

CAN VS. BE ABLE TO: WHY 'SEMIMODALS' ARE NOT MODALS

RYAN FISHER & PATRICK J. DUFFLEY  
*Université Laval*

5

PRE-PRODUCTION COPY - NOT FOR CITATION OR DISTRIBUTION

IN ONE OF THE MOST INFLUENTIAL STUDIES ON THE TOPIC, Palmer (2001:4) presents modal auxiliaries as one of the ways in which languages deal grammatically with the cross-language category of “modality,” a category of the clause which denotes “the status of the proposition that describes the event.” The main division within this category is that between 10  
realis, defined as portraying situations “as actualized, as having occurred or actually occurring, knowable through direct perception,” and irrealis, which portrays situations “as purely within the realm of thought, knowable only through the imagination.” Palmer recognizes that he is taking a typological approach to modality and using categories that are external 15  
to the linguistic system of any particular language—he calls them “notional” (18) to avoid becoming embroiled in the semantics-pragmatics debate. The present paper intends to go beyond the purely notional-typological level, building on the insights offered by the realis/irrealis distinction, but delving into the language-specific semantics of English in order to propose a more complete and explanatorily adequate account of the difference in behaviour 20  
between the modal auxiliary CAN and its closely related semi-modal cousin BE ABLE TO.

In two insightful papers (1977, 1980), Palmer explores the contrast between these two items, a treatment which is summarized and completed with a few further minor observations in the second edition of his *Modality and the English Modals* (1990). In the 1977 article Palmer offers an explanation for the fact that we cannot normally say \**I ran fast and could* 25  
*catch the bus*, suggesting that CAN is not used if there is an implication of actuality (i.e., that the event actually took place) in the past, because the factual status of the event is known and a modal verb is inappropriate in such circumstances. In contrast, it is perfectly possible to say *I ran fast and was able to catch the bus*, because BE ABLE TO is not a modal. In the negative, *I ran fast, but I couldn't catch the bus* is also acceptable, since negative modality implies 30  
negative actuality (if the subject couldn't perform the event, it follows that he didn't).

In his 1980 paper, Palmer completes the picture, bringing out two further points. He shows first of all that while the strongest restriction on CAN in terms of the expression of actuality is with single actions in the past, even in other circumstances BE ABLE TO may be preferred if the implication of actuality is strong. This occurs in the present tense in cases such as: 35

- (1) By bulk buying in specific items, Lasky's are able to cut prices on packages by as much as 30 per cent or so. (Palmer 1980:96)

It also occurs in the past with reference to repeated actions in cases where *could* is the more usual form (cf. *I could get up and go to the kitchen whenever I wanted to*) if there is a very strong implication of actuality: 40

(2) Most people worked harder than me during the University, of course, and when it came to exams, they were able to draw not just upon two weeks of knowledge. They were able to draw upon three years of knowledge. (*ibid*:95)

5 The second point which Palmer makes is that the conditioning factor permitting the occurrence of CAN referring to single actions in the past must be extended from ‘negation’ to include the notions of the almost nonoccurrence of the event in the past (3), its occurrence under difficult circumstances (4), the restriction of its occurrence to a particular agent (5), and the questioning of its occurrence (6).

10

(3) I could just reach the branch. (95)

(4) He was laughing so much he could hardly get a word out. (94)

(5) Well she was the only one of the family there who could do it. (94)

(6) Could you reach it? (92)

15

Palmer concludes that:

[I]t is fairly clear, then, that it is semantic negation that is at issue, but in a wide sense, to include not only nonoccurrence of the event but also its occurrence under

20

difficult circumstances, or its “almost nonoccurrence.” (1980:95)<sup>1</sup>

While Palmer’s account is basically sound, the relation which it suggests between the facts of usage described and the meanings of the linguistic units involved needs to be made much clearer, especially in the case of BE ABLE TO. Neither the complete range of its uses nor the contribution of the language-specific semantic content of the words constituting this phrase to its expressive capacities in discourse are considered, and so no real explanation is offered as to how or why it is suppletive for CAN in certain contexts. Westney (1995), in an attempt to build on Palmer’s actuality hypothesis, claims that actuality must be seen as an idiosyncratic feature of the meaning of the BE ABLE TO construction. This requires clarification, however, since actuality can also be associated with other semi-modal constructions such as HAVE GOT TO and HAVE TO (cf. *I had to leave work at 3:30 yesterday to pick up my daughter Rosalie*), and so would not seem to be merely idiomatic with BE ABLE TO. As for CAN and COULD themselves, it is possible to develop even further some of Palmer’s intuitions about the restrictions which apply to these forms as a consequence of the type of meaning expressed by the modals.

35

As a first point, the notions of “semantic negation” or “affective contexts” (1990:9495), which Palmer rightly invokes as a factor allowing the use of COULD, are much broader than his examples show. COULD also occurs in reference to single past actions in many other types of contexts:

40

---

<sup>1</sup> Cf. Palmer (1990:95): “What all of these have in common is that the possibility is in some way qualified, or that there is a suggestion that the event almost did not take place, or that it took place in a minimal way.”

- (7) The most the doctor could tell me was that she was very sick.  
 (8) I had to wait outside for two hours before I could talk to her.  
 (9) He got to the scene of the accident sooner than I could.

The best term to describe the full range of the uses of **COULD** referring to a single past action is Jacobsson's notion of "nonassertiveness" (1974:62). This can embrace both the interrogative usage illustrated in (6) and the uses given in (7)–(9) above: in (7) the doctor couldn't tell the speaker any more than that she was very sick; in (8) the speaker couldn't talk to her for two hours; in (9) the speaker couldn't get to the scene of the accident sooner than the other person referred to.

Seeing these uses of **COULD** in terms of nonassertiveness has the further advantage of allowing one to relate them to the modal uses of the verbs **NEED** and **DARE**. As Jacobsson has shown for **NEED**, such usage occurs only in contexts in which the notion (of necessity or obligation) is "not asserted but denied, questioned, conceded (in concessive clauses), or represented as a mere conception rather than as a positive fact" (1974:62). He gives the following list of contexts where he has found **NEED** auxiliary, which encompasses all of the usage where **COULD** doubles for **WAS ABLE TO** in reference to single actions in the past:

- (10) a. Need I be present?  
 b. He needn't come.  
 c. No one need know.  
 d. I need hardly say how glad I am.  
 e. He need only state his opinion clearly.  
 f. Standards are lower than they need be.  
 g. His book covers most that need be said on the subject.  
 h. I have half an hour to spare before I need go.  
 i. It is embarrassing that such a truth need be stated at all.  
 j. However much need be said, let it wait.

In contrast, the auxiliary use of **NEED** does not occur in pure affirmative contexts:

- (11) \*She need be operated on right away.

This behaviour closely resembles that of **COULD** which, as seen above, is not used in assertive contexts when the infinitive evokes a single action actually realized in the past.

Based on the pattern of usage found with **NEED** and **DARE**, it has been argued that the notion of nonassertiveness is a basic and inherent characteristic of the modal auxiliaries, which are defined semantically by their opposition to reality (Duffley 1994). Usage with **COULD** in reference to single actions in the past provides another confirmation of this analysis and so should be correlated with the modal uses of **NEED** and **DARE**, which Palmer's treatment of the question neglects to do. **NEED** and **DARE** denote notions which refer a priori to realities, and it is only when they are not asserted that they can be construed as mere potentialities equivalent to the modal auxiliaries. **CAN/COULD**, on the other hand, evoke



5 **Figure 1.** *Conceptual relations involved in modal usage of DARE.*

potentiality by nature, and the restrictions we have observed on the use of *COULD* when its infinitive refers to a single past action are the product of a semantic conflict between the notion of past possibility and that of past realization: if the action really happened at a particular moment in the past, what is the point of treating it as merely possible? On the other hand, it does make sense to evoke a single action in the past as having been impossible, thereby implying that it was not performed.

While the explanation for the use of *COULD* in the negative is quite easy to follow, the motivation for the acceptability of *COULD* in non-assertive contexts where the actuality of the infinitive's event is implied, such as (3)–(5) and (7)–(9) above, requires deeper consideration. Palmer suggests that it is the fact that in all of these uses “the possibility is in some way qualified, or there is a suggestion that the event almost did not take place, or that it took place in a minimal way,” which accounts for there being a “focus on the modality, whereas in the examples where *could* would not be used the focus of attention is more on the actuality itself” (1990:95–96). This certainly points in the right direction but it still leaves a number of questions without a precise answer. Why does the suggestion that the event almost did not take place call to mind the notion of possibility? Why does restricting the predication to a subset of a larger set of subjects by means of a word like *only*, as in (5) above, cause a “focus on the modality” even though the subject singled out actually did perform the event denoted by infinitive?

Since it involves the conditioning factor of non-assertiveness, modal usage with *DARE* and *NEED* sheds considerable light on this question. Consider (12) below:

- (12) But launching a new cigarette in today's climate of disapproval requires finesse.  
 30 The company is advertising its Premiers as 'cleaner smoke', the furthest it dare go.  
 (*The Economist*, Sept. 17, 1988, p. 33)

This use obviously implies that the company in question dare not go any further than it has, but the sentence would still seem to assert real daring up to that point. Such cases can best be analyzed by taking into account the fact that some limit or restriction is imposed on the daring here. In (12), therefore, daring is declared to be impossible beyond the limit of advertising the new cigarette as a cleaner smoke, a situation which can be depicted by the diagram in **Figure 1** so as to visualize the conceptual relations which it involves.

The important point here is that daring is discussed in terms of its possibility beyond the limit. Since this is the case, it is only natural that it should also be evoked in terms of its possibility inside the limit as well. It is argued here that the same thing occurs with *COULD* in uses such as (5) above: if doing the action is evoked as having been impossible for the other members of the family, it is logical for it to be discussed in terms of its possibility for

the one person who was actually able to do it. As for (3), it is the fact that the adverb *just* implies the possibility of nonoccurrence which leads the speaker to see the occurrence in terms of its possibility rather than its reality (moreover, to say that a single past action was just barely possible implies that it actually occurred). The idea that an event occurred only with considerable difficulty (cf. *hardly*) also calls to mind the state of affairs prior to its existence when it was a mere possibility which very well might not have become an actuality due to the obstacles to its occurrence. 5

To sum up, there is a semantic aversion to using COULD when the infinitive's event corresponds to a single action which really took place in the past, since if a singular action really happened, one does not see the point of representing it as merely possible. This aversion is not felt, however, when the possibility of a singular action in the past is not asserted straightforwardly but only in a qualified manner.<sup>2</sup> This can involve the assertion of possibility up to a certain limit, or its restriction to certain subjects, notions which both imply impossibility outside of these limits or restrictions, and therefore lead the speaker to view the situation in terms of possibility rather than reality. It can also involve the assertion of a possibility which just barely escaped the realm of impossibility, or did so only with difficulty, which also leads quite naturally to viewing the actually realized past action in terms of its possibility rather than its reality. 10 15

We can now turn our attention to BE ABLE TO, the construction which is suppletive for COULD in past contexts where a singular action was actually realized. The first point which must be made is that a sentence such as (13) is in fact ambiguous without a context: 20

(13) He was able to lift 200 pounds.

This sentence can imply that the person in question actually did lift this weight on a particular occasion, as in Palmer's examples; but it can also evoke the subject's merely having the ability to lift 200 pounds, a sense which overlaps with that of COULD expressing past ability. This is perhaps clearer in (14), said by a professor who had just finished giving the explanations required for the homework being assigned: 25 30

(14) You are now able to do Exercise 9.

Here BE ABLE TO clearly does not imply actuality.<sup>3</sup> The real question to be asked in comparing COULD and WAS ABLE TO is therefore this: why can WAS ABLE TO evoke either 'he 35

<sup>2</sup> It goes without saying that there is no resistance to evoking the past possibility of a single event whose actualization is denied or questioned.

<sup>3</sup> This would also be the interpretation Palmer's example *Are you able to take that on?* (1990:92). Further examples found include: 40

- (a) He was convinced that we were all perfectly able to write, spell and figure, but that we were making a show of being misinformed... (LOB R04 109 10)
- (b) Christ Jesus... who is able to preserve you without sin ... (BUC D16 0750 8)

did x' or 'he had the ability to do x'/'it was possible for him to do x', while COULD can only evoke the notion 'he had the ability to do x'/'it was possible for him to do x'?

The answer to the question of why COULD is restricted to the idea of having an ability/possibility lies in the nature of the modal auxiliaries, which evoke mere states of potentiality. This is amply confirmed for past COULD by Palmer's (1990:93ff) observations on the contexts where COULD can be used (besides negatives and interrogatives) and by his comments on the meaning of COULD in these contexts:

1. "Could may be used [in the affirmative] if there is no implication of actuality, but only a statement of possibility:

(15) I was plenty scared. In the state she was in she could actually kill."  
(Ehrman 1966:50)

2. "If there is an indication not of a single action, but of successive or habitual actions, *could* may be used:

(16) I could get up and go to the kitchen whenever I wanted to.

- (...) It may be that could is permissible in such circumstances because there was a general possibility over a period of time, not a possibility that resulted in a single definable action."

3. "[...] it must be admitted that the following are rather unlikely to occur (except in a habitual sense):

(17) ? I ran fast and I could just/almost/nearly catch the bus.

(18) ? I could catch the bus because I ran fast.

- The reason seems to be that catching a bus is a momentary activity and there is no continuing possibility of this action. By contrast, one can be in a position to reach a branch or enter a house for some time [cf. *I could just /almost / nearly reach the branch; I could reach the branch because it was loaded down*]. It seems that the occurrence of *could* with *just*, *nearly*, *hardly*, *because...* etc. depends on the nature of the event and the duration of the possibility.

- It is where *could* is closest to *did* that it becomes most unacceptable."

Even a case which Palmer (1990:96) describes as "more difficult to explain" due to the fact that "there is a clear implication of actuality" can be accounted for in terms of a state of possibility:

- (c) He finally concluded that I was able to take care of myself. (Oxford English Dictionary, sub *able*)  
(d) [H]e was fully able to take care of her. (F. Scott Fitzgerald 1925:178)  
(e) I'm well able to handle this alone, Mr. Slocum. (O'Neill 1947:543)

(19) Jane darling, I'm so glad you could make it.

Here the host is not merely, as Palmer suggests, "saying he is glad that the guest has arrived;" he is rejoicing at the fact that this fortunate event was *possible*. The use of COULD also allows the speaker to allude discreetly to his/her guest's efforts at surmounting any obstacles to her coming which could have made the latter impossible. 5

As regards the capacity of BE ABLE TO to evoke both the sense of having an ability and that of exercising it, the explanation lies in the language-specific meanings of the component parts making up this construction. The adjective *able*, although it evokes a quality like all adjectives, denotes a quality which implies activity. Comparison of the noun phrases *an able man* and *a capable man* shows that whereas *capable* can describe a man who has talent but is doing nothing with it, *able* depicts the man as active or at least as ready for immediate action. This nuance shows through in the complement forms used after these two adjectives. *Capable* is construed with the preposition *of* followed by a noun or an *ing* form (*he was capable of murder*); *of* signifies that murder is a point of reference on the basis of which the capacity is defined (an abstract version of the notion of origin). *Able* is followed by the preposition *to* (idea of movement) plus the infinitival form of the verb (which is less nominal than the *ing* form or a noun and denotes the full actualization of an event in abstraction from any specific place in time). The meaning of the BE ABLE TO construction thus involves an orientation towards the actualization of the infinitive's event which is not present in the *capable of* construction with *of* followed by a gerund or noun. 10 15 20

Indeed the meaning of the preposition *to* makes a very important contribution to the resultant meaning of the BE ABLE TO phrase. When used with the infinitive, this preposition evokes the latter's event as the endpoint of a movement in time of its potential actualizer from a before-position to an after-position (cf. Duffley 1992:16–17). In the construction under study, the potential actualizer is identified by the subject of BE, the before-position corresponds to the latter's possession of the ability denoted by ABLE and the after-position to the (prospective or subsequent) realization of the infinitive's event. Furthermore, because *to* evokes the idea of a movement, this movement can be construed as either nonrealized (as in *I wanted to leave early*) or realized (as in *I managed to leave early*). The two senses of (13) above can thus be explained on the basis of the meaning of *able*, which involves orientation towards immediate action, and the meaning of *to*, which introduces the idea of a movement conceivable as either realized or nonrealized, notions that are not present in the construction with COULD, which as a modal merely denotes a state of possibility with respect to the event expressed by the infinitive following it. 25 30 35

Nor should one neglect to take into account the meaning of the verb BE in the ABLE TO construction. At first sight, the uses under consideration here might seem to involve the stative copula sense. However, it must not be forgotten that BE plus adjective can also be used by a speaker to evoke a dynamic sense, as in one reading of (20): 40

(20) He was sick twice.

A dynamic sense is also present in the use of BE plus adjective in the progressive construction:



(21) He was being childish.

It will be proposed here that it is the dynamic sense of BE which occurs in the BE ABLE TO construction when the latter evokes the infinitive event as actually realized. Confirmation of this hypothesis is provided by uses such as (22) below, a use involving the progressive construction much like (21) above:

(22) We are gradually being able to get our message across.

10 To complete the discussion of CAN and BE ABLE TO, it is worth noting a parallel phenomenon with MUST and HAVE TO. Palmer points out two relevant facts (1980:98) concerning this pair, but once again his observations require some refinements and, above all, call for explanations. He remarks, first of all, that HAVE GOT TO and HAVE TO can imply actuality in the present whereas MUST does not:

15

(23) It's a slow walk down, he's got to fight his way through the crowds.

(24) It's a slow walk down, he must fight his way through the crowds.

In (23) "it may well be that the person referred to (a boxer coming to the ring) is actually engaged in fighting his way through the crowd; with the second using *must*, this interpretation would be impossible." The second pertinent fact is that "the past tense forms of HAVE TO and HAVE GOT TO imply actuality," as in (25) below, and are almost never used to evoke mere past necessity:

25 (25) We had to make a special trip to Epsom to collect the bloody thing,

This latter observation must be qualified however; it quite commonly the case in contexts such as (26) that HAD TO evokes the necessity in the past of an event whose actualization is not necessarily implied.

30

(26) Why did they leave so suddenly?

They had to go to the dentist's.

This means that we observe the same two possibilities with HAVE TO as with BE ABLE TO, and so can oppose MUST and HAVE TO in similar terms to those applied to CAN and BE ABLE TO above. Accordingly, (27) can imply either 'he has an obligation to leave at 4:00' or 'he leaves every day at 4:00 because he is obliged to'.

40

(27) He has to leave at 4:00.

MUST, on the other hand, as Palmer observes, can only refer to the future, i.e., it evokes merely a state of necessity or obligation in the present and cannot imply the actuality of the infinitive's event as a result of the obligation:



(28) He must leave my class every day at 4:00.

The explanation for the capacity of HAVE TO to evoke the actualization of the event expressed by the infinitive is similar to that for BE ABLE TO. Like the latter, the HAVE TO phrase contains the preposition *to*, with the consequent possibility of a double construal according to whether the movement signified by the preposition is conceived as realized or nonrealized. Moreover, the first component of the construction, the verb HAVE, is also conceivable in both a stative and a dynamic sense, as shown by the existence of uses such as:

(29) I am having to take an aspirin every two hours.<sup>4</sup>

It is plausible therefore that this dynamic sense combines with the realized construal of the movement signified by *to* when HAVE TO signifies actuality of the infinitive's event as in (25) above. Before any definitive conclusions can be made, however, a proper investigation of the semantic contribution of the meaning of the verb HAVE is necessary, an undertaking which is beyond the scope of the present study.

The authors hope that this paper has completed Palmer's insights into the nature of the modal auxiliaries and their contrast with the semimodals in English by going more deeply into the level of language-specific explanation. Palmer's intuitions about the modal auxiliaries being opposed to the notion of actuality have been clarified by showing more clearly their relation to the semantic category of nonassertiveness and the analogy with the modal uses of NEED and DARE. More importantly, however, the meaning-capacity of the BE ABLE TO and HAVE TO constructions has been explained by the semantic contribution of each of their component parts to the resulting meaning of these phrases in discourse, thereby providing a framework of analysis for the numerous other semimodal constructions in English. The analysis proposed here provides further evidence that these so-called 'semimodals' (Palmer 1990:25) are not modal auxiliaries, while at the same time allowing us to understand how they are so closely related in meaning to the latter. Because BE ABLE TO contains the notions 'be', 'ability', and 'movement', it can be practically equivalent to the state of possibility evoked by CAN when the movement signified by *to* is conceived as nonrealized and BE has its stative sense. Ironically enough, however, the use which has merited it the appellation 'semimodal', where it is suppletive for COULD in reference to single actions in the past, is precisely a case where CAN is not appropriate because it is a modal, whereas BE ABLE TO is utilizable because it is not.



## REFERENCES

BROWN UNIVERSITY CORPUS OF AMERICAN ENGLISH [BUC]. 1964. Providence RI: Department of Linguistics, Brown University.

<sup>4</sup> Cf. Haegeman (1980) for further examples.

- DUFFLEY, PATRICK J. 1992. *The English infinitive*. London: Longman.
- . 1994. *Need and dare*: The black sheep of the modal family. *Lingua* 94:213–43.
- EHRMAN, MADELINE. 1966. *The meanings of the modals in present-day American English*. The Hague: Mouton.
- 5 HAEGEMAN, LILIANE. 1980. *Have to* and progressive aspect. *Journal of English linguistics* 14:1–5.
- JACOBSSON, BENGT. 1974. The auxiliary *need*. *English studies* 55:56–63.
- NEHLS, DIETRICH. 1986. *Semantik und syntax des englischen verbs*, vol. 2. Heidelberg: Julius Groos.
- 10 LANCASTER/OSLOBERGEN CORPUS OF BRITISH ENGLISH [LOB]. 1978. Oslo: Department of English, University of Oslo.
- O'NEILL, EUGENE. 1947. *The plays of Eugene O'Neill*. New York: Random House.
- PALMER, F. R. 1977. Modals and actuality. *Journal of linguistics* 13:1–23.
- . 1980. *Can, will, and actuality*. In *Studies in English linguistics for Randolph Quirk*, ed. S. Greenbaum, G.N. Leech & J. Svartvik, 91–99. London: Longman.
- 15 ———. 1990. *Modality and the English modals*. London: Longman.
- . 2001. *Mood and modality*. Cambridge: Cambridge University Press.
- SCOTT FITZGERALD, F. 1925. *The great Gatsby*. New York: Scribners.
- WESTNEY, PAUL. 1995. *Modals and periphrastics in English: An investigation into the semantic correspondence between certain English modal verbs and their periphrastic equivalents*. Tübingen: Niemeyer.
- 20



25

30

35

40