We Did What We Could: An Experimental Study of Actuality Inferences in Dialogues with Modal Verbs

Lori A. Moon
University of Illinois at Urbana-Champaign
4080 Foreign Languages Building, 707 South Mathews Avenue
Urbana, IL 61801 USA
aralluna@gmail.com

Abstract

Modal verbs sometimes license actuality inferences, as in the sentence, *'Mayra knew that Jeff could see the mountains'*, which licenses the inference that Jeff did see the mountains. While advances have been made into modeling the interpretation of modal verbs with actuality inferences, far less has been done to determine, in a given discourse, which reading interlocutors are most likely to share. Previous work with corpora (Moon, 2011) indicated that, in uses of modal verbs in complement clauses, the type and tense of the matrix verb and the presence of intensifying adverbs (e.g., *so carelessly*) were sentence-level linguistic features which correlated with actuality inferences.This study runs and analyzes experiments with human subjects for statistically significant correlations among the features observed in corpora and the presence of actuality inferences. It is found that sentence level features do significantly bias subjects towards a particular reading and, furthermore, that the interaction between modal verbs and the types of matrix verb with which they occur is statistically significant.

1 Introduction

Theoretical semantic interpretations of modal verbs have posited multiple readings, one of which is selected by an agent based on the context in which the modal occurs (e.g., (Kratzer, 1981),(Veltman, 2005)) or the syntactic constituent to which the modal applies (e.g., (van der Auwera & Plungian, 1998),(Boland, 2006)). Although the choice among readings can be constrained by lexical properties of the modal verb, the readings available for a given modal verb still differ significantly in their temporal and inferential properties. For example, on one reading of sentence (1) below, Jeff did see the mountains,[1] and a realized ability is expressed.

1. *Mayra said that Jeff could see the mountains.*

On another reading, Mayra is being reported to have suggested that Jeff see the mountains and there is no felicity requirement that Jeff has seen the mountains before the time the utterance is made. Although sentence (1) is a completely natural sounding utterance, without additional context, it allows variability in temporal and inferential meaning which needs to be resolved in order for the correct semantic interpretation to proceed. Existing semantic theories can interpret each reading, but they have very limited ability to constrain or predict which reading interlocutors are most likely using in a given discourse.

In a discourse context which allows multiple readings of a modal verb, do sentence-level linguistic features facilitate interlocutors in conveying and interpreting matching modal verb readings? If so, does each relevant feature make a uniform contribution in biasing interlocutors towards a particular interpretation, or are there, rather, interactions among features making combinations of features have a greater effect than the sum of their individual contributions? Three specific research questions within these big questions are investigated in this study: In complement clauses such as in sentence (1) above, 1) Does the presence of a modal verb interact with the particular type of matrix (embedding) verb? 2) Does changing the tense of the modal verb from past

[1] The high likelihood of actuality inferences given sensory verbs has been observed by (Vendler, 1957) and others.

to present lower the likelihood of readings with actuality inferences across matrix verb type? and 3) Does information in the embedded clause making the actuality inference stronger (i.e., intensifying adverbs) increase the likelihood of a reading with actuality inferences equally across matrix verb type?

1.1 Background

Actuality inferences are important for natural language processing because they attribute non-modal meaning to modal verb uses.[2] For human interlocutors, an actuality inference updates the common ground of the discourse with (defeasible) information about the actual world, as opposed to the hypothetical or purely epistemic information introduced by other readings of modal verbs. Actuality inferences also carry temporal inferences with them: If an agent attributes an actuality inference to a modal verb use, it follows that the time of the described event precedes the utterance time.[3]

The problem of interpreting modal verbs in automated tasks has proven to be difficult. State of the art systems such as MacCartney 2006 and MacCartney et al. (2009:57) use lexical association with one inference pattern. The FraCas test set (Cooper et al., 1996), on which more linguistically rich entailment models are tested, only contains two uses of *would* and none with *could*. When textual inference tasks are discussed, modal interpretation is often either being overlooked or making inadequate predictions.

Uses of modal verbs with actuality inferences are particularly useful for automated inferencing systems to detect due to the information they contribute to the actual world of the discourse as well as the significantly lessened computational modeling required. The independent variables tested with subjects in this study were chosen with consideration for their tractability as encodable features for automated tasks.

[2](Hacquard, 2006), unlike (Bhatt, 1999), claims that modal verbs with actuality entailments in French and Italian do need to be modeled with possible worlds. Although her reasons could be argued to apply to the English data, they are somewhat dependent on her particular theory of event and modal scope interaction.

[3]Some accounts such as (Marneffe, Manning and Potts, 2012) treat actuality inferences as including future uses, provided there is speaker certainty (e.g., paraphrasability with *will*.)

1.2 Preliminary Work

In a previous study (Moon, 2011), 375 instances of *could* and *would* from an 80,000 line corpus of weblogs were hand-classified according to whether or not their most prominent reading was one with an actuality inference or one without an actuality inference. Within the actual or non-actual uses, examples were classified according to coarse-grained taxonomic categories such as *counterfactual, hypothetical, dispositional*, and *habitual in the past*, following various modal taxonomies (e.g., (Portner, 2009), (Ziegeler, 2000), (Celce-Murcia & Larsen-Freeman, 1999)).

Examples were sorted into syntactic categories in order to allow the classification problem to be potentially described by construction-based rules (cf Fernandez et al. 2007). The most prominent syntactic categories which emerged were: **Conditional (Cond):** instances in the consequent clause of a conditional sentence. **Free-Standing (FS):** instances in a main clause, possibly with adjuncts or coordination. **Relative Clause (RC):** instances in a clause headed by a noun phrase. **Complement Clause (CC):** instances in a clause which is the complement of a verb. **Other (OT):** instances in adjunct or coordinating constructions which seemed to exhibit exceptional properties.

For sentences with the modal verb *would*, the syntactic construction was found to be a strong factor in determining whether a given use of the modal was likely to have an actuality inference or not. For *could*, however, the syntactic construction was less helpful. As shown in figure 1, many constructions were almost equally likely to encode actual or non-actual uses.

Additional feature exploration was done in the corpus data within construction type in order to determine which linguistic expressions correlated with actuality inferences. For complement clauses with *could*, certain properties of the matrix verb type, matrix verb tense, and intensifying adverb phrases appeared to correlate with actuality entailments.

In this study, the generalizations from corpus data mentioned in (Moon, 2011) are treated as independent variables in an experimental study with human subjects. It is proposed that, if the same features which emerged in corpora are varied in controlled linguistic contexts with statistically significant correlations to the presence of ac-

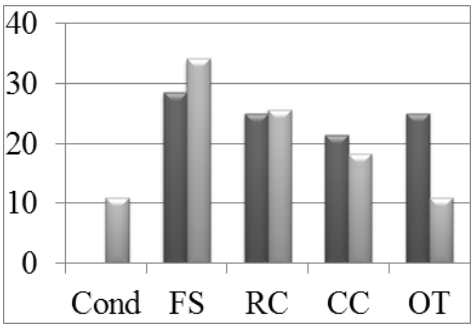

Figure 1: Given an actual occurrence of the modal *could* (black) or a non-actual occurrence (gray), this chart shows the likelihood that a particular construction will be used (figure from Moon 2011). Cond = Conditional, FS = Main Clause, RC = Relative Clause, CC = Complement Clause, OT = Other.

tuality inferences, their validity as features determining the modal verb reading is strongly corroborated.

2 Research Question 1: Matrix Verb Type and Modal Verbs

It was observed in corpus data (Moon, 2011) that the matrix verb type appeared to be a strong determining factor in actuality entailments. Although Levin verb classes (Levin, 1993) were first considered, it was the three general classes defined in Karttunen's work (Karttunen, 1973), (Karttunen, 1974) which emerged as forming separate categories with respect to actuality inferences.

Karttunen proposed that three types of matrix verbs respond differently regarding whether or not they allow presuppositions in their scope to project. Factive verbs, and most other verbs, are considered to be *holes* to presuppositions, allowing presuppositions in their scope to percolate. Verbs of saying and telling are considered to be *plugs*, relativizing any presuppositions in their scope to the beliefs of the reported speaker. Propositional attitude verbs, such as *believe*, are claimed to alternatively allow or disallow the percolation of presuppositions in the clauses in their scope. Karttunen's work on presupposition projection remains very influential in both implementations (e.g.,(Clausen & Manning, 2009)) and experimental work (e.g., (Smith & Hall, 2011)).

Actuality inferences would be difficult to categorize as presuppositions. In English, they are pervasively cancelable: Allowing continuations which re-construe the modal reading as one without an actuality inference as in sentence (2):

2. *Mayra knew that Jeff could see the crime taking place,* {*but he was actually sleeping at the time/if he wanted to*}.

To speak of actuality inferences projecting through embedded clauses would also not be accurate, since the unembedded modal clause in itself rarely has an actuality inference. Therefore, Karttunen's theory is not straightforwardly tested in this study, rather, it provides a classification of embedding verb types which has proven helpful for presupposition projection and also correlates with presence or absence of actuality inferences.[4]

By alternating the matrix verb class with the presence or absence of a modal verb, we test whether the verb classes behave as holes, plugs, or filters regarding the assertion in their scope in cases such as sentence (3) where the modal is omitted.

3. *Mayra* {*knew/thought/said*} *that Jeff saw the crime taking place.*

When the modal is included, as in sentence (2) above, there are two possible outcomes: The first hypothesis is that the modal verb will make a uniform contribution across matrix verb type in lowering the likelihood of a reading with actuality inferences. In this case, the contribution of each feature is uniform and can be composed to represent the overall likelihood of a particular reading. An alternative hypothesis is that the presence of a modal verb will have a greater effect with some matrix verb types than with others. If the second hypothesis holds, then it follows that the contribution of the modal verb and the contribution of matrix verb type in determining a modal verb reading cannot be described by any simple combination of the contribution of each part. Rather, the semantic theory must represent the fact that the interaction of modality and matrix verb type is greater than the combination of its parts.

[4] One application of this observation is that, in automated systems such as that of (Clausen & Manning, 2009), the regular expressions used to locate factive verbs for representing presupposition projection can be used to detect embedding verbs with modal verbs in their complement clause and encode predictions regarding actuality inferences.

3 Research Question 2: Past under Past Embeddings versus Past under Present Embeddings

There are various phenomena which are said to exhibit *sequence of tense* (sot) effects (Abusch, 1997), (Ogihara, 1995). This study examines one specific type of sot phenomena: Past under past embedded clauses in which the central past tense modal verb, *could* (Quirk et al., 1985) (Crouch, 1993) occurs under a past tense embedding verb.

The modal verb *could* allows a *simultaneous reading*: A reading in which the reported event occurred at the time of the subject of the main clause's reported speech act:

4. *Yvette said that Jill could jump on the trampoline for twenty minutes.*

In sentence (4), there is a simultaneous reading on which Jill has the ability to jump on the trampoline for twenty minutes, also called the *generic* reading (Bhatt 1999:173). At the time of Yvette's reported speech act, Jill had the ability to jump on the trampoline for twenty minutes. It is even possible for the (defeasible) inference to be made that Jill has, at the time of Yvette's saying that she could jump on the trampoline for twenty minutes, already jumped on the trampoline for twenty minutes in the past at least once.

Sentence (4) also involves a *sequential reading*: A reading in which the potential event of Jill jumping on the trampoline for twenty minutes occurred before Yvette's act of reporting Jill's ability. This is sometimes called an *episodic* reading (Bhatt 1999:173).

Cross-linguistically, some languages which have morphological aspectual marking on their modal verbs distinguish generic from episodic readings in that only the latter can occur with the perfect aspect (Bhatt, 1999). In English, however, both the episodic and generic readings of sentence (4) license actuality inferences, provided that, on the episodic reading, the utterance time is taken to be after the time of the potential event in the scope of the modal (i.e., Jill's jumping on the trampoline for twenty minutes). Such readings might be described as being *metaphysical* or *historical* (Condoravdi 2002:77-84)[5] in that they describe a past prediction about an event which was future at the time of reported speech, but past relative to the time of the utterance. The potential event in the scope of the modal might or might not have occurred at the speech time. On the reading in which Jill did jump on the trampoline for twenty minutes, the prediction of the reported speaker held and, on the reading in which Jill did not jump on the trampoline, the prediction of the reported speaker has proven to be incorrect. Only the former reading licenses an actuality entailment.

It is hypothesized that, when subjects accept an actuality inference in past under past embeddings, they are accepting a historical (or possibly generic) reading. In accepting a historical reading, they are rejecting the reading in which the past under past sentences describes a prediction about an event which follows the speech time.

In past under present embeddings, the three verb classes are still used, but with the embedding verb is in the present tense rather than in the past tense. As discussed in section 4.1, past under past embeddings allow various sot phenomena, each of which locates the embedded (telic) event as preceding the speech time. It is hypothesized that, by using a present tense embedding verb, such as sentence (5), the modal will be interpreted as referring to a future prediction rather than a comment on a past event.

5. *Rika understands that Jordan could report her to the authorities.*

The modal in sentence (5) most prominently describes a hypothetical future possibility rather than Jordan's past ability.

Given the readings available, it is hypothesized that readings with actuality inferences do not occur in past under present embeddings. This effect is hypothesized to be uniform across matrix verb type.

4 Research Question 3: Intensifying Adverb Phrases

One feature which emerged and which has not received a lot of attention with respect to actuality entailments were intensifying adverbs such as *so emphatically* or *so carelessly*. Such expressions were used in the corpus data most frequently with actual readings. In general, intensifying adverb constructions modify an actual event. However, non-actual readings with *so + ADVERB* are licit

[5](Condoravdi, 2002) does not include *could* among her modals for the past, however, it seems that this reading of *could* is included in her description of metaphysical uses of non-root modals (2002:78).

as in sentence (6):

6. *Judson believes that Imri could so quickly forget his parents.*

Sentence (6) can be uttered, for example, as a speculation about Imri's anticipated mental state during his future absence. The reading without an actuality inference, without a background context, does not, however, seem to be the most prominent one.

The presence of intensifying adverb phrases is hypothesized to increase the number of subjects accepting readings with actuality inferences across matrix verb type.

5 Methods

It was not feasible to test every possible value of the proposed features (3 x 2 x 2 x 2 = 24 conditions), so only select combinations were tested, resulting in twelve test conditions:

(ia) factive verb, past, no modal: Marian was amazed that Dewayne silenced the witness.
(ib) factive verb, past, modal: Marian was amazed that Dewayne could silence the witness.
(ic) factive verb, past, modal, so-phrase: Marian was amazed the Dewayne could so effectively silence the witness.
(id) factive verb, present, modal: Marian is amazed that Dewayne could silence the witness
(iia) saying verb, past, no modal: Marian said that Dewayne silenced the witness.
(iib) saying verb, past, modal: Marian said that Dewayne could silence the witness.
(iic) saying verb, past, modal, so-phrase: Marian said that Dewayne could so effectively silence the witness.
(iid) saying verb, present, modal: Marian says that Dewyane could silence the witness.
(iiia) propositional attitude verb, past, no modal: Marian believed that Dewayne silenced the witness.
(iiib) propositional attitude verb, past, modal: Marian believed that Dewayne could silence the witness.
(iiic) propositional attitude verb, past, modal, so-phrase: Marian believed that Dewayne could so effectively silence the witness.
(iiid) propositional attitude verb, present, modal: Marian believes that Dewayne could so effectively silence the witness.

5.1 Subjects

A total of 41 subjects were tested, with five results discarded due to acquisition of English as their primary language later than three years of age. Subjects were recruited through two introductory linguistics courses as well as through the author's personal contacts. All subjects were undergraduate students at the University of Illinois at Urbana-Champaign except one individual living in another US city. Twenty-six of the subjects were female and ten male. The age range was from 18-41 years old (average 21). All but one subject reported their English proficiency as level '5' on a scale from 1-5 where 1 was 'beginner' and 5 'advanced' (one reported 4). Two bi-lingual speakers considered themselves to be more proficient in a language other than English (Spanish and German). Their results were included since they claimed to use English as their primary language prior to age three and self-reported their proficiency as 'advanced'. Other languages spoken by subjects but considered less primary than English included Danish, French, Hindi, Japanese, Korean, Mandarin Chinese, Taiwanese, and others. Only four subjects were not born in an English-speaking country but moved to one by age three.

5.2 Apparatus

Experimental results were collected via online surveys through SurveyGizmo Student Edition. Subjects completed the survey in two parts which could be taken at separate times at the subjects' convenience. Thirty-six subjects were tested via online surveys (12 scripts, 3 subjects each). Subjects were presented with a scenario containing the target sentence and asked whether the predicate embedded under the scope of the modal had already occurred at the time when the dialogue was taking place. Subjects answered 'yes' or 'no' and then rated their certainty on a five-point scale.

Finding a context in which subjects can potentially read the modal as having an actuality inference and potentially read it as being predictive is difficult. Each dialogue involved a scenario in which two agents, Ann and Jacob, took discourse turns. Ann's role was to present a topic which Jacob then commented on either requesting more information or affirming what was said. The target data were always presented preceded by the discourse particle, *well*. An example is in figure (2).

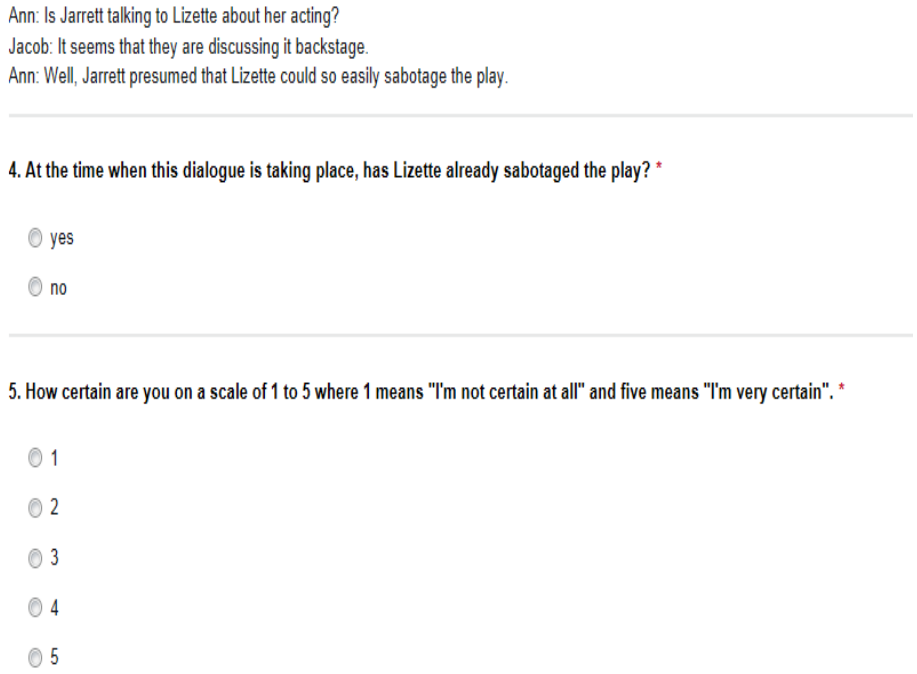

Figure 2: Example of SurveyGizmo interface

In each of the 36 scenarios (three examples of each of the twelve conditions tested) in which the independent variables were manipulated, there was, at a minimum, a historical and a predictive reading available.[6]

On the historical reading of the scenario in figure (2), Ann is reporting Jarrett's presumption after the time during which Lizette could have sabotaged the play. The discourse particle is read as introducing an affirmation that Jarrett's presumption has some relation to the present unspecified state of affairs. On a predictive reading, the discourse particle is read as providing additional information explaining the situation discussed in the first and second turns.

By the use of *already* in the question, the actuality inference is suggested to the subject. However, for it to be possible, the subject must read the modal report as historical, not predictive. If the subject can only get a predictive reading, then she will not respond that there is an actuality inference. This bias was the same across conditions.[7]

6 Results

Results were run for 'yes' or 'no' responses only (with degree of certainty measured separately), as shown in figures 3 and 4 . A sample sentence is given for each of the twelve conditions and indexed in the results table in figure 3 in section 5 above.

Figure 4 shows the basic trajectory of subjects' responses by showing the percentage of 'YES' responses to each condition.

[6]The complete set of scenarios used in testing is available at `https://netfiles.uiuc.edu/lcoulter/ExperimentalScripts`.

[7]When subjects in a pilot study were presented with sentences without a discourse context, particular problems were encountered with future readings. The pilot used two other conditions: past temporal adverbs and future temporal adverbs. It was predicted that future temporal adverbs, as in *Mary said that Jarrett could cook the fish tomorrow*, subjects would overwhelmingly report no actuality inference. However, it was found that a significant number of subjects accepted an actuality inference. It was surmised that subjects attributed an additional level of evidentiality to the data, assuming that whoever was reporting the data was reporting it after the quoted *tomorrow* had passed. Perhaps some notion of Grice's relevance maxim (Grice, 1975) can then explain the assumption subjects made that there would be no point in reporting an overtly future prediction unless it had come to pass. The addition of a brief dialogue as a control condition helped ensure that subjects evaluated the possibility of an actuality inference at the utterance time.

	(a) past, -modal		(b) past, +modal		(c) past, +modal, so		(d) present, +modal	
	YES	NO	YES	NO	YES	NO	YES	NO
(i) Factive	**100%**	**0%**	**61.11%**	**39.81%**	**75.92%**	**24.07%**	**49.07%**	**50.92%**
certainty (stdev)	4.69 (0.55)	N/A	3.93 (1.09)	3.69 (0.61)	4.22 (0.79)	4.18 (0.46)	4.13 (0.81)	4.18 (0.46)
(ii) Say/Tell	**97.22%**	**2.77%**	**18.51%**	**81.48%**	**47.22%**	**52.77%**	**13.88%**	**86.11%**
certainty (stdev)	4.51 (0.75)	1.33(n/a)	3.60 (1.38)	3.73 (1.05)	3.71 (0.96)	3.42 (1.32)	3.88 (0.8)	3.75 (0.92)
(iii) Propositonal	**54.62%**	**45.37%**	**18.51%**	**81.48%**	**31.48%**	**68.51%**	**2.77%**	**97.22%**
certainty (stdev)	3.97 (1.06)	3.4 (1.04)	3.58 (0.87)	3.85 (1.1)	3.69 (1.02)	3.57 (1.16)	3 (2.82)	4.20 (0.90)

Figure 3: Results of study of complement clauses with *could*. Alphanumeric labels correspond to sample conditions

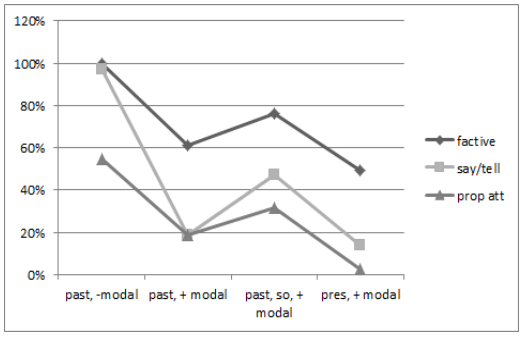

Figure 4: Plot of percentage of subjects accepting actuality inferences by condition and verb type. Y axis = percentage of subjects giving a 'yes' response (e.g., 100% means all subjects responded with 'yes' to all three samples of the condition).

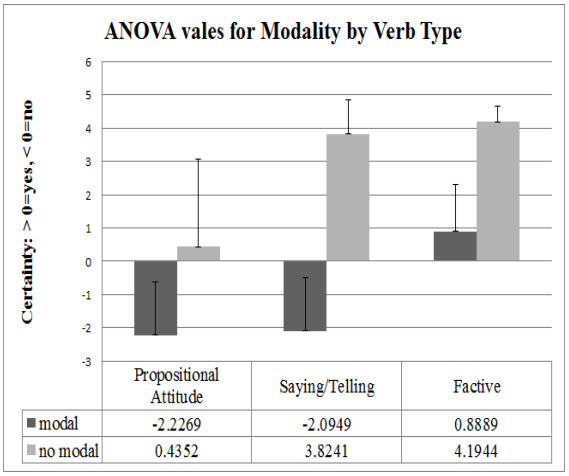

Figure 5: Effects of the modal verb *could* on actuality inferences by verb class.

6.1 Matrix Verb Type and Modal Verbs

Regarding the research question of how the matrix verb interacts with the presence or absence of a modal verb, it was found that assertions project through verbs of saying or telling on par with holes such as factives.

Column (a) in figure 3 is taken to illustrate the isolated effects of past under past embeddings in which no modal is present. Sot are known to occur in such contexts, which allow a sequential or a simultaneous past reading. Without the modal present, predictive readings would not be possible. The internal argument of the embedded verb referred to a telic event in all the data, so habitual readings were less likely.

The results are that factive verbs (YES: 100%) and verbs of saying and telling (YES: 97.2%) result in subjects almost always reporting actuality inferences, but propositional attitude verbs show a near even split among interpretations (YES: 54.62%). These results are somewhat surprising compared to those predicted in (Karttunen, 1973), (Karttunen, 1974) for presupposition projection. Without the presence of a modal verb, an actuality entailment exists in the embedded clause, however, subjects do not treat information in the scope of verbs of saying and telling as less actual. Whether human subject judgments on presuppositions are the same in this context is a topic for future experiments. As will be seen in the results of column (b), introducing uncertainty in the embedded clause drastically changes subjects' responses.

For column (b), past under past embeddings with *could*, the percentage of subjects making an actuality inference greatly decreases for all embedding verb types, but much more sharply for verbs of saying and telling (NO: 81.48%), which pattern identically to propositional attitude verbs (NO: 81.48%). Factive verbs remain majority actual (YES: 61.11%).

Column (b) gives evidence for the role of evidentiality in interpretation: When another person's views are words are being reported, the presence of the modal makes the report epistemically

uncertain. On epistemically uncertainty readings of *could*, it is the subject of the main clause who is unsure, but it can often be inferred that the person reporting the speaker's view is epistemically uncertain. The effects of evidentiality can be seen if co-indexing occurs as in *'I said that I could cook the fish'* which can be easily read with the actuality inference that the speaker actually did as she said.

For ANOVA analysis, the binary 'yes' or 'no' responses and five-point certainty scale were converted to an even ten-point continuum of 0.5-interval values from -4.5 to 4.5 ('very certain' 'no' and 'very certain' 'yes', respectively).[8],[9]

A 2x3 ANOVA was run crossing modality (2 levels: Presence or absence of the modal) and verb class (3 levels: Factive, saying or telling, and propositional attitude), as show in categories (i-iii) a-b of figure 3.

This measure was run to determine the effect of the presence of the modal on the interpretation of the three verb classes as shown in figure 5.

The verb type was found to be statistically significant ($F(2, 70) = 119.6; p < 0.001$) with factive verbs raising the likelihood of assertions projecting and propositional attitude verbs lowering the likelihood. The presence of a modal verb was also statistically significant ($F(1, 35) = 319.07; p < 0.001$) in lowering the likelihood of actuality inferences across matrix verb type. The interaction between verb type and the presence of the modal verb was also statistically significant ($F(2, 70) = 53.55; p < 0.001$), with verbs of saying or telling being affected most strongly by the presence of a modal verb greatly lowering their likelihood of being interpreted as having an actuality inference.

6.2 Past versus Present Tense Matrix Verb

Column (d) in figure 3 sought to see to what extent the tense of the embedding verb interacted with actuality inferences in complement clauses. It was found that, under present embeddings, about half of the subjects still accepted the actuality inference under factive verbs (NO: 50.9%). However, with verbs of saying or telling (NO: 86.11%) and propositional attitude verbs (NO: 97.22%) the percentage of subjects rejecting actuality inferences increased to nearly absolute agreement.

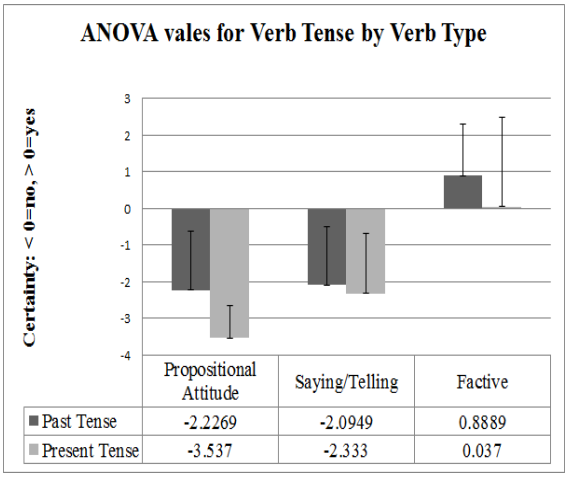

Figure 6: Interaction of tense with other features.

Column (d), along with column (b), show that the combination of verb tense and embedding verb type are the strongest features determining whether or not a modal verb licenses a prominent reading in which an actuality inference holds.

A 2x3 ANOVA was run crossing tense of the matrix verb (2 levels: Past and present) and matrix verb type (3 levels: Factive, saying or telling, and propositional attitude) as in (i-iii), (b and d) of figure 3. The measure did not show a statistically significant interaction between verb type and tense (6), but both tense and verb type were found to be statistically significant in themselves ($F(1, 35) = 11.68; p < 0.003, F(2, 70) = 76.08; p < 0.001$, respectively) in that present tense lowers the likelihood of a reading with an actuality inference.

6.3 Intensifying Adverbs

Column (c) in figure 3 adds an intensifying adverb phrase to the past under past modal embedding to see if sentence-level features in the embedded clause can coerce subjects to make actuality inferences. If so, then it shows more clearly which embedding verbs strongly prevent actuality inferences.

It was found that all embedding verb types led more subjects to accept actuality inferences, but only factives (YES: 75.92%) were above the 50% threshold. Verbs of saying were very close to an

[8]The resulting ten point scale placed the origin at the midpoint between an answer of 'yes' with certainty of one and an answer of 'no' with certainty of one.

[9]See (McKinstry, Dale, and Spivey, 2008) for experimental evidence that the divide between 'yes' and 'no' responses is more of a continuum than binary.

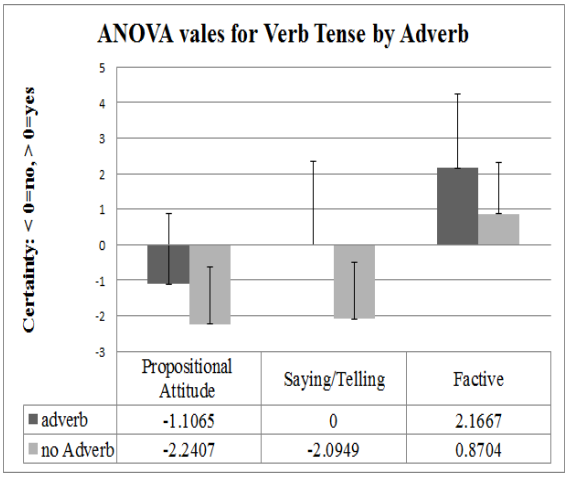

Figure 7: Interaction of 'so + adverb' phrases with other features.

even split (NO: 52.7%), with propositional attitude verbs remaining least actual (NO: 68.51%).

A 2x3 ANOVA of intensifying adverb phrases (2 levels: Presence or absence of the adverb) given verb type (3 levels: Factive, saying or telling, and propositional attitude) was also run. This measure showed the adverb to be statistically significant ($F(1, 35) = 31.05; p < 0.001$) in raising the likelihood of a reading with an actuality inference when it was present. Verb type ($F(2, 70) = 94.68; p < 0.001$), again, was statistically significant, but the interaction of the two was not statistically significant.

7 Discussion

Although English does not have overt morphological markings helping distinguish uses of modal verbs with actuality inferences from uses without, there are a number of sentence level linguistic features which assist in limiting the most likely reading. Furthermore, when embedded in a discourse context allowing both historical and predictive readings, subjects systematically interpret the discourse markers and context in a way that facilitates the most likely reading of the modal based on sentence level features.

Matrix verb type was the most prominent feature in determining actuality inferences. This fact cannot be straightforwardly explained by the theory from which the verb classes were chosen (Karttunen, 1973), (Karttunen, 1974), but it rather shows that levels of evidentiality and intuitions about the limitations of epistemic knowledge influence modal verb readings.

The second most prominent determining feature was the past tense of the main verb. Sot phenomena are predicted to occur in past under past embeddings, making the possibility of actuality inferences clear. However, the uniform trajectory across verb type, as shown in figure 4, indicates that embedding verbs play a much stronger role than tense in determining the reading.

In response to the big research questions, it was found that sentence level features do indeed help determine which reading interlocutors are most likely to attribute to a given use of a modal in that the features tested: Matrix verb type, matrix verb tense, and intensifying adverbs all had a statistically significant effect on the modal verb reading.

It was found that some features have a uniform contribution to what reading interlocutors are most likely to choose. When the matrix verb is in the present tense, all matrix verb types have a lower likelihood of being interpreted as having actuality inference. Similarly, intensifying adverbs have an effect across verb type of increasing the likelihood of a reading with actuality inference.

Other features, such as the interaction between modal verbs and the matrix verb type can not be represented as a simple combination of the matrix verb type combined with the contribution of the modal. Rather, verbs of saying or telling are much more affected by the presence of a modal than factive or propositional attitude verbs. These results present particular challenges for any theory attempting to represent a probabilistic bias towards one reading in terms of uniform contributions of constituents in a compositional semantic approach.

The results of our study have an impact on theoretical models of the role of sentence level features on modal verb interpretation in discourse. They also have applications for automated models of modal verb interpretation, providing tractable features correlating with a strong likelihood of modal verb readings with actuality inferences.

In future work, we will test features correlating with actuality inferences in other syntactic constructions and with other past central modal verbs. We will also work to improve our discourse scenarios in order to measure when human subjects make actuality inferences rather than measuring whether they accept or reject them when they are suggested.

Acknowledgments

The author's committee was helpful in supervising this research: Professors Tania Ionin, Shalom Lappin, Peter Lasersohn, and José Meseguer. Professor Michael Spivey's workshop at the 6th annual Experimental Methods in Computational Linguistics boot-camp was indispensable in learning to run ANOVA on the data. The audience at NASSLLI 2012 who attended the related talk at Reasoning and Interaction (RAIN) gave excellent feedback as well as Professor Elizabeth Allyn Smith, and members of the University of Illinois at Urbana-Champaign Experimental Pragmatics Group. Finally, warm thanks go to the anonymous reviewers for helpful suggestions and encouragement on this project.

References

Dorit Abusch. 1997. Sequence of Tense and Temporal de Re. *Linguistics and Philosophy*, 20(1): 1-50.

Rajesh Bhatt. 1999. *Covert modality in non-finite contexts*. PhD thesis, University of Pennsylvania.

Annerieke Boland. 2006. *Aspect, Tense and Modality: Theory, Typology, Acquisition, Volume I*. Center for Language and Communication.

Marianne Celce-Murcia and Diane Larsen-Freeman. 1999. *The Grammar Book: An ESL/EFL Teacher's Course.* Heinle and Heinle Publishing Company. Boston, MA.

David Clausen and Christopher D. Manning. 2009. Presupposed content and entailments in natural language inference. *Proceedings of the 2009 Workshop on Applied Texual Inference, ACL-IJCNLP 2009.* pages 70-73.

Cleo Condoravdi. 2002. Temporal Interpretation of Modals: Modals for the Present and for the Past. Chapter in *The Construction of Meaning*, pages 59-88. CSLI Publications.

Robin Cooper, Dick Crouch, Jan van Eijck, Chris Fox, Josef van Genabith, Jan Jaspars, Hans Kamp, David Milward, Manfred Pinkal, Massimo Poesio, and Steve Pulman. 1996. *Using the Framework*. The FraCaS Consortium, January 1996.

Richard Crouch. 1993. The temporal properties of English conditionals and modals. PhD thesis, University of Cambridge.

Raquel Fernandex, Jonathan Ginzburg, and Shalom Lappin. 2007. Classifying non-sentential utterances in dialogue: A machine learning approach. *Computational Linguistics*, 33(3):397-427.

H. P. Grice. 1975. Logic and conversation. *Syntax and Semantics, 3: Speech Acts*, ed. by P. Cole & J. Morgan. New York: Academic Press.

Valentine Hacquard. 2006. Aspects of modality. PhD thesis, Massachusetts Institute of Technology.

Lauri Karttunen. Spring 1973. Presuppositions of compound sentences. *Linguistic Inquiry*, 4(2):169-193.

Lauri Karttunen. 1974. Presupposition and linguistic context. *Theoretical Linguistics*, 1:181-194.

Angelika Kratzer. 1981. The notional category of modality. Chapter in *Words, Worlds and Contexts*.

Beth Levin. 1993. *English Verb Classes and Alternations*. University of Chicago Press.

Bill MacCartney. June 2009. Natural language inference. PhD thesis, Stanford University.

Bill MacCartney, Trond Grenager, Marie-Catherine de Marneffe, Daniel Cer, and Christopher D. Manning. 2006. Learning to recognize features of valid texual entailments. *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*, pages 41-48.

Marie-Catherine de Marneffe, Christopher D. Manning, and Christopher Potts 2012 Did it happen? The pragmatic complexity of veridicality assessment. *Computational Linguistics*, 38(2):301-333.

Chris McKinstry, Rick Dale, and Michael J. Spivey 2008 Action Dynamics Reveal Parallel Competition in Decision Making. *Psychological Science*, 19.1:22-24.

Lori Moon. 2011. Discriminating among 'actual' and 'non-actual' uses of *could* and *would* for improved text interpretation. *Building Representations of Common Ground with Intelligent Agents: Technical Report from the AAAI Fall Symposium*.

Toshiyuki Ogihara. 1995. Double-access sentences and references to states. *Natural Language Semantics*, 3:177-210.

Paul Portner. 2009. *Modality*. Oxford University Press.

Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartivik. 1985 *A Comprehensive Grammar of the English Language*. Longman.

Elizabeth Allyn Smith and Kathleen Currie Hall. 2011 Projection Diversity: Experimental Evidence. Proceedings of the Workshop on Projective Meaning, ESSLLI 2011.

Johan van der Auwera and Vladimir Plungian. 1998 Modality's Semantic Map. *Linguistic Typology*, 2:79-124.

Frank Veltman. 2005 Making Counterfactual Assumptions. *Journal of Semantics*, 22.2:159-180.

Zeno Vendler. 1957. Verbs and times. *The Philosophical Review*, 66(2):143-160.

Debra Ziegeler. 2000. The role of quantity implicatures in the grammaticalisation of *would*. *Language Sciences*, 22:27-61.