

Children's Understanding of Modal Verbs And its Relation to Theory of Mind*

Avril Kenney

MIT

May 17, 2008

Introduction

In this project, we investigate how young children, who are in the process of acquiring language, interpret the meanings of modal verbs in English. We are particularly interested in the word “will”, because of its use in Theory of Mind tests. Modal verbs (words such as “must”, “may”, “can”, etc.) can have two possible interpretations, and which of the two is chosen depends on the semantic and pragmatic context of its usage. In the deontic interpretation, the modal is expressing a definite fact, whereas in the epistemic interpretation, it expresses a fact as understood by the speaker. For example, the sentence “he must eat” could be read either deontically as “it is necessary for him to eat”, or epistemically as “from the available evidence, it is a necessary conclusion that he eats”.

Previous studies have established that children begin using the epistemic meanings significantly later than the deontic, but a reason has not yet been determined for this phenomenon. Some researchers (e.g., Papafragou 1998) suggest that young children cannot get the epistemic reading because they lack the cognitive capacity to represent alternative versions of the world as real, so the notion of possibility is simply not comprehensible. What we hypothesize is that children's inability to interpret epistemic modals is due to their not having acquired verb raising, the only syntax compatible with that modality.

If true, this would have significant implications for how the presence of a Theory of Mind is determined. A common Theory of Mind test is the false-belief test, in which the subject sees a situation along with an observer, then the situation is changed in the observer's absence, and the child has to predict what the observer's behavior will be. The idea is that if a child has a Theory

* This project was guided and supervised by Nadya Modyanova. Most of the ideas presented in this paper are hers.

of Mind, he will comprehend that the observer is unaware of the change, and so he will say that the observer will behave based on what the situation was initially, even though he himself knows that it has changed and therefore that such behavior is incorrect. Previous findings have shown that children below the age of about three or four years fail these tests; however, the subject's predictions are often elicited by asking what the observer "will" do, and if the word "will" is being understood in the deontic sense, children will respond with what they think the observer *should* do, rather than what the person is *going* to do, simply because they are misinterpreting the question.

The goal of the test that we made is to assess understanding of the words by portraying different situations in which they are used and asking the subjects about their expectations. Using cardboard cutouts and scenery drawn on paper, we filmed short videos of interactions between cartoon characters in which one of them makes a statement using a modal verb. After this statement, the video is paused and the subject is asked to predict what happens next, a judgment he will (presumably) make based on what the characters have said. He is then shown the remainder of the scene, in which his prediction is either fulfilled or unfulfilled, and he is asked whether he is surprised and whether the character lied in making the statement. We hypothesize that in epistemic usages, children will predict fulfillment of the modal statement even when it is not supported by context (as if it were deontic), and in the unfulfilled cases they will be more surprised and more likely to say that the character lied than adults.

In conjunction with the modals experiment, we are also running a Theory of Mind test that uses a minimal amount of language and no modal verbs. In this test, an object is placed under one of two bowls, and an observer indicates (by touching or pointing, without words) which bowl he thinks contains the object, after which the subject has to indicate his own opinion of where the object is. Different trials vary such factors as whether the subject sees where the

object is, whether the observer sees, whether the observer leaves, whether the object or the bowls are moved, and when the observer is asked to indicate where he thinks the object is. This test will show whether the children tested can use the information they have observed to conclude that the observer's indication might not be correct. The ability to reason in this way should be indicative of having Theory of Mind.

If our idea about modals is correct, then what we expect to find is that children's performance on the nonverbal Theory of Mind test, measured as tendency to indicate the same bowl as the observer indicates despite having seen the bowls switched, should not necessarily correlate with their performance on the modals test, measured as tendency to be surprised at unfulfilled modal statements regardless of whether the statements are supported by context. We would also expect performance on both of the tests to be correlated with age; children below the age of three or four years old should fail the modals test, and children below some younger age threshold should fail the Theory of Mind test.

Background on Theory of Mind and its Development

Theory of Mind is a topic of considerable interest in the field of psychology, for it is one of the features that have enabled humans to develop to such sophisticated levels, yet the concept itself is defined quite vaguely. The general idea is that a person has Theory of Mind if he is able to represent the mental states of others and understand that they are different from his own. In some cases, it seems easy to see the contrast between its presence and absence on a high level: if a person claims that everyone else knows what he knows, then he would be considered not to have Theory of Mind. But at the same time, other behaviors suggest that even infants have some understanding that their own minds are separate from others'. Confusion in determining the presence of Theory of Mind arises from the lack of a clear definition and the confounding with other cognitive abilities in tests devised for young children.

There are many different ways of testing for Theory of Mind, but the tests must necessarily get simpler as the subjects get younger, eventually reaching a point where it is not at all clear that what is being tested is the same as what was originally being searched for. With older subjects who have specifically social cognitive deficits but otherwise normal intelligence, such as people with autism, the normal false-belief tests can be used, directly asking them what they think. What is tested in these is as close as possible to the common conception of Theory of Mind. With younger children, it would seem that language can still be used because they are capable of responding to it, but (as our study aims to find out) their comprehension might not be coinciding perfectly with the speaker's. And with infants too young to talk, it is impossible to tell what processes are occurring in their minds to cause their behavior, so the proxies such as preferential looking that are used to determine interest are not necessarily measuring what they are intended to, not to mention that it is difficult to get very accurate measurements in babies.

Although it would be convenient to have a single test that could categorize a person as either having Theory of Mind or not, it is unlikely that that can happen, because the more research that is done on the subject, the more it appears that Theory of Mind is not really a distinct mental capacity but rather a combination of abilities that express themselves at different places along the developmental continuum. As they grow, children pass various milestones that each in their own way indicate progress toward mature social cognition, and there is not necessarily one point at which it can be declared that they have acquired Theory of Mind. By the age of one year old, children point to objects and follow the direction of others' gaze, showing that on some level they know other people's attention might be directed differently from their own. Sometime in their second or third year, children develop the ability to play pretend and to lie, indicating that they can conceive of alternate realities, but not necessarily that they think these could be true or believed to be true by someone else.

Another limit to Theory of Mind and its being tested is that young children have much less working memory than adults, making them unable to manipulate several facts at the same time. It is possible that children are capable of understanding each of the individual components required for representing another person's mental state but incapable of actually doing so because they cannot put the different pieces together. For example, in order to pass our test with the bowls (described in detail below), the subject must use rather complicated logic: the observer was absent when the bowls were switched, so he will point to the bowl on the side where the token was before he left the room, but because the bowls were switched the token is now under the one on the opposite side from where it was earlier, so the correct choice is whichever bowl the observer does not choose. This abstract reasoning is likely the product of a mental faculty separate from the social aspect of Theory of Mind, but it nevertheless affects related behavior.

Background on Modal Verbs and their Acquisition

Modal verbs, in English, are auxiliary verbs that express the modality of a statement. This means, roughly, that a statement with a modal verb added to it changes from being a declaration of a fact to some sort of qualification of that fact. While there have been observed to be many different types of modality, to a first approximation all English modal verbs are capable of being interpreted as either epistemic or deontic, depending on the linguistic context and external circumstances. In an epistemic reading, the modal applies to the speaker's view of the proposition's truth-value, whereas in a deontic reading, the modal indicates some requirement or obligation involving the elements in the proposition.

Syntactically, modals are a type of Inflection and are used with a non-finite form of the main verb. It has been argued that the two interpretations involve different syntax, with deontic modals behaving as control verbs and directly interacting with the subject, while epistemic modals behave as raising verbs (e.g., Jackendoff 1972, cited in Holt & Modyanova 2004). In

many cases, this makes intuitive sense.

1. a. John must crack an egg. (deontic)
- b. John must have cracked an egg. (epistemic)

In (1a), it would seem that *must* is theta-marking *John*, so the whole phrase *must crack an egg* is something that applies to John, whereas in (1b), *must* applies to the whole proposition of John having cracked an egg. (1a) could be paraphrased as “It is a requirement of John that he crack an egg,” and (1b) as “It is a necessary assumption of the world that John cracked an egg.” However, this account does not hold for all sentences.

2. a. An egg must crack.
- b. An egg must be cracked.

Both examples in (2) can be interpreted deontically, yet the modal *must* cannot be applying to the subject because *an egg* does not specify any particular egg upon which to impose the requirement. Therefore, the modal *must* again be applying to the whole proposition, as “It is a requirement of the world that an egg crack.”

The use of modals in spontaneous speech shows gradual development in children starting sometime in the later part of the second year. Wells (1979 and 1985) (cited in Papafragou 1998) showed that deontic use of modals is seen significantly earlier than epistemic. By the age of 3;3, children in his sample were using several different modals deontically, but even by the age of 5 many of them were not using modals epistemically at all. Holt and Modyanova (2004) point out that in acquiring language, children might make the assumption that if two sentences have the same phonetic form, any difference in their meanings must arise from differences in linguistic structure. Children do not acquire raising until the age of about seven years (Hirsch and Wexler 2007), yet they use modals in their speech, so they must be analyzing them differently from how adults do, possibly as the simpler control verbs proposed above. Thus, they have only one structural option for sentences involving modal verbs, and this grammar is only compatible with

the deontic interpretation, so that is the only meaning they can understand.

For both epistemic and deontic readings, the meanings of modal verbs fall along a continuum of strength (necessity versus possibility for epistemic, obligation versus permission for deontic), with some overlap but very little variation for any given word itself, implying that this aspect is part of each word's lexical entry and not dependent on context. As children get older, they make finer distinctions along the necessity-possibility scale. Moore, Pure, and Furrow (1990) show that children between three and six show a significant positive correlation between age and ability to distinguish degree of certainty in pairs of modals. Importantly, this effect was also found for non-modal verbs (e.g., "think" versus "know") and adverbs (e.g., "probably" versus "maybe"), suggesting that it is not a result of the linguistic construction but of the children's developing ability to understand speaker uncertainty.

It has also been shown (Noveck, Ho, & Sera 1996) that children's interpretation of epistemic modals can take context into account. When presented with two mutually incompatible statements and no situational information about their truth, children would choose to believe the one with the modal that expressed more certainty. But when there was additional nonlinguistic information available, the children would choose the statement that agreed with their knowledge, even if it used a less forceful modal. This finding demonstrates that children are able to incorporate the pragmatic context into their understanding of linguistic propositions involving modal verbs.

Argument and Objectives

Understanding of epistemic modals is often thought to be related to the understanding that there can be alternative possible realities, since their usage is in specifying the truth-value of a particular such reality. For this reason, it could be argued that children's failure of Theory of Mind tests is not due to their inability to interpret modal verbs as epistemic but rather that the

same cognitive capacities underlie both Theory of Mind and comprehension of epistemic modals. However, we reject this conclusion because, as discussed above, children show signs of having Theory of Mind and being able to represent speaker uncertainty before they are able to use epistemic modals. We propose that children are in fact interpreting all modal verbs as using a control syntax, which only makes sense for the deontic reading.

Through this experiment, we hope to provide evidence that children below the age of three or four can pass nonverbal Theory of Mind false-belief tests, but that they are unable to interpret modal statements as epistemic despite clear contextual information.

Methods

The modals test consists of eight test scenarios involving four different modal verbs – one *may*, one *can*, two *must*, and four *will* – and one control scenario for each of the four verbs. The control scenarios are written to be as ambiguous as possible regarding whether the intended meaning is epistemic or deontic, so any unexpected results in the test scenarios should be explained by performance in the controls. For example, it has been observed that children are more likely to respond affirmatively to yes-or-no questions because they are eager to agree with the questioner, regardless of their actual belief of the answer. We have also tried to account for this fact by making some of the modal statements negative and some positive. For each of the eight test scenarios, there are four versions – fulfilled and unfulfilled for deontic and epistemic – and for the control scenarios there are two versions, fulfilled and unfulfilled.

Whether the modal is to be interpreted as deontic or epistemic is indicated by the context in which it occurs. In deontic scenarios, the modal statement is often similar to a command, and its truth is supported by circumstantial evidence. In epistemic scenarios, it is made clear that the modal statement represents the speaker's beliefs, and it does not necessarily coincide with what might be concluded from the circumstances alone.

Each scenario involves a mother and a child, and some of them also contain other characters (a father or a grandmother). The backgrounds are drawn on white printer paper in color, with enough detail to be identifiable but not realistic-looking, and the characters are greyscale printouts of characters from the television show “Family Guy.” The characters’ mouths and body parts cannot move independently, but it is assumed that the viewer will be able to determine who is speaking by their different voices and the way they address one another. The scenarios are videotaped, and are shown to the subjects on a computer.

At the beginning of each scenario, there is a short conversation (a few sentences at most) between two of the characters, in which one of them makes a statement using a modal verb. For an adult listener, it should be clear from the context and the phrasing whether the statement is meant as epistemic or deontic. The video is paused, and the subject is asked what he thinks will happen next (or, in some cases, a more specific question pertaining to the exact statement made by the character). Then the rest of the video is played, with either the fulfilled or the unfulfilled outcome (which outcome is played for each scenario is determined beforehand, as explained below, and does not depend on the subject’s prediction). The subject is asked whether he is surprised at what happened and whether the character who made the modal statement lied.

Each subject is tested with only one version of each of the twelve scenarios, but they are presented in the same order for every subject, alternating between “will” and other modals. The versions are put together such that each child gets a total of two fulfilled and two unfulfilled controls, one of each condition (i.e., deontic fulfilled, etc.) with “will,” and one of each condition with other words. Currently we do not have enough scenarios to present each of the four types with each of the four verbs to every subject, so we will have to use cross-sectional analysis and assume that age is the only relevant factor that varies among subjects.

For the Theory of Mind test, there are two bowls, which are placed upside-down in two

unchanging locations on a table. The experimenter has a bag of plastic tokens, and during each trial, he places a token under one of the bowls. There is another person, the “communicator,” who can always see where the experimenter initially places the token. At the end of each trial, the subject is asked to touch the bowl where he thinks the token is, and if he is correct, he is given the token (he will get stickers for the tokens he has at the end of the experiment).

In the practice trials, the communicator does not ever leave the room, and the experimenter does not move the token. Sometimes there is a screen placed in front of the bowls so that the subject cannot see where the token is hidden. The experimenter asks the communicator to indicate which bowl contains the token, and in these trials his indication is always correct since he has been watching the whole time. In the cases where there is no screen, the subject will be able to see that the communicator always indicates correctly and is not trying to trick him.

In the control trials, there is always a screen at the beginning, and it is removed after the experimenter hides the token. The communicator then leaves the room and the experimenter either switches the places of the bowls (so the subject still does not know the location of the token) or takes the token out and moves it to the other bowl. The communicator is asked to indicate, either before leaving or after returning, where he thinks the token is. In cases where the bowls are switched, rather than the token, the communicator’s indication occurs before he leaves. Performance in these control trials will show whether the subject actually understands what is going on and is using the expected reasoning to make his choices.

In the test trials, the screen is still present as in the controls, and the switch is always of the bowls and not the token, so the subject never sees where the token is hidden and must rely entirely on the indication of the communicator, which always occurs after the bowls have been switched. For most of the trials, the communicator leaves the room and thus does not see the

switch, but sometimes he is present the whole time. If a subject has Theory of Mind, then he should indicate the bowl that the communicator did not indicate in the trials where the communicator left the room and did not see them switched; if he does not have Theory of Mind, then he should always indicate the same bowl as the communicator, even when he knows that the communicator did not see the bowls switched.

Results

Some of the original versions of the scenarios were informally piloted on MIT undergraduates by reading the scripts out loud, and in these trials the results were as expected for adults. All of the subjects predicted fulfillment for the deontic cases, but at least one subject predicted unfulfillment for one of the epistemic cases. Subjects were more likely to be surprised by unfulfillment and consider it lying in the deontic versions than in the epistemic or ambiguous (control) ones. These results show that the context provides adequate information for epistemic and deontic meanings to be differentiated from each other, so if children's responses do not differ between the two, then it is likely because they consider them to have the same meaning.

The final set of twelve video scenarios was run on five adults (MIT students) and thirteen children (ages 3;7 to 5;4). The results were more complicated than had been expected, with significant variation in responses based on which modal was being used.

For the adults, their responses to "may" scenarios were exactly as predicted: only in the deontic cases did they expect fulfillment, be surprised at unfulfillment, and consider unfulfillment lying. With "can" and "must," they almost always expected fulfillment but rarely said they were surprised or that a character lied. For "will," interestingly, they tended to expect unfulfillment and be surprised at fulfillment, yet they still claimed that the character lied only in unfulfilled cases.

The children's responses were much less consistent. They almost always predicted

fulfillment, though slightly less for “will” than for the other modals. For “can,” they were generally surprised at unfulfilled outcomes and said the character lied. For “may,” they were sometimes surprised at unfulfillment and indicated lying, but only for epistemic cases. For “must,” about half of the time they indicated surprise and lying for both fulfilled and unfulfilled cases. For “will,” they were surprised about half the time for either outcome, though somewhat less for deontic unfulfilled, and indicated lying significantly more in the unfulfilled cases.

Possible Problems

The scenarios were constructed primarily with the purpose of being able to distinguish between epistemic and deontic, and there was not much consideration of other facts that could potentially be important. The stories are different for each modal verb; in order for the modals to be compared directly, we would have to make versions of every scenario with each of the four words. Some of the scenarios use “you” as the subject of the modal statement, some use “I,” some use “he,” and some use the expletive subjects “there” or “it.” In the epistemic scenarios, while it would usually be clear to an adult that the modal is expressing the speaker’s belief, there is not always adequate contextual information making this clear, so it would be possible for children to predict fulfillment from factors other than the character’s statement.

Another complicating issue is that expectation of fulfillment is not perfectly correlated with belief of obligation. We were assuming that the only factor determining subjects’ fulfillment predictions would be whether the character was required to do something (as set up by a deontic modal) versus believed to do something (as set up by an epistemic modal), and that they would always predict fulfillment for the deontic situations. But in some cases, especially with young children, this was an oversimplification. For example, when a character says he does not want to finish his sandwich because he is full, children would sometimes predict that he would not eat the sandwich even when his mother said he “must” do it.

Additional problems arose during the administration of the test. It was not always clear to the children what was meant by “what happens next,” so they sometimes had to be asked specifically whether what the character had claimed would be true (using the present tense and no modals). Some of them did not seem to understand the question of whether they were surprised, first trying “yes” and then trying “no” when asked to give a reason for their answer, as if they were seeking confirmation for the correct response. Some children also had a hard time ignoring the previous scenarios’ outcomes, predicting, for example, that the character would not clean his room because he did not do so the last time.

Many of the younger children did not know the relevant meaning of the word “lie,” interpreting it as “fall down” or “go to sleep” (as in “lie down”), so they could not be assessed on that part of the test. When the word was explained to them, it was not clear that they grasped the concept, repeating it back as “only joking” or not being able to paraphrase it at all. Some appeared to know the meaning of “tell the truth,” but since this phrase does not assume anything about the speaker’s belief, it could not be used as an opposite or substitute for “lie.” A few of the subjects did not know what the word “surprised” meant, and it was difficult to tell whether they understood the question because some would enthusiastically answer “yes” even when they did not know what it was asking.

Conclusion

It is still not clear from these results what the cause is for the timing of children’s acquisition of modals. They certainly suggest that children’s grammar is doing something differently from adults’, but it remains to be seen exactly what this is. With further refinement and testing on more subjects of different age groups, and comparison within subjects with the nonverbal Theory of Mind test, this experiment could provide insight into the nature of the relationship between children’s developing cognitive and language abilities.

References

- Butler, J. 2003. A minimalist treatment of modality. *Lingua* 113: 967-996.
- Hirsch, C., Wexler, K. 2007. The late acquisition of raising: What children seem to *think* about *seem*. In S. Dubinsky and B. Davies (eds.) *New horizons in the analysis of control and raising*. New York, NY: Springer.
- Holt, A., Modyanova, N. 2004. What Children (and Linguists) Do and Don't Know about Modals: The Case of "Will" and Implications for Theory of Mind. MS.
- Jackendoff, R. 1972. *Semantic interpretation in generative grammar*. Cambridge: MIT Press.
- Kosslyn, S., Rosenberg, R. 2006. *Psychology in Context*. Boston: Pearson Education, Inc.
- Moore, C., Pure, K., Furrow, P. 1990. Children's understanding of the modal expressions of certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development*, 61: 722-730.
- Noveck I., Ho S., Sera M. 1996. Children's understanding of Epistemic Modals. *Journal of child language*. 23:621-643.
- Papafragou, A. 1998. The Acquisition of Modality: Implications for Theories of Semantic Representation. *Mind & Language*, 13 (3): 370-399.
- Papafragou, A., Ozturk, O. 2007. Children's Acquisition of Epistemic Modality. *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, ed. Alyona Belikova, Luisa Meroni, and Mari Umeda, 320-327. Somerville, MA: Cascadilla Proceedings Project.
- Perner, J., Leekam, S. R., Wimmer, H. 1987. Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5(2): 125-137.
- Wells, G. 1979. Learning and Using the Auxiliary Verb in English. In V. Lee (ed.), *Cognitive Development: Language and Thinking from Birth to Adolescence*. London: Croom Helm, 250-270.
- Wells, G. 1985. *Language Development in the Pre-school Years*. Cambridge University Press.