuse modal theories of the future with analyses of indeterminism in context of branch-
ing time (Thomason, 1970; Belnap et al., 2001; MacFarlane, 2003). The basic point to keep in mind is that any modal analysis of the future must be a thesis about the meaning of certain lexical items. The branching time hypothesis is emphatically not a thesis about the meaning of any lexical item, and it is independent of the meaning of expressions like will. For this reason, MacFarlane (2003, 2014) dubs the level at which supervaluationist (and relativist) theories operate postsemantics. The point is best seen in action as we consider the logical consistency of accounts that divorce the question of modality from the question of branching. It is logically consistent to analyze will as a modal against a picture on which worlds have linear structure—just as you might do for any other modal. It is also logically consistent to analyze will with linear clauses in the spirit of (5-a)-(5-b) against a branching background, as is the point of supervaluationist semantics. 

Are there are languages that lack a future tense? Given our set-up, the answer is not just yes, but obviously yes. As noted in section 2, while English typically recruits inflectional morphology to draw the distinction between past and present, it does not do so for expressing futurity, and instead recruits lexical items, such as will to capture future displacement. Indeed, the English tense system is ordinarily described as only having two tenses: past and non-past. For example, Chomsky’s (1957) formal representation of the grammar of English lumps will with the modals and not the tenses. 

English is far from being the only language that succeeds in future displacement without recruiting any future tense markers, and specifically without recruiting inflectional marking on the verb. Dahl and Velupillai (2013) report that about half (50.4%) of 222 surveyed languages are like English in this respect. Needless to say, these languages differ amongst themselves in the types of devices they recruit (see Bochnak, 2019, for an introductory perspective on this variety).

We ought to be careful, however, to avoid the simple identification between lacking a grammatical system for future marking and more substantive claims about the connection between modality and futurity. Such identifications might fail in both directions. To start, there are devices of future displacement that are not tense but still do not appear to involve modality. In many languages, including Italian and German, sentences in the simple present tense manage to easily express futurity (Grönn and von Stechow, 2016; Bochnak, 2019). Suppose that my tennis racquet has a frayed string. Unlike English, Italian allows:

For more on these themes, see chapters 2 and 3 of Cariani (2021).

While there is not much commentary around this within the confines of Syntactic Structures, the reasons for this classification do not appear to be semantic. One such possible non-semantic reason is the historical and morphological relationship between will with the bona-fide modal would. That will is a modal in any stronger sense than this requires more specific arguments, and I shall turn to those in the next section.
Quella corda si rompe nella prima ora di gioco.

"That string breaks within the first hour of play

[That string will break within the first hour of play]

In English, this kind of use of unvarnished present tense to introduce futurity is only allowed in a much more restricted class of “scheduled” events.

Federer plays Nadal tonight.

These uses are called futurates, and demand their own analysis that is linked to, but independent of, the analysis of future reference. (see Kaufmann, 2005; Copley, 2008, 2009; Rullmann et al., 2021 for analyses of futurates) What is distinctive about languages like Italian and German is that future displacement without any devices of future displacement is far more pervasive, and stretches beyond examples which are plausibly described as “scheduled”. The upshot: just because some future reference is not powered by tense it need not follow that it is powered by modality.

In the converse direction, we need not automatically assume that the modal future hypothesis is false for languages that have a future tense. Classifying an expression as belonging to certain lexical and syntactic categories (e.g., that it is, or isn’t, a tense) is not sufficient to classify it semantically. Importantly, under our present characterizations of tense and modality, nothing prevents an expression from exemplifying both. Once we have diagnostics for modality, we will need to consider separately whether the bound morphemes that characterize future tense in these languages ought to be also understood as modals.

This insight is important as we turn our attention to the universal perspective.

I noted that 50.4% of the world-languages do not recruit bound morphemes as devices of temporal displacement. Evidently, the remaining 49.6% do recruit such bound morphemes. Languages are not uniform in whether they have tenses, and those that do are not uniform about which tenses they have. If the modal future hypothesis were the claim that no languages can have tenses, it would appear to be be falsified by these languages.

However, our earlier conclusion looms large at this point of the dialectic: our characterizations of tense and modality allow some expressions to be both tense markers and modals. An entertainable universal version of the modal future hypothesis might consist of the claim that devices of future displacement across languages, whether tense or not, also have a modal component. I do not know of attempts to tackle this kind of universal investigation, but I am substantially more
bullish towards the possibility that it might have be answered in the affirmative. The core diagnostics for modality I identify in the next section seem to carry over to (at least some) languages that have future tenses. Still, any kind of universal investigation of this sort demands a kind of cross-linguistic research that I am not equipped to carry out on my own and that I have not seen carried out by other authors. (But see Tonhauser[2015] and Bochnak[2019] for relevant, cross-linguistically informed discussions.)

4 Deeper into the provincial modal future thesis

In section 3, we raised but deferred the question of how to argue that English will is a modal. To my mind, the clearest diagnostic that favors the classification of will as a modal is the availability of subordinated readings. Modal subordination is a linguistic phenomenon whereby modals in a discourse can be restricted in their interpretation by prior elements of the discourse. At the most basic level, the diagnostic is just that will patterns with modals in triggering modal subordination. Past tense does not allow this. A classic example that shows modal subordination at work with modals is Roberts’s discourse:

(11) A wolf might come in. It would eat you first!

In each group of examples below, the a-sentence has will showing modal subordination, while the second shows the unavailability of modal subordination with the past tense (The first group of examples is drawn from Cariani (2021) while the second is from Roberts[1989], p. 683.)

(12) a. If Katie travels to Berkeley, she will shop at Amoeba records.
    She will buy a boxed set and a dozen used LP’s.
    b. If Katie traveled to Berkeley, she shopped at Amoeba records.
        # She bought a boxed set and a dozen used LP’s.

(13) a. If Edna forgets to fill the birdfeeder, she will feel very bad.
    The birds will get hungry.
    b. If Edna forgot to fill the birdfeeder, she felt very bad.
        # The birds got hungry.

The ‘#’ marker attached to the b-sentences indicates that they are either defective, or if non-defective that they are not interpreted with a subordinated reading.

Roberts[1989] is the locus classicus for modal subordination. Klecha[2014], Cariani and Santorio[2018], Cariani[2021] all claim that modal subordination is a critical diagnostic for modality.
These examples show clearly that will can inherit a restriction from prior elements in the discourse, in a way that past tense simply does not allow.

In (12-a) and (13-a), the restricting proposition comes from a conditional antecedent. In each of these cases, the propositional content of a declarative sentence is used to restrict the modal occurring later in the discourse. However, there are remarkable and important examples in which modal subordination happens across clause types (all these examples are from Cariani 2021 §3.4). The second sentence of (14) is naturally interpreted as claiming that if one does throw paper towels in the toilet, it will clog and overflow. A restricting proposition is extracted from the imperative clause occurring earlier in the discourse.

(14) Please do not throw paper towels in the toilet. It will clog and overflow.

The same phenomenon can be illustrated by questions. Suppose I am reading the story of Cinderella to my children, but we have to stop before the end of it. Here are three possible discourses they might add as commentary.

(15) a. Does she stay at the ball past midnight? The carriage will turn into a pumpkin!
    b. Did she stay at the ball past midnight? # The carriage turned into a pumpkin.
    c. Did she stay at the ball past midnight? The carriage could turn into a pumpkin!

These examples of modal subordination across clause types are dialectically important. They help overcome a certain style of objection against the examples in (12-a) and (13-a). The objection is that with the right intonation, it is possible to get apparently subordinated readings even with past tense. For example, one might hear something like the continuation in (13-a) as appropriately restricted, if one’s intonation made it clear that we are in some sense “continuing” working under the hypothesis from the previous sentence in the discourse. However, there is an easy alternative explanation that does not involve treating these as cases of modal subordination. We all agree that it is fine to say If Edna forgot to fill the birrfeeder, she felt very bad and the birds got hungry. Now it seems to me that it is possible to hear the right kind of hesitation in (13-a) as a suppressed and, resulting in a speech that sounds like modal subordination but is not. This suggests that these apparently subordinated readings do not provide convincing evidence that past tense allows modal subordination. If that’s true, cases of modal subordination across clause-types become important test cases, because they allow us to control for this phenomenon. After all, in those cases, the subordination is not well
understood as producing a reading equivalent to the (surface) conjunction of the two elements of the discourse. However, I do not know of any way of making the past-tense variants of modal subordinations across clause types go through felicitously.\(^8\)

The availability of modal subordination is not the only argument in support of the modal future hypothesis. Two other arguments that are worth flagging. While these arguments are dialectically weaker, they complete the picture that is established by the modal subordination arguments. Furthermore, they combine with the the modal subordination arguments to shape modal analyses of future discourse. The first of these auxiliary reasons for endowing \textit{will} with a modal semantics is that it has \textit{bona fide} modal uses (Palmer, 1987; Enç, 1996; Huddleston and Pullum, 2002; Cariani and Santorio, 2018; Cariani, 2021). Imagine being at a restaurant at the end of a meal, seeking to compliment the chef. You might inquire with a waiter, and, after a short wait, receive the response:

\begin{enumerate}
\item[(16)] The chef will be in the kitchen right now.
\end{enumerate}

The distribution of these present-directed uses of \textit{will} is somewhat limited in English. For example, this use of \textit{will} is not compatible with eventive prejacent. That is, except in very special circumstances,\(^{17}\) is defective:

\begin{enumerate}
\item[(17)] They will win right now.
\end{enumerate}

The key observation here is that this use of \textit{will} appears to have a modal component. If that is agreed, it would be surprising if \textit{will} behaved as a modal in some environments, while not behaving as a modal in others\(^9\).

---

\(^8\)In a different theoretical context, Starr (2014) highlights some examples that might make trouble for the dialectic I have just developed.

\(^{17}\)If these examples—\textit{advertising conditionals}, as Starr labels them—are construed as involving modal subordination then they might show some kind of subordination across clause types targeting past tense clauses without overt modals. As I understand Starr, he is not suggesting that the examples in \textit{advertising conditionals} are instances of modal subordination, but examples of enriched interpretations that are derived by more general principles of discourse coherence. But the dialectic here may admittedly become more complex once we start asking whether those principles of discourse coherence might account for garden variety cases of modal subordination. I won’t pursue this thread here.

\(^9\)It might be objected here that what \textit{will} has in common with \textit{must} is only a kind of evidential requirement. Both \textit{will} and \textit{must} are evidential, in that they demand a certain kind of evidence for their prejacent. Although these evidential requirements are not exactly the same (Winans, 2016), they have enough in common that it might explain why we feel that \textit{They will win right now} sounds approximately similar in meaning to \textit{The chef must be in the kitchen right now}.
Yet another one of Klecha’s (2014) arguments for treating will as a modal is that like modals and unlike past tense, it suppresses acquaintance inferences for predicates of personal taste.\(^{10}\)

\[\text{(18)} \quad \begin{align*}
\text{a. This movie will be great, but I haven’t seen it.} \\
\text{b. This movie must be great, but I haven’t seen it.} \\
\text{c. #This movie is great, but I haven’t seen it.}
\end{align*}\]

The non-modal claim in (18-c) is defective, unlike the modal claims in (18-a) and (18-b). A natural explanation of this is that, because of their evidential requirements, the modals modulate the acquaintance inference (Anand and Korotkova, 2018; Willer and Kennedy, forthcoming; Cariani, 2021).

One dialectical limitation of this argument is that the acquaintance inference might be modulated by will and must for reasons that do not establish that will is modal. For example, it appears possible at first sight to have a theory of the lexical meanings of these two expressions such that both must and will have similar (but ultimately distinct) evidential requirements; yet, the former is a modal and the latter is not.\(^{11}\) Whether this possibility is genuine depends on how we spell out our theory of evidentiality. According to mainstream theories, evidentiality is intrinsically connected with modality (for a survey of these connections see Murray, 2021), such that we could not ascribe evidential constraints to an expression without also ascribing to it some degree of modality. Evidently, resolving this controversy would take us too far afield.

5 The semantics of dual tense-modal operators

The implicit suggestion of a modal future hypothesis is that will and similar expressions carry both temporal and modal content. This section articulates how that might be reflected in a semantic theory. Semanticists who study will and its cognates standardly assume that will is not a primitive lexical item. Instead, it is the result of combining present tense and a tenseless modal morpheme, written out as ‘woll’ (Abusch, 1985). This morpheme is understood as the common modal element between will and would (which expression is in turn is analyzed as combining past tense with woll). As a result of this common decomposition, the semantic analyses discussed in this section are responsible for their predictions about both modals.

\(^{10}\)For more on acquaintance inferences, see Ninan (2014); Anand and Korotkova (2018); Willer and Kennedy (forthcoming).

\(^{11}\)Something like this is the position of Winans (2016).
5.1 Two styles of modal semantics

The literature offers up two fundamentally different semantic analyses for will: The universal semantic[12] and the selectional semantics[13] . The universal theory treats will as a necessity operator, endowed roughly with the semantics that Kratzer assigns to necessity operators in classic works such as [Kratzer, 1981, 1991, 2012].

In the interest of modularity and brevity, I skip over many details of domain generation in Kratzer semantics, and simply assume that we have a function that inputs a modal base $f$, an ordering source $g$ and a world of evaluation $w$ and outputs a domain. I write this as $\text{domain}(f, g, w)$. The denotation of will maps these arguments to true if every world in the domain makes the proposition true.

**Universal starting point**

\[
[w_{\text{oll}}] = \lambda A \lambda w. \forall v \in \text{domain}(f, g, w), A(w)
\]

Different versions of the universal analysis endorse different strategies for choosing evaluation parameters that are relevant for domain generation. They also differ in whether they ascribe to will an element of homogeneity, which is suggested by [Copley (2009)] as a way of capturing the fact that will seems to satisfy a kind of excluded middle principle. As [Thomason (1970)] puts it, either it will rain or it won't rain has the force of a tautology.

The selectional analysis proceeds from the idea that will does not have any quantificational force. Unembedded occurrences of will only answer for their truth to the actual world. For example, if, on Tuesday, I say Richmond FC will win Saturday’s match, the truth of what I asserted depends only on what goes in the future of the world of utterance [Huddleston and Pullum, 2002; Cariani, 2021].

From a formal point of view, this is modeled by assuming that its denotation essentially involves a selection function—similar to the selection functions that are deployed in Stalnaker's analysis of conditionals. So let sel be a function that inputs a set of worlds (the set of worlds we are selecting from) and a single world (the world from whose perspective we are selecting). Assume that sel satisfies two of Stalnaker’s classic conditions: success (unless $A$ is empty, if selecting from $A$, sel ought to select an $A$-world) and centering (if possible, when selecting from the perspective of $w$, sel ought to select $w$ itself).

**Selectional starting point**

[12] Versions of this kind of modal analysis, with different conceptions of modal bases and ordering sources, are articulated among others by Condoravdi (2002), Kaufmann (2005), Copley (2009), Giannakidou and Mari (2017), Todd (2022) attempts a bold defense of universal analyses of will from many standard objections.

[13] Selectional theories are built off an extended analogy between will and Stalnaker’s semantics. These are defended by Cariani and Santorio (2018), Cariani (2021), Kratzer (2020).
The point of [Cariani and Santorio, 2018] and [Cariani, 2021] is that the selectional starting point provides a much superior foundation to build an analysis of the future. Here, however, I do not want to wage that battle, or take that victory lap, one more time. Instead, in this survey, I simultaneously compare a few ways of building on both kinds of modal foundation, the selectional and the universal.

5.2 A crude recipe for future orientation

We have entertained two broadly modal analyses of will (and, implicitly, would). Our goal however was to design plausible semantic values that would manage both the modal component and the temporal one. Towards this goal, one might consider integrating both approaches with linear tense. (In reading the formulas below, recall that italic bold metalinguistic variables, like “p”, range over functions from world-time pairs to truth-values.)

Universal modal, existential times

\[
[woll_{f,g}] = \lambda p \lambda w \lambda t. \forall v \in \text{domain}(f, g, w), \exists t' > t, p(w, t')
\]

Selectional modal, existential times

\[
[woll_f] = \lambda p \lambda w \lambda t. \exists t' > t, p(\text{sel}(f(w), w), t)
\]

These theories exemplify how one might easily integrate a modal semantics with temporal displacement.

There are several problems with these analyses, however. Here is a minor one: the existential quantifier restricts the eligible values of \( t' \) to those times that occur after \( t \). This is problematic if we wanted to give a unified analysis of ‘ordinary’ future-directed will and present-directed will, as in (16). The issue could be sidestepped by an analysis that replaced ‘\( > \)’ with ‘\( \geq \)’. That style of analysis however incurs the intellectual debt of explaining why the distribution of present-directed will is so limited.

The real reason to doubt these analyses is that they predict scope distinctions between negation and will, and no such distinctions are observed (MacFarlane, 2014; Schoubye and Rabern, 2017; Cariani and Santorio, 2018; Cariani, 2021). While this is plausibly an argument in favor of selectionist views, sophisticated
proponents of quantificational accounts do not generally deny the scopelessness data and seek to account for it by enriching a universal analysis, e.g. with homogeneity constraints ([Copley][2009]). However, if we introduce existential quantification over times, as in ([22]) we endanger the commutativity of negation and will. Consider a future-oriented version of Partee’s example:

(23) Val will not turn the stove off.

Under ([22]) this sentence ought to have two readings, depending on whether negation scopes above will (interpretation: no future time in the selected world verifies the prejacent) or below will (interpretation: some future time in the selected world does not verify the prejacent). Only the first of these approximates the attested reading of ([23])

There is a more general way of presenting this problem. The challenge for defenders of these theories is how to square these two plausible ideas:

**Universality.** From the point of view of temporal evaluation, sentences like ([23]) have the force of (possibly restricted) universal quantifiers. In particular, the truth of ([23]) requires that the relevant time interval altogether lack an event of turning off the stove.

**Commutativity.** Intuitively, will and negation commute.

I find it pretty intuitive that ([23]) has universal truth-conditions, but the observation can be cemented by noting how these sentences involving will and an eventive prejacent interact with temporal adverbials. Note that in the non-negated form, we cannot validly infer from a claim that is restricted to an interval to a claim that is restricted to a proper subset of that interval—i.e., ([24-a]) does not entail ([24-b]).

(24) a. Val will turn off the stove tomorrow
    b. Val will turn off the stove tomorrow morning

By contrast, in the negated form these inferences are generally valid:

(25) a. Val will not turn off the stove tomorrow
    b. Val will not turn off the stove tomorrow morning

The universality of ([23]), ([25-a]) and ([25-b]) is predicted by theories that assume that will quantifies existentially over times, together with the assumption that the negation scopes over it. What’s lost, under that theory, is commutativity.

---

A similar problem applies to the quantificational analysis in ([21]). On the reading in which negation takes narrow scope under will it is predicted to mean that in every relevant world, there is a future time at which Val does not turn the stove off.
5.3 Future orientation without quantification

The solution to this problem is account for the existential behavior of will by intervening elsewhere in the clause, and not in the denotation of will. In particular, if the existential quantification is locked into a low position, it will be guaranteed to sit below negation. In this section, I present a semantic framework for the modal/temporal interaction that implements this kind of structure.

Start with the idea the temporal dimension of evaluation is an interval, not a single temporal instant. Intervals are non-empty, convex sets of instants. There are attempts to provide linguistic arguments for using intervals (e.g. Bennett and Partee, 1972). Here, however, I do not rely on these arguments and I just acknowledge that recruiting intervals provides greater expressive power and a better foundation for the kind of development I have in mind. It will be justified by overall the soundness of that development itself, rather than by any specific argument.

At the core of our semantics are sentence radicals—basic description of events before tense or aspect is applied. Let us use Val will turn off the stove as our running example of a full clause. Write the radical corresponding to this clause as Val turn off the stove. This radical denotes an event which is describable as a “turning off”, and the participants in the event are, in order, Val and the stove. Let us introduce some notation to keep track of this sort of information:

\[
[\text{Val turn off the stove}] = \lambda e. \text{turn off}(\text{Val, the stove})(e)
\]

The notation in \([26]\) is not to be viewed as a substantive bit of theory about lexical semantics. Instead, it is a relatively transparent bookkeeping device, standing in place for work that needs to be carried out in a crucial, but separate, module of linguistic theory.

Between sentence radicals and full clauses there is a whole region of intermediate layers. Aspect, modality, negation and tense all live in this region. Roughly speaking, we assume that aspect applies directly to radicals and takes the narrowest scope. Modals and negation can occur in any scope with respect to each other, but above radicals. Finally, tense sits at the top of the clause. Structurally then, we envision the clause as having the structure tense > operators > aspect > radical. (Here ‘operators’ denotes a variety of operators, such as modals, negation, and perfect.) We won’t formalize the underlying syntax at any finer level of grain than this.

---

15To my knowledge, this particular solution to the puzzle is first found in Condoravdi (2002), although it is interwoven there with many other theses which might make it hard to distill it.

16One might consider here a theory of intervals that does not make them sets of points. For examples of such theories on the logic side, see Humberstone (1979) and ?.
The only aspect to be considered is perfective aspect. We assume a minor modification \(27\)-b of a classical semantics for perfective aspect given in \(27\)-a (see Beck and von Stechow 2015, Kratzer 1998, Klein 1994).

\[
\begin{align*}
(27) & \quad \text{a. } [\text{PF}] = \lambda P \lambda I. \exists e [\tau(e) \subseteq I \land P(e)] \\
& \quad \text{b. } [\text{PF}] = \lambda P \lambda w \lambda I. \exists e \in w [\tau(e) \subseteq I \land P(e)]
\end{align*}
\]

The central semantic function of perfective aspect is to input a property of events coming from the radical, existentially quantify its event argument, and abstract away the world argument. In our running example:

\[
\begin{align*}
(28) & \quad [\text{PF}(\text{Val turn off the stove})] = \\
& \quad \lambda w \lambda I. \exists e \in w [\tau(e) \subseteq I \land \text{turn off}(\text{Val, the stove})(e)]
\end{align*}
\]

I will claim that the existential quantifier in \(28\) is the existential we were seeking to find in the logical form of sentences like \(23\). The important thing to note is that it ranges over events, not times. It is a crucial assumption of this approach that the combination of aspect and radicals forms a sentential core that negation, quantifiers and modals cannot break through.

The next ingredient in the account comes from the illuminating discussion in Condoravdi (2002), which assigns to modals two jobs: one is the standard job of quantifying over worlds; the other is extending the interval of evaluation into the future. Given an interval, \(I\), let \(I^+\) be the result of adding all the instants that are later than all the instants belonging to \(I\). (i.e., \(I^+ = \{t | \exists t' \in I, t' \leq t\}\)). We can then suppose must to have this kind of semantic value:

\[
(29) \quad [\text{must}_{f,g}] = \lambda p \lambda w \lambda I. \forall v \in \text{domain}(f, g, w), p(w, I^+)
\]

As a result of this analysis, we predict the following truth-conditions for Val must turn off the stove:

\[
(30) \quad [\text{must}_{f,g}(\text{PF}(\text{Val turn off the stove}))] = \\
\quad \lambda w \lambda I. \forall v \in \text{domain}(f, g, w), \exists e \in v [\tau(e) \subseteq I^+ \land \text{turn off}(\text{Val, the stove})(e)]
\]

\[\text{In the system of Beck and von Stechow }2015, \text{the clause for perfective is actually }27\text{-a and world abstraction is performed by a separate head named }\text{modl}.\]

\[\text{In this respect, the framework differs radically from that of Hacquard }2006\text{ in which aspect can, in some instances, move to locations that outscope some modals and negation. There are, of course, some advantages to Hacquard’s approach, but it would take us too far afield to explore all the tradeoffs involved. I will here record it as a commitment of my solution to the puzzle that the verb always needs to bind with aspect before anything else.}\]
Informally, every world in the domain of must contains an event that is a turning off of the stove by Val and is located somewhere in the future with respect to the interval of evaluation.

A complete account of the meaning of this sentence should also clarify why it so strongly suggests a deontic interpretation for the modal. Here I will leave this issue in the background while noticing that it is a critical component of what Condoravdi (2002) was out to achieve (Hacquard, 2010).

The analysis in (30) displays the general type of semantic value that the universal analysis of will casts on *Val will not turn the stove off*. As for the selectional analysis, we can adapt it to the Condoravdi framework as in (31) and derive its output in the example in (32).

\[(\text{31}) \quad [\text{woll}_f] = \lambda p \lambda w \lambda \mathcal{I}. p(\text{sel}(f(w), w), \mathcal{I}^+)\]

\[(\text{32}) \quad [\text{woll}_f(\text{p}(\text{Val turn off the stove}))] = \\
\lambda w \lambda \mathcal{I}. \exists e \in \text{sel}(f(w), w)[\tau(e) \subseteq \mathcal{I}^+ \land \text{turn off}(\text{Val, the stove})(e)]\]

To complete the anatomy of the clause, we must add tense. For simplicity, we assume an indexical present tense, according to which ‘pres’ denotes a very narrow interval consisting of the time of utterance.\footnote{There are many ways in which this indexical analysis of the present lands short of the mark. Indeed, to cover its staggering variety, Grönn and von Stechow (2016) theorize that present tense in English is ambiguous. However, my ambition here is not to land on a theoretically satisfactory characterization of present tense, but to lay down enough basics that my key points about the puzzle concerning will can be understood.} Call this interval ‘now’. One thing that present tense might do is input a temporal proposition, and output a temporal proposition whose temporal coordinate is indexically fixed to the time denoted by ‘now’.

\[(\text{33}) \quad [\text{pres}] = \lambda p \lambda w \lambda \mathcal{I}. p(w, \text{now})\]

However, because tense is assumed to be at the top of the clause, we could exploit its occurrence in that position and assume (again with Condoravdi 2002) that it saturates the temporal argument, thus converting temporal propositions (i.e., functions from worlds and intervals to truth-value) into eternalist ones (i.e., functions from worlds to truth-values).

\[(\text{34}) \quad [\text{pres}] = \lambda p \lambda w. p(w, \text{now})\]

The upshot is that, while the theory traffics in temporal propositions at the sub-clausal level, it may be possible to associate each clause with eternal propositions.
Thus Condoravdi’s idea that modals operate interval extension may be viewed as neutral between temporalist and eternalist propositions.

We have all the ingredients to state the truth-condition for \textit{Val will turn off the stove} within the quantificational \([35-a]\) and the selectional \([35-b]\) analysis:

\[
\begin{align*}
(35) & \quad \text{a. } [\text{pres } \text{woll}_f g(\text{pf}(\text{Val turn off the stove}))] = \\
& \quad \quad \quad \quad \lambda w. \forall \tau(e) \in \text{DOMAIN}(f, g, w) \exists e \in \tau(e) \subseteq \text{now}^+ \land \text{turn off}(\text{Val, the stove}(e))
\end{align*}
\]

\[
\begin{align*}
(35) & \quad \text{b. } [\text{pres } \text{woll}_f (\text{pf}(\text{Val turn off the stove}))] = \\
& \quad \quad \quad \quad \lambda w. \exists \tau(e) \in \text{sel}(f(w), w) \subseteq \text{now}^+ \land \text{turn off}(\text{Val, the stove}(e))
\end{align*}
\]

Because negation is only allowed to sit outside the sentential core—consisting of \text{pf}(\text{Val will turn off the stove})—the result of evaluating \textit{Val will not turn off the stove} has (i) universal force and (ii) the same interpretation whether negation scopes above or below \text{woll}. Here is the selectional truth-condition:

\[
\begin{align*}
(36) & \quad [\text{pres } \text{woll}_f \text{ not } (\text{pf}(\text{Val turn off the stove}))] = \\
& \quad \quad \quad [\text{pres } \text{not } \text{woll}_f (\text{pf}(\text{Val turn off the stove}))] = \\
& \quad \quad \quad \lambda w. \forall \tau(e) \in \text{sel}(f(w), w) \subseteq \text{now}^+ \rightarrow \text{not turn off}(\text{Val, the stove}(e))
\end{align*}
\]

Informally, no event in the present or future of the selected world is an event of turning off the stove whose agent is Val.

To conclude this discussion, it is worth comparing this solution to an alternative. Recall Partee’s idea that tenses might productively be analogized to pronouns, and the two ways of vindicating this analogy. Some of these ways are fully compatible with the approach I just sketched. For example, it is within the spirit of the theory that \text{woll} carries with it a variable that further restricts the interval that is relevant to its evaluation. As a result, sentences like \textit{it will snow} might demand a bit more than just the occurrence of a snowing event at some future time or other. This restriction could be imposed by a pronoun-like item, with all of the typical features of such items (e.g. the ability to generate bound readings by being anchored to a prior element of the discourse).

What could be a genuine rival to my proposal is if instead we entirely gave up on the apparatus of interval extension, and assumed that future shifting was entirely achieved by a pronoun-like device. Imagine for instance, attaching a variable \(i\) to the argument structure of \text{woll} such as the interpretation of \(i\) according to the assignment function \(g\) is the future interval of evaluation.

\[
(37) \quad [\text{woll}_f, i] = \lambda p \lambda I \lambda w. p(\text{sel}(f(w), w), g(i))
\]
From the point of view of commutativity, this analysis is just as good as \((35-\text{b})\). In fact, I doubt that there is much to choose on purely empirical grounds: with enough supplementation from context, this analysis might be bent to make the exact same predictions as an analysis based on \((35-\text{b})\). Let us then go beyond purely empirical criteria of theory choice. If as a point of systematic meaning an expression has the capacity to stretch intervals of evaluation forward into the future, it seems desirable to me to write that effect into the semantics of the expression. For that reason alone, I find the interval extension analysis preferable.

6 Fake past as modal past?

We have been elaborating on the idea that future tense has a modal component. A large literature in semantics has emphasized that there appear to also be modal interpretations of past tense. In particular, many languages express the counterfactuality of conditional antecedents with a layer of past tense that does not appear to require a temporal interpretation. To have one example to fix ideas, consider:

\[(38)\] If Ivan were to leave before lunch, he would get to the airport in time.

The point illustrated by such (ubiquitous) examples is that the past tense in the antecedent does not appear to be interpreted temporally (Iatridou 2000 refers to this as fake past). One prominent theoretical analysis of this phenomenon is that, somehow, the past tense marker gets reinterpreted as a mood marker and gets given a modal interpretation. Schulz (2014) calls this the past-as-modal account and provides an elaboration of the strategy, as does Mackay forthcoming. The main rival to this hypothesis is the view according to which the past tense in these conditionals is indeed given the standard temporal interpretation. Schulz calls this the past-as-past analysis (Arregui 2009; Ippolito 2013; Khoo 2015).

Suppose now that the past-as-modal analysis is correct. Does it at all undermine our claim that there is something special about the connection between futurity and modality? One might reason that, if this were the case, both past and future can be associated with modal interpretation. Perhaps, then, the idea of a “modal future” hypothesis should be replaced by the idea of a “modal tense” hypothesis, according to which any tense (and perhaps more generally any device of temporal displacement) can, on occasion receive a modal interpretation.

I do not think that this line of reasoning is sound. Even under the modal theory, the ways of fake past are extremely different from the ways of the modal future. To start, there are crucial distributional differences between fake past and future. Unlike future, fake past is limited to a small number of constructions. After all, any will sentence can occur in a modally subordinated context, but it’s not the case...
that any occurrence of past tense can get a modal interpretation. Furthermore, there is a deep disanalogy when it comes to their temporal interpretation: The distinctive thing about the modal future is that it it is simultaneously both modal and a device of future reference. Fake past does not exhibit this kind of duality: if the past-as-modal theory is correct, the fake past in the antecedent of *if Ada went to the store, she could meet Jana* has no temporal interpretation at all. And if the past-as-past theory is, then it does not have a modal interpretation.

In sum, there are two claims that are constitutive of a modal future hypothesis. One: devices of future reference are modals. Two that there is a fundamental linguistic asymmetry between future-tense claims and past tense claims. Both of these claims can be true even if the expressions that ordinarily function as past tense were to occasionally get modal interpretations.

### 7 Conclusion

Modality and futurity are tightly interwoven, but distilling their connection requires careful consideration of fundamental categories in the analysis of language. The observation that English *will* is not a tense, is not by itself an argument for it being a modal. Conversely, the observation that some languages have future tenses does not by itself undermine the modal future hypothesis, since it is conceptually possible for a single expression to be both modal-like and tense-like. After running through standard arguments for ascribing to *will* some element of modality, we showed this in detail by developing an analysis of *will* on which it operates at this dual level.

Bibliography


**Belnap, Nuel, Perloff, Michael and Xu, Ming** (2001), *Facing the Future*, Oxford University Press.


Cariani, Fabrizio (2021), The Modal Future, Cambridge University Press.


Chomsky, Noam (1957), Syntactic Structures, DeGruyter Mouton.

Comrie, Bernard (1976), Tense, Cambridge University Press.


Dowty, David R (1979), Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague’s PTQ, Reidel.


von Fintel, Kai and Heim, Irene (2011), ‘Notes on intensional semantics’, MIT.


Hockett, Charles D et al. (1960), *The origin of speech*, volume 1, JSTOR.


Mackay, John (forthcoming), ‘Modal interpretation of tense in subjunctive conditionals’, *Semantics and Pragmatics*.


Todd, Patrick (2022), *The Open Future*, Oxford University Press.


