



# A Reanalysis of the Uses of *Can* and *Could*: A Corpus-Based Approach

Lauren Whitty<sup>1,2</sup> 

Received: 1 February 2019 / Accepted: 11 June 2019 / Published online: 18 June 2019  
© Springer Nature Switzerland AG 2019

## Abstract

This study offers an in-depth analysis of the English modal auxiliaries CAN and COULD, using both spoken and written components of the British National Corpus. An examination of previous corpus-based studies of the modal auxiliaries CAN and COULD highlights discrepancies in the terminology utilised and the main categories associated with CAN and COULD, as well as insufficient surrounding context for a confident categorisation and a lack of clarity in explanations for classification. Based on findings from a new investigation of these modal auxiliaries in the BNC, I argue for a wider range of usage categories for CAN and COULD. The categories identified here differ from those reported in previous studies, as the present study differentiates categories of use beyond the traditional distinction between ‘ability’, ‘possibility’ and ‘permission’. This study offers transparency on categorical criteria and the usage category assigned to individual tokens and demonstrates expanded context is an essential requirement in the semantic and pragmatic (re)analysis of corpus data.

**Keywords** English modal auxiliaries · Corpus analysis · Modal uses · British National Corpus

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s41701-019-00058-9>) contains supplementary material, which is available to authorized users.

---

✉ Lauren Whitty  
lauren.whitty@yahoo.com

<sup>1</sup> Victoria University of Wellington, School of Linguistics and Applied Language Studies, PO Box 600, Wellington 6140, New Zealand

<sup>2</sup> Present Address: City University of New York, Graduate Center, Linguistics Program, 365 Fifth Avenue, Room 7407, New York, NY 10016, USA

## Introduction

The purpose of this study is to investigate the occurrences of the central modal auxiliaries CAN and COULD<sup>1</sup> in corpus data, with a view to identifying the uses of these modals, as well as the frequencies of each use. The analysis presented here draws on existing corpus-based studies as well as a new detailed examination of the British National Corpus, henceforth BNC (Davies 2004). The main usage categories associated with CAN and COULD in previous studies on modality are ‘ability’, ‘possibility’, ‘permission’, and sometimes ‘epistemic’, with some variation in terminology. While ‘ability’ and ‘permission’ were the most consistent in their labels, their boundaries of classification differed, and were rarely delineated; the same discrepancies for boundaries hold for ‘possibility’ and ‘epistemic’ as well. These inconsistencies, as well as a lack of transparency, encouraged me to investigate the uses of CAN and COULD in a more contemporary corpus, the BNC.

Though many studies utilise the term ‘meaning’ (e.g. Coates 1983; Mindt 1995; Leech 2004; Römer 2004), I prefer the term ‘use’, with respect to modal verbs. Larsen-Freeman (2003: 34) presents the idea of ‘the three dimensions’—form, meaning and use—which she applies to ‘language in communication’. Larsen-Freeman (2003: 36) employs the following questions to help differentiate between the dimensions:

*Form:* How is the unit formed?

*Meaning:* What does it mean (its essential meaning)?

*Use:* When and why is it used?

Existing polysemic studies tend to assign multiple meanings to CAN and COULD, such as ‘ability’, ‘possibility’ and ‘permission’ (e.g. Coates 1983; Mindt 1995; Facchinetti 2002; Römer 2004), while some monosemic accounts have linked the term ‘possibility’ with *can* and *could* (e.g. Joos 1964; Papafragou 2000a; Westney 1995). I follow that the overall, or ‘essential meaning’, for CAN and COULD is ‘possibility’, with the current study’s focus on the uses of CAN and COULD, which include ‘ability’, ‘possibility’ and ‘permission’, as well as other uses that I will introduce. This focus on ‘use’ (as opposed to meaning) is more appropriate for this study as the descriptions of use are not only semantic but pragmatic as well. There is an abundance of previous work on both (e.g. Coates 1983; Collins 2009; Joos 1964; Klinge 1993; Mindt 1995; Palmer 1990; Perkins 1983), which has informed the analysis presented here.

The present study draws on the spoken and written components of the BNC to gain a better understanding the range and frequency of different uses of CAN and COULD, as well as utilising large amounts of context required for qualitative analysis and providing robust parameters of criterial classification to aid this analysis. This study differs from previous studies in its application of additional usage

<sup>1</sup> CAN represents spoken and written *can/can't/can not/cannot*. In the same way, COULD represents spoken and written *could/could not/couldn't*.

**Table 1** Summary of usage categories for CAN and COULD from selected corpus-based studies

Linguist/s	Traditional category			
	Ability	Possibility	Permission	Epistemic
Coates (1983)	Ability	Possibility	Permission	Epistemic
Mindt (1995)	Ability	Possibility/high probability	Permission	Certainty/prediction Necessity Inference/deduction
Biber et al. (1999)	Ability		Intrinsic-permission	Extrinsic-possibility
Facchinetti (2002)	Ability	Dynamic possibility	Deontic	Epistemic
Römer (2004)	Ability	Possibility	Permission	Hypothetical meaning
Collins (2009)	Dynamic	Dynamic	Deontic	Epistemic

categories for corpus analysis and its reliance on the inclusion of expanded context, which allows for a more transparent analysis.

### Previous Corpus-Based Investigations of *CAN* and *COULD*

Many previous investigations of *CAN* and *COULD* have been corpus-based, and each has had varying views on the usage categories associated with *CAN* and *COULD*. Perkins (1983: 10) notes that ‘the number of modalities one decides upon is to some extent a matter of different ways of slicing the same cake’. Table 1 shows how the general usage categories described above, ‘ability’, ‘possibility’, ‘permission’ and ‘epistemic’, are treated in six key corpus-based studies of the modal auxiliaries *CAN* and *COULD*.

Coates (1983: 86) applies the categories ‘ability’, ‘possibility’ and ‘permission’ to *CAN* in her study of the Lancaster Corpus and the Survey of English Usage. For *COULD*, she also distinguishes an additional ‘epistemic’ category. Mindt (1995: 75) draws on fictional texts in British English and analyses *can* and *could* using the combined categories ‘ability’, ‘possibility/high probability’, ‘permission’, ‘certainty/prediction’, ‘necessity’ and ‘inference/deduction’. Biber et al. (1999: 491) apply the categories of ‘ability’, ‘extrinsic-possibility’ and ‘intrinsic-permission’ to examine *can* and *could* in the Longman Spoken and Written English Corpus (LSWE). Facchinetti (2002) analysed *can* and *could* in the International Corpus of English—Great Britain (ICE-GB). Facchinetti’s frequency investigation is subdivided into very specific genres (e.g. direct conversations, telephone calls) and examines the meanings of these modal auxiliaries using a sample of 10% of the corpus. In addition to the traditional categories noted above, Facchinetti (2002: 242) also has a category titled ‘implication’, which ‘is strongly connected to the pragmatic context and intrinsically conveys the illocutionary force of requests, suggests, and proposals.’ Similar to Coates (1983), Römer (2004: 188) applies the categories ‘ability’, ‘possibility’, and ‘permission’ to *can* when analysing the spoken part of the BNC. For *could*, she also includes ‘hypothetical meaning’. Like Facchinetti, Collins (2009: 98, 109) looked at *can*

and *could* in the ICE-GB. Collins distinguishes the categories: ‘dynamic’, which includes ‘theoretical possibility’, ‘ability’ and ‘dynamic implication’; ‘deontic’ (permission); and ‘epistemic’.

From the above, discrepancies are evident not only in the terminology used, but in the number of categories used as well. For instance, for spoken *can* in the BNC, Römer (2004) includes three categories of meaning, ‘ability’, ‘possibility’ and ‘permission’, and offers no criterial information or descriptions of her categories. Mindt (1995: 75-6) does not use the term ‘epistemic’ but distinguishes: ‘possibility/high probability’, ‘certainty/prediction’, ‘necessity’ and ‘inference/ deduction’, with no explanation of the differences in his classification. Biber et al. (1999: 485) describe their category ‘extrinsic-possibility’ as ‘epistemic’ with three meanings within—possibility, necessity, or prediction, but no further clarification on what each signifies. Because linguists’ decisions about in what category to put a token differ from framework to framework, these decisions can contribute to the variances in reported frequencies.

A further complexity in previous studies of these modal auxiliaries is that there seems to be a lack of agreement in the categorical criteria for classifying individual tokens. For example, one criterion Coates (1983: 89) uses for identifying instances of ‘ability’ is ‘the possibility of the action is determined by inherent properties of the subject (this includes what the subject has learnt [...])’. Coates (1983: 89) offers the instance, ‘*I can only type very slowly as I am quite a beginner*’ as a ‘typical’ example. However, Collins’ (2009: 104) examples of ‘ability’ include, ‘And now I *can* see the Prime Minister, John Major’, which violates Coates’ criterion, as the ‘seeing’ in this instance is not the speaker’s ‘inherent’ or ‘learnt’ capability to see; the person did not undergo extreme circumstances (e.g. blindness to sight in only a moment). Further evidence that this instance does not fit Coates’ criteria for ‘ability’ comes from the text that follows, which is not included in Collins (2009) but found in the ICE-GB: ‘There’s a cheer from the crowd as he arrives’ (S2A-019 30). This additional context clarifies that it is possible for the speaker to see the Prime Minister due to his arrival, and not because of the speaker’s sight.

The difficulty with the example from Collins (2009) and other similar instances is twofold. Some studies list a usage category with examples representing that category, yet very little, or no, explanation is provided in support of the classification (e.g. Biber et al. 1999; Kennedy 2003; Mindt 1995; Römer 2004). For example, when describing the ‘core meanings’ of the modals, Kennedy (2003: 186–187) explains ‘Each modal has a core meaning or meanings’ and provides the following example for *could*: ‘We *could* leave it here. (possibility/permission)’. As with many of the previous studies, expanded context for the instance is also lacking. This makes it difficult for readers to have enough information to understand why a token has been assigned to a particular category. Furthermore, the differences in category labels and distinctions make it difficult to relate previous studies’ findings to one another. To address these issues, the present study offers clear criteria for usage categories and makes explicit connections to these criteria when presenting examples. Additionally, ample context is included with each modal token, to make the analyses as transparent as possible.

**Table 2** Forms collected for CAN and COULD from the BNC

Form	Mode	Quantity	Total
<i>Can</i>	Spoken	100	400
<i>Can</i>	Written	100	
<i>Can't</i>	Spoken	100	
<i>Can't</i>	Written	100	
<i>Could</i>	Spoken	100	200
<i>Could</i>	Written	100	
Total		600	

**Table 3** Comparison of BNC and sample data for COULD

	Spoken			Written		
	Sample (%)	BNC (%)		Sample (%)	BNC (%)	
<i>Could</i>	82	80	FE=0.8572, N.S.	93	91	FE=0.795, N.S.
<i>Couldn't</i>	18	20	FE=0.8572, N.S.	7	9	FE=0.795, N.S.

## Methods

This study relied on the BNC, which is comprised of data collected between the 1980s and early 1990s. I chose to use the BNC due to its integrity. The BNC has been referred to as a ‘finite, balanced, sampled corpus’ (Leech et al. 2001: 1) and ‘exceptional in that it is fairly ‘balanced’ yet very large’ (Biber et al. 1999: 27). Categories of use associated with the central modal auxiliaries CAN and COULD were identified using the BNC. The usage categories for CAN and COULD were developed through an iterative process, using the literature and previous studies for guidance at the outset, but later heavily reliant on the modal auxiliaries in use and in context within the BNC.

For determining usage frequencies at least 200 instances of each for CAN and COULD, including affirmative and negative contexts in spoken and written mode, were collected from the BNC, by way of random sampling. Support for this amount of tokens comes from Römer (2004: 186) who analysed 200 instances of modal auxiliaries in the spoken component of the BNC. Table 2 provides a breakdown of the forms and modes collected.

In the table above, 100 instances of each form were collected, with results for *can* including *can not* and *cannot*, and results for *could* including *could not* and *couldn't*. The negative form *couldn't* was not collected separately, which is different from the case with *can't*, as *couldn't* is included in the search for *could* by design; the search function of the BNC did not allow for a search for *could* only. However, this raises the issue of how many instances of *could* and *couldn't* occur in my sample sets compared to the overall BNC. Table 3 below shows that these were very similar.

I used Fisher’s Exact (FE) test to test for significance. In all four cases I found that there are no significant differences. The data set in this study thus contained a

total of 400 tokens of CAN and 200 tokens of COULD. Utterances containing the tokens were uploaded into the concordance software package WordSmith Tools 6.0 (Scott 2012).

## Expanded Context

Context played a significant role in analysing instances, as well as reporting instances. The issue regarding the amount of context used to help understand an instance's use will be discussed first, while the matter of reporting context is addressed further below. The analysis of modal tokens in this study began with 100 words of surrounding context; however for the majority of tokens, additional context beyond that (usually between 150 and 200 words), was required in order to reach a more confident classification. Outside this expanded context, I also utilised the BNCweb's audio component, which was available for some of the spoken instances (see the discussion of example (10) below). Examining further context in many cases allowed for a more robust usage analysis, as illustrated in (1) below.

In example (1), the italicised words are those that were not included in the initial 100 words of context. The context was expanded to include the italicised words in order to arrive at a more confident categorisation of this instance of *could*.

- (1) [A] *Social conditions certainly affected Mrs. Daisy Sawyer's choice of furnishing. She remembers setting up her first home, right after the war.*
- [B] *They were docketed that we had, after the war, to buy our furniture, because there wasn't much furniture around, we was only allowed so many per family. And once you spent those docketed, you just had to go and buy secondhand if you wanted any more. There's quite a few around us because then we was not going onto a council estate from the one room, and we were all in the same boat together. We was all having a hard time, a rough time, and doing what we **could** to make our homes look respectable and nice for people to come into. (spoken BNC, radio broadcast, 1985–1994)*

In the instance above, without the expanded context, 'doing what we could' can be read as either having an 'ability' or 'external possibility' reading (see "[Ability](#)" section and "[External Possibility](#)" section below). In an 'ability' reading, the subjects are doing everything they are internally capable of, and in an 'external possibility' reading, the subjects are doing everything that external circumstances allowed. The additional context in italics favours an 'external possibility' reading with support coming from the context, 'Social conditions certainly affected Mrs. Daisy Sawyer's choice of furnishing'. The majority of BNC instances in the study are similar to the above in that they required an abundant amount of context in order to arrive at a more convincing analysis.

Since modal auxiliaries are largely used to express speaker's stance, I believe there is an extra responsibility on linguists to include expanded context to show each instance with as much context as possible not only in their analysis but also in their reporting. When reading previous studies, what I found most striking, and often frustrating, was the lack of access I had to the context around an instance as the majority of these studies report instances with only limited context. Though the

importance of context is unmistakably acknowledged in literature (see also Leech and Coates 1980), linguists still often present extracts of one sentence or shorter.

In Biber et al. (1999: 492), ‘*He goes, I can’t swim*’ is given as an example of ‘ability’. This is taken from the BNC, which allowed me to search for further context around it. While it is the case that the expanded context for this particular instance supports an ‘ability’ reading, when looking at the other instances of ‘*can’t swim*’ in the BNC, not all are ‘ability’. For example, the instance of ‘*you can’t swim freely when you get tangled up in roots*’ (BNC, *The Art Newspaper*, 1985-1994) reflects ‘external possibility’ (see “[External Possibility](#)” section below), as being ‘tangled up in roots’ makes it not possible to swim. Often times, with expanded context included, an instance’s use can be made clearer to readers. The importance of reporting expanded context for example instances is thus crucial, and looking beyond sentence level is also emphasised in Larsen-Freeman’s (2003: 67) work. Because of the issues above, context beyond the concordance lines has been used for analysis in this study, and, where necessary, examples in this study are reported with an expanded amount of context in an effort to provide the fullest representation possible.

## Categories of Use for *CAN* and *COULD*

From examining instances of *CAN* and *COULD* in the BNC, the following usage categories (semantic and pragmatic) were established: ‘ability’, ‘external possibility’, ‘permission’, ‘epistemic possibility’, ‘directive’ or ‘commissive’, ‘volition’ and ‘phrase’, together with ‘ambiguous’ and ‘unclear’ categories. Each category and its criteria are explained below, including the ‘linguistic substitution check’, adapted from Yin’s (2014: 91) ‘linguistic substitution test’ and Hermerén’s (1978: 89) ‘paraphrases’. This check serves multiple purposes: it aids the researcher in maintaining consistency when classifying tokens and bridges the gap between the researcher and the reader in facilitating a mutual understanding of usage categories. Furthermore, each usage category is supported with a BNC example.

### Ability

Going by Table 1, this category is the most consistent in its label, yet as discussed in “[Previous Corpus-Based Investigations of \*CAN\* and \*COULD\*](#)” section above, and further elaborated in this section, the criteria applied in the existing literature vary.

The criteria used for ‘ability’ in this study are given below. Criteria (a) and (b) have been adapted from Coates (1983: 89).

- (a) Subject is animate (or indirectly animate);
- (b) The possibility of the action is determined by the internal competence of the subject (at the moment of action);
- (c) Linguistic substitution check: *be capable of* (Quirk et al. 1985: 222; Leech 2004: 74).

Example (2) is an example of *can't* in 'ability' use.

- (2) [A] and er, she said he's a bit a (pause) she said he's suffering from Alzheimer's disease  
 [B] mm  
 [A] and half way through a composition he *can't* remember who he's talking to and things like that  
 [B] mm (spoken BNC, conversation, 1992)

Instance (2) meets the above 'ability' criteria: (a) 'he' is animate; (b) the possibility of 'remembering' is determined by the internal competence of the subject (at the moment of action); (c) linguistic substitution check: *...he is not capable of remembering who he's talking to and things like that.*

There are two main differences between the criteria for 'ability' in this study, compared to previous studies. The first difference is what qualifies as an 'animate subject', which is part of criterion (a). The present study takes the view that an 'animate' subject 'has life' (Sinclair 2006: 48), and conversely views an 'inanimate' subject as 'one that has no life' (Sinclair 2006: 731). However, many linguists (Ehrman 1966; Coates 1983; Palmer 1990; Facchinetti 2002; Collins 2009) include instances with inanimate subjects in their 'ability' category. For example, Coates (1983: 92) includes the subject *tape recorder*, '*The plane has a built in stereo tape recorder which can play for the whole 4 h it will take to fly to Majorca*', in her ability category on the grounds that its ability is 'inherent'. The present study excludes inanimate subjects, as allowing some inanimate subjects to be included would add an additional layer of subjectivity. In this study, instances with inanimate subjects are categorised as 'external possibility' (see "[External Possibility](#)" section below). With regard to 'indirectly animate' subjects which is also part of criterion (a), these are those which are internal to an animate subject and cannot be separated from having life (e.g. ego). An example from my data set using 'ego' is:

... this is why some people erm (pause) presumably er feel better in groups, perhaps that they get something out of a group that their own ego *can* not provide, ... (spoken BNC, London School of Economics, 1885–1994)

The other difference in my criteria is the inclusion of 'at the moment of action' in (b). In this study, 'at the moment of action' plays a significant role in the analysis of instances, affecting those assigned to 'ability' that would have otherwise been analysed as 'external possibility' (see "[External Possibility](#)" section below). Without including 'at the moment of action' in the criterion, instances such as (2) above could be interpreted as 'external possibility' on the basis that 'Alzheimer's' is an external circumstance. However, taking 'at the moment of the action' into consideration allows readers to interpret this as *in this particular instance, the subject is not capable of remembering, but there are other times when the subject is capable of remembering.* This creates a focus on the internal competence of the subject at the time of speaking or writing, and allows for abilities to change.



## External Possibility

Often referred to as only ‘possibility’, this study found a need to include ‘external’ as reasons for ‘possibility’ may be ability, permission or external circumstances, which are the focus of this section. In instances of ‘external possibility’, the possibility of the action is dependent upon external circumstances and not upon the internal capability of the subject (‘ability’) or permission from ‘human authority/rules and regulations’ (‘permission’) (see “[Ability](#)” section above and “[Permission](#)” section below, respectively). The criteria used for ‘external possibility’ are:

- (a) The ‘possibility’ of the action is dependent upon external circumstances;
- (b) Linguistic substitution check: *Due to external circumstances, it is possible for x to...* (Hermerén 1978; Leech 2004; Collins 2009)

Instance (3) is an example of an ‘external possibility’ use.

- (3) It says video or a maximum 640 by 480 pixel rectangular region of the screen can be sent in near real time. Users *can* adjust frame speed, image quality and window size, giving control over how much data is sent over the network. (written BNC, Computergram International, 1985–1994)

In instance (3), the subject is human, and applying the criteria for ‘external possibility’ to this instance: (a) the possibility of adjusting the frame speed, image quality and window size is dependent upon the functionality of the video conferencing program; and (b) linguistic substitution check: *due to the functionality of the video conferencing program, it is possible for users to adjust frame speed, image quality and window size...* ‘Be capable of’ (‘ability’) is a less likely interpretation as this instance is not focusing on the internal competence of the subject to operate the program. And ‘permitted’ (‘permission’) is not plausible because this instance is not focusing on the subject being allowed to operate the program. It is clear in this instance that it is not the case that the writer is referring to users being capable of adjusting, or being allowed to, it is the case that the program facilitates the adjustments being made.

As discussed in the previous section, inanimate objects would meet the criteria for ‘external possibility’, as in the example provided below:

- (4) Virtual PreVUE *can* emulate a system with thousands of users, PreVUE-X tests for bugs in X-Windows applications. (written BNC, Unigram, 1985–1994)

In (4): (a) the possibility of emulating a system with thousands of users is dependent upon the functionality of the Virtual PreVUE system; and (b) linguistic substitution check: *due to the functionality of the system, it is possible for PreVUE to emulate a system with thousands of users.*

## Permission

Of the usage categories identified, ‘permission’ seems to me to be the most straightforward. Coates (1983: 88) highlights how close the ‘permission’ and ‘possibility’ meanings are for CAN and states, “there is no non-arbitrary way to draw the line between ‘Permission’ and ‘Possibility’”. Leech (2004: 73) agrees and conveys this by creating a diagram in which there is a two-way arrow from ‘possibility’ to ‘permission’. However, for my own analysis, I created distinct sets of criteria for ‘permission’ and ‘external possibility’ (see “Ability” section above), which help to disambiguate these two uses.

The criteria used for ‘permission’ are as follows:

- (a) Subject is animate (Coates 1983: 87);
- (b) *x* receives (or has) permission from human authority/rules and regulations to perform *y* (adapted from Coates 1983: 87–88);
- (c) Linguistic substitution check: *BE + permitted* (Hermerén 1978; Coates 1983).

In instances of ‘permission,’ the subject seeking or given permission is ‘animate’, with ‘animate’ reflecting the same quality as discussed above in “Ability” section, ‘has life’. Coates (1983: 88) also found that the instances in her permission category contained animate subjects. In criterion (b), the first part of this explanation, ‘*x* receives (or has) permission from human authority/rules and regulations’ was influenced by Coates (1983: 87–88) and the second part, ‘to perform *y*’, was added for clarity. Instance (5) is an example of a ‘permission’ use.

- (5) A person under 16 can not consent to surgery but the parent *can* on his behalf. (written BNC, *Criminal Law*, 1992)

Instance (5) above meets the ‘permission’ criteria as follows: (a) ‘parent’ is animate; (b) ‘parent’ is granted permission from rules/regulations to ‘consent to surgery’; and (c) linguistic substitution check: *A person under 16 cannot consent to surgery but the parent is permitted to on his behalf.*

## Epistemic Possibility

While interpretations of the epistemic use of CAN have changed over the past few decades, what remains constant, and current, is that linguists have various views on epistemic CAN. Huddleston and Pullum (2002) assert that CAN can be used epistemically but ‘restricted to non-affirmative contexts’. In contrast, Coates (1995) and Collins (2009) provide examples in their work where CAN is used in epistemically in affirmative contexts. Coates’ (1995: 63) example is from a Symposium, ‘*we hope this coding system can be useful* [to other linguists working in the field]’, and Collins (2009: 99) notes that though Coates believes this

affirmative use is found only in American English, he has found otherwise, as in an example from his ICE-GB data set, ‘I think they *can* have but I’m not sure’.<sup>2</sup>

This study considers ‘epistemic possibility’ to be a speaker’s, or writer’s, conclusion of certainty concerning the factuality of a past, present or future situation based on logical probability; this degree of certainty ranges from ‘certain’ to ‘possible’ (Holmes 1983; Celce-Murcia and Larsen-Freeman 1999; Papafragou 2000b; Kennedy 2003; Collins 2009). The criteria used for ‘epistemic possibility’ are as follows:

- (a) Speaker’s, or writer’s, level of certainty towards a situation (Holmes 1983; Celce-Murcia and Larsen-Freeman 1999; Kennedy 2003; Collins 2009);
- (b) Linguistic substitution check(s):
  - (1) *it is possible that* (Hermerén 1978; Sweetser 1982; Coates 1983; Palmer 1990; Leech 2004; Collins 2009)
  - (2) *it is certain that* (when *can’t* supplies the negative form of epistemic MUST) (Coates 1983; Holmes 1983).

Instance (6) below is an example of the speaker expressing his or her degree of certainty at a level of ‘possible’, while in instance (7), the speaker is expressing his or her degree of certainty at a level of ‘certain’.

- (6) It seems likely that our dreams are attempts to make sense of experiences which are vivid but inconsequential (as suggested by John Hughlings Jackson in his Perceptual Release Theory) and this will be dealt with later in the chapter. Maury’s dream *could* be a good example of this sort of attempt at comprehension of a series of striking but unconnected images and sensations, a form of the “effort after meaning” that the Cambridge psychologist Sir Frederick Bartlett ascribed to normal waking memory processes in the first half of this century. (written BNC, *Sleep and dreaming*, 1989)

Instance (6) is analysed as ‘epistemic possibility’ in that the speaker is stating his or her level of certainty, which is ‘possible’ (as opposed to ‘certain’), that *Maury’s dream is a good example of this sort of attempt at comprehension*. The linguistic substitution check for this is: *It is possible that Maury’s dream is a good example of this sort of attempt at comprehension*.

- (7) I don’t think you could afford to do so. You told me you have to manage on what you earn—dashed bad luck, I know—but for your own sake, you need to face facts. However well they pay chaps like you to look after horses, it *can’t* be all that much! Do you have any idea how extravagant that sister of mine is? The pater gives her fifty pounds a year as a dress allowance and she has nearly always spent it by the end of the first month. (written BNC, *The Spinning Wheel*, 1993)

Instance (7) is classified as ‘epistemic certainty’: (a) the speaker in the written story is expressing his or her level of certainty (‘certain’) that the listener does not get paid that much; and (b) linguistic substitution check: *it is certain that it is not all*

<sup>2</sup> A reviewer of this paper notes a ‘dynamic’ interpretation (not ‘epistemic’) of the example instances from Coates (1983) and Collins (2009).

*that much...* Contextual support comes from ‘I don’t think you could afford to do so’. Though not shown in the BNC, in BNCweb (*The British National Corpus, Version 3 (BNC XML Edition) 2007*), ‘afford’ is italicised for emphasis. Additionally, the context ‘you told me you have to manage on what you earn—dashed bad luck’, indicates, as readers can assume, that ‘what you earn’ is not very much.

While the previous usage categories—‘ability’, ‘external possibility’, ‘permission’ and ‘epistemic possibility’—were semantically motivated, the next three usage categories—‘directive or commissive’, ‘volition’ and ‘phrase’—are pragmatically motivated.

### Directive or Commissive

During the analysis of modals in the BNC, there were instances where the modals were performing another sort of job, different from the categories presented above, which led me to include the usage category ‘directives’ or ‘commissives’. Though the instances below carry modality in that the speaker is framing them in terms of ability and possibility, they do not fit the same substitution checks as their other ‘more modal’ counterparts. According to Searle (1979: 13), ‘the illocutionary point of these [directives] consists in the fact that they are attempts [...] by the speaker to get the hearer to do something’. He continues to explain that these ‘attempts’ range from suggesting to insisting. Also included in this are commissives, which are ‘those illocutionary acts whose point is to commit the speaker [...] to some future course of action’ (Searle 1979: 14).

The criterion for ‘directive’ and ‘commissive’ consists of linguistic substitution checks as follows:

- (a) Linguistic substitution check: *I want* subject to (directive) and *I intend to* (commissive)

Instance (8) is an example of a ‘directive’ and (9) is an example of a ‘commissive’.

(8) *Could* you put you’re [sic] head back a bit, move! (spoken BNC, conversation, 1991)

- (9) [A] We’ll drop her back.  
 [B] Well we can, we *can*  
 [A] Unless you want her earlier  
 [B] pick her up about seven.  
 [A] yeah that’s fine. (spoken BNC, conversation, 1991)

Instance (8) is a ‘directive’ in the form of a request, and instance (9), a ‘commissive’ in which the speaker commits to ‘pick her up’. The linguistic substitution checks for these are, respectively, *I want you to put your head back a bit* and *I intend to pick her up*.

## Volition

In my BNC data, I found instances where the meaning was more than possibility, the speaker or writer was stating personal desire, or ‘volition’. The speaker or writer is stating what he or she wants. Though it is common for linguists to include a ‘volition’ category connected to modal auxiliary uses, this category is usually associated with *will*, *would* and *shall*, not CAN and COULD as found in this study. The criterion for classifying an instance as ‘volition’ is:

- (a) Linguistic substitution check: *want to...*

The following instance is a form of social hedging. Speaker A is a young girl and Speaker B is an adult male. In the instance below, the speaker is taking advantage of the impreciseness of *can't* to communicate his preference.

- (10) [A] Hold me up.

[B] Oh no I can't Katie, I *can't*, you're too heavy. Up you get, ooh you're a big lump now, you are getting a big girl. (spoken BNC, conversation, 1992)

In (10), the young girl is requesting to be picked up and Speaker B says, ‘I can't, you're too heavy’, followed immediately by ‘up you get’. The ‘up you get’ is the speaker picking the girl up, which is audible on the recording from BNC*web*. This instance is volitional in the sense that Speaker A does not really want to pick the girl up; it is not a question of the speaker's ‘ability’, ‘external possibility’, or ‘permission’ constraints. The linguistic substitution check for this instance is: *I don't want to [pick you up]*.

## Phrase

In the process of this analysis, I came across various instances that included the modal auxiliaries in this study, but similarly to the ‘directives’ or ‘commissives’ and ‘volition’ categories, they did not fit the linguistic substitution checks, and seemed to have different meanings altogether. Sometimes referred to as *idioms* or *idiomatic expressions*, this study prefers the term ‘phrase’ for this group of instances. ‘Phrase’ comes from Sinclair (2006: xviii) who describes ‘phrases’ as ‘groups of words which are used together with little variation and which have a use of their own’. The criteria for classifying an instance as a ‘phrase’ are:

- (a) The modal auxiliary + verb create a new meaning (e.g. ‘can't say’ = *don't know*);  
 (b) The verb meaning does not occur without the relevant modal. (e.g. the meaning *don't know* from ‘can't say’ is not conveyed with ‘say’ only).

Instance (11) is an example of a ‘phrase’ from the BNC.

- (11) [A] It's just (pause) and it's so, it's like it was raining the other day when we went to go out (pause) and just *could* not be bothered. It was really  
 [B] Mm.  
 [A] pelting down. Oh it was when we were gonna go to the cinema and I phoned Marion and I said no, I really cannot be bothered, it's just pouring down with rain, had to (spoken BNC, conversation, 1993)

In (11), 'could not be bothered' can be paraphrased as *I did not want to*. Support for this comes from Sinclair (2006: 155), who describes, 'If you say that you **can't be bothered** to do something, you mean that you are not going to do it because you think it is unnecessary or because you are too lazy'. Other phrases found in my data include *can't help*, *can't bear*, *can imagine* and *couldn't believe*.

### Ambiguous and Unclear

Palmer (1990: 22) claims that ambiguity occurs when, 'it is not possible to decide in a particular context between two possible meanings of a form' and only one applies to the situation. This needs to be distinguished from what this study calls 'unclear' instances, where 'no firm decision could, even in principle, be made' (ibid). From the BNC data, an examples of an 'ambiguous' instance is shown in (12), and an example of an 'unclear' instance is shown in (13).

- (12) [A] after the first fast bend he went to overtake and I could see a car coming the other way and he was like er running along side this pick-up but it was like he couldn't drop back and he *couldn't* make it either  
 [B] Mm  
 [A] so he just kept going, so the pick-up jammed its anchors on, there was a Rover it was coming the other way, the Rover went up the verge (spoken BNC, conversation, 1985-1994)

Instance (12) is 'ambiguous' between an 'ability' and 'external possibility' reading. In an 'ability' reading: (a) he is animate; (b) the possibility of making it was determined by the internal competence of the subject; and (c) linguistic substitution check: *...he couldn't drop back and he was not capable of making it either*. In this reading, 'he' couldn't make it due to his internal capability (at the moment of action). For example, it may be his mental capabilities holding him back. In an 'external possibility' reading: (a) the possibility of making it was dependent upon external circumstances (e.g. speed capacity of vehicle); and (b) linguistic substitution check: *Due to external circumstances (e.g. the speed capacity of vehicle), it was not possible for him to make it*. This instance does not have enough context to determine one reading over another. More context or outside information would be required to determine to which usage category it belongs.

Instance (13) is an example of an 'unclear' instance.

- (13) [A] Its not worth it for you is it, you might have to hang on till after Christmas and try and (unclear)  
 [B] I might phone them up and just, ask them if there's any jobs going after Christmas sort of thing (pause)  
 [A] *Can* use the good (unclear) the things, I mean, for a, for us who haven't got permanent jobs yet (spoken BNC, conversation, 1991)

It is natural for humans to want to understand the messages others convey. Halliday and Hasan (2014: 54) discuss how a hearer or reader 'will go to enormous lengths to interpret' a text. However, however, the instance above does not convey any meaningful reading. Readers could add their own context to (13) and assume that the speaker is permitting the hearer to use the 'good' *x*, or in another scenario that it is possible to use the 'good' *x*, but these are just fabricated scenarios as there is not enough context available in the corpus to analyse its use.

Having identified the main usage categories for CAN and COULD of 'ability', 'external possibility', 'permission', 'epistemic possibility', 'directive' or 'commissive', 'volition' and 'phrase', as found in the BNC, I now move on to the analysis of usage frequencies.

## Usage Frequency Findings

This section reports on the frequencies of use for spoken and written CAN and COULD for all of the usage categories identified in "[Categories of Use for CAN and COULD](#)" section. The figures in Tables 4 and 5 below show the usage frequency counts for spoken and written CAN and COULD in the BNC. In Table 4 spoken *can* and spoken *can't* are combined into spoken CAN, and written *can* and written *can't* are combined into written CAN. I do this for the following reasons: (a) other corpus-based studies (e.g. Coates 1983; Collins 2009; Römer 2004) that I reference in comparison to my own work combine *can* and *can't* into the modal category of CAN; and (b) *could* and *couldn't* are combined already as per the search parameters in the BNC. Further below, Table 6 reflects the combined spoken and written data for CAN and COULD.

For spoken and written CAN, this study confirms, in line with the majority of other studies, that 'external possibility' is the most frequent usage. Where my usage findings diverged for CAN is that most often 'ability' is the second most frequent use reported; however, in the spoken context for CAN I found the 'directive'/'commissive' category to be the second most frequent, which is different from the written context, in which it was the fourth most frequent. Also, where other studies have found 'permission' to be the third most frequent use with CAN, I found it to be the third most frequent in written context but not spoken, where it was fifth. This shows that in these higher frequency usage categories, the spoken and written registers play a major role.

Table 5 below shows the usage frequency counts for spoken and written COULD in the BNC. There was no need to weight the data for COULD since *could* and *couldn't* forms were already combined per the search functionality in the BNC.

**Table 4** Usage frequency counts for spoken and written CAN in the BNC

Category of use	Spoken CAN (BNC <sup>w</sup> )		Written CAN (BNC <sup>w</sup> )	
	Count	% within data set (%)	Count	% within data set (%)
External possibility	123	61.4	148	73.8
Ability	14	7.1	26	13.0
Directive/commissive	22	11.2	6	2.8
Phrase	17	8.5	4	2.0
Permission	10	4.8	10	4.8
Volition	1	0.5	2	1.2
Epistemic possibility	1	0.7	1	0.3
Ambiguous	2	1.0	2	1.1
Unclear	10	4.8	2	1.0
Total	200	100	201	100

The superscript 'w' in BNC<sup>w</sup> indicates that the sample data sets from the BNC are weighted. For detailed weighting calculations, see Electronic Supplementary Materials

**Table 5** Usage frequency counts for spoken and written COULD in the BNC

Category of use	Spoken COULD BNC <sup>s</sup>		Written COULD BNC <sup>s</sup>	
	Count	% within data set	Count	% within data set
External possibility	42	42%	57	57%
Epistemic possibility	20	20%	13	13%
Directive/commissive	18	18%	0	0
Ability	8	8%	13	13%
Phrase	3	3%	6	6%
Permission	0	0	4	4%
Volition	2	2%	1	1%
Ambiguous	2	2%	6	6%
Unclear	5	5%	0	0
Total	100	100%	100	100%

The superscript 's' in BNC<sup>s</sup> indicates that the data percentages from the BNC are taken from a sample

For spoken and written COULD, similar to CAN, the spoken and written contexts have an impact on frequency counts. 'External possibility' is the most frequent usage category used, followed by 'epistemic possibility'. For spoken COULD, 'directive'/'commissive' is the third most frequent, and for written COULD, 'ability' is the third most frequent. 'Epistemic possibility' as the second most frequent use is a noticeable difference from CAN, where 'epistemic possibility' was the least



**Table 6** Usage frequency counts for CAN and COULD in the BNC

Category of use	CAN (BNC <sup>w</sup> )		COULD (BNC <sup>w</sup> )	
	count	% within data set (%)	count	% within data set (%)
External possibility	286	71.4	110	55.1
Ability	47	11.8	25	12.4
Epistemic possibility	2	0.4	28	13.9
Phrase	13	3.3	11	5.6
Directive/commissive	18	4.5	5	2.3
Permission	19	4.8	7	3.5
Volition	4	1.0	1	1.1
Ambiguous	4	1.1	11	5.5
Unclear	7	1.7	1	0.6
Total	400	100.0	199	100.0

frequently used. ‘Directive’/‘commissive’ shows the greatest difference between spoken and written use, where spoken COULD has an 18% frequency but for written COULD it is zero. COULD is used this way in spoken texts to “soften” requests (Celce-Murcia and Larsen-Freeman 1999: 145).

In order to provide an overall comparison of CAN and COULD to previous studies, Table 6 combines the written and spoken data.<sup>3</sup> Table 6 will be used for comparison to other studies in the discussion in the following section.

### Traditional Categories

In the following subsections, I discuss the four categories—‘external possibility’, ‘ability’, ‘epistemic possibility’ and ‘permission’—which, in title, are similar to the categories found in the previous studies. I focus on the differences in frequency findings in comparison to the previous studies and the possible reasons for these differences.

### External Possibility

While most corpus-based investigations of CAN (e.g. Hermerén 1978; Coates 1983; Mindt 1995; Leech et al. 2001; Collins 2009) found the ‘possibility’ use (in its broadest sense) to be the most frequent, others report different findings (e.g. Bald 1990; Biber et al. 1999; Facchinetti 2002; Römer 2004). When examining the London-Lund Corpus of spoken English (also known as the Survey in Coates’ work), Bald (1990: 354) included CAN in those modal auxiliaries that ‘dominantly occur as epistemic’. In the LSWE corpus, Biber et al. (1999) found equal use between

<sup>3</sup> For detailed weighting calculations, see Electronic Supplementary Materials.

what they call ‘extrinsic-possibility’ and ‘ability’ in conversation and *can* used more for ‘ability’ than ‘extrinsic-possibility’ in academic contexts.

While Bald (1990) and Facchinetti (2002) found ‘epistemic possibility’ to be the most frequent meaning for COULD, other corpus-based studies found ‘possibility’ to be most frequent (Coates 1983; Mindt 1995; Biber et al. 1999; Römer 2004; Collins 2009). Though external possibility predominates in the present study, there are differences to other studies in the category percentages, which I believe can be linked to linguists’ differing criteria for each meaning usage category, as well as the different corpora used.

In this study, ‘external possibility’ was not only the most frequent usage category, but also significantly more frequent compared to the other usage categories, which was not the case in previous studies. Due to the varying classifications and criteria for categories, it is not possible to carry out more detailed comparisons with the previous studies that found this to be the most frequently used category in their respective corpora. For example, Collins (2009) reports findings for his ‘dynamic’ category, which includes, ‘external possibility’, ‘ability’ and ‘directive’, while Mindt (1995) uses the categories ‘possibility/high probability’, with no clear distinction between the two.

A closer comparison can be made to Römer (2004: 188), who analysed the BNC spoken corpus and found ‘possibility’ to be the second most frequent category, and provides a figure of 31.5%. In the spoken part of my data set, I found a weighted usage frequency of 61.4% for external possibility CAN (66% for spoken *can* and 48% for spoken *can’t*), which is nearly double that of Römer’s figure. Again, this is most likely due to the different parameters in criteria for classification.

## Ability

In Table 6, we see that ‘ability’ was the second most frequent usage category for CAN, and the third most frequent for COULD; however, the percentage differences are far below those for ‘external possibility’, with ‘ability’ CAN at 11.8% and ‘ability’ COULD at 12.4%. These percentages are much lower than found in previous studies.

For CAN, Römer (2004: 188) found ‘ability’ to be the most frequent category in her study on the spoken part of the BNC, at 36%, yet I found the BNC spoken portion of my data to have a weighted frequency use of 7.1% for ‘ability’ CAN (3% for spoken *can* and 19% for spoken *can’t*). Using the same spoken data set, Römer (2004: 189) reported the ‘ability’ use for COULD at 34%, whereas I found a frequency of 8%. These are both remarkable differences.’ Mindt (1995: 81) also reports quite a high percentage of ‘ability’ instances for COULD. This may be due to his inclusion of, for example, the following instance of ‘ability’: ‘*with the driver’s seat pushed fully back, I could almost straighten my legs*’. This would be classified as ‘external possibility’ in this study, rather than as expressing the internal competence of the subject, with the external circumstances being the positioning of the driver’s chair.

## Epistemic Possibility

Linguists vary on whether or not they include an ‘epistemic’ category for CAN. While Facchinetti (2002) and Collins (2009) do, Coates (1983, 1995) and Römer (2004) do not. Coates’ (1983: 102) exclusion of an epistemic category for *can* may be due to the scarceness of examples found, only one in all her data; however, she does claim that ‘the meaning of examples with Epistemic *can*’t corresponds to that of Epistemic MUST’. In Coates’ (1995: 63) later work, she includes an example of ‘epistemic’ *can*, but does not include the modal auxiliary *can* in her ‘epistemic’ category.

With regard to ‘epistemic possibility’, in the present study, its frequency of use differs greatly between CAN and COULD, with this usage category least common for CAN, but the second most frequent for COULD. Some studies also report the category ‘epistemic possibility’ for COULD (Coates 1983; Facchinetti 2002; Collins 2009); however, in others, the category labels are different. For example, Mindt (1995: 83) uses the category ‘inference/deduction’ and though Römer (2004: 189) does not have an ‘epistemic’ category, she does have a category, similar to Mindt, titled ‘inference/deduction,’ which has a zero count for spoken COULD.

## Permission

Although ‘permission’ is often mentioned as one of the main meanings for CAN and COULD in discussions in the literature, frequencies are usually quite low, as they were in the current study. In both parts of their corpus, conversation and academic, Biber et al. (1999: 491) found ‘permission’ used less frequently than ‘extrinsic-possibility’ and ‘ability’, by about half in the conversation part of their corpus, and ‘rarely expressed in academic writing’. In the present study, I found permission’s weighted use to be 3.3% for CAN and 3.5% for COULD. This low frequency may be due to the lower frequency of the use of permission in everyday language as a whole, in contrast to the other uses. Given that ‘permission’ frequencies are generally reported as being low, or at least lower than other uses, it is surprising that it continues to be described as one of the central uses for CAN and COULD, along with ‘possibility’ and ‘ability’.

The frequency differences noted above between my study and others discussed are most likely due to varying parameters in categorical criteria, and may also be attributed to the amount of context considered in analysis. While analysis of different corpora is a part of corpus-based research and a part in which differences in findings are to be expected, when the criteria for each usage category are not explicitly stated, this makes it difficult for readers to understand a linguist’s reasoning for the inclusion of a token in a usage category. Without having further insight into other linguists’ criteria, the way in which I can contribute to helping readers better understand my own analysis is to support my category classifications with explanations of reasoning so that readers can understand my justification for instances included in each usage category.

## Additional Categories

In the following subsections, I discuss the categories—‘directive’ or ‘commissive’, ‘phrase’ and ‘volition’—which are outside those found in previous studies. One explanation for this is that perhaps in previous studies the categories were not present in the data sets, yet, more plausibly, it is most likely the case that the categories were present but the considerations for the analysis of modal instances were different and these were subsumed under larger semantic categories. Below, I focus on the need for an analysis of CAN and COULD to include these categories.

### Directive or Commissive

Coates (1983), Palmer (1990), Facchinetti (2002) and Collins (2009) recognise ‘directives’ in their studies, but do not categorise them as a main category of use. Coates (1983: 98) refers to them as ‘covert imperative[s]’, while Collins (2009: 104), Palmer (1990: 86) and Facchinetti (2002: 242) classify ‘directives’ as ‘dynamic implication’, with Facchinetti using their functions (e.g. suggestion, request) for identification. In all four studies, these meanings are subsumed under larger categories of meaning and not given due attention. Facchinetti (2002: 236, 239) first includes this category in ‘dynamic modality’ and it is only at the end of her paper that she introduces the term/category ‘implication’.

The BNC data in this study provide a frequency case for having a separate ‘directives’/‘commissives’ category of use. In my weighted data, 4.5% of instances of CAN and 2.3% instances of COULD were used as a ‘directive’ or ‘commissive’, and for spoken COULD alone, 18% of instances were used as a ‘directive’ or ‘commissive’. Furthermore, there is a functional case for having this category. ‘Directives’ and ‘commissives’ stand out because they call for a special response from the hearer/reader. With regard to directives in interrogative form, Palmer (1990: 191) states, “It would be perverse for the addressee to take them as simply questions about his ability or willingness to act and to reply ‘Yes’ but take no action”. In response to such a directive, the norm would be for the hearer to respond with a ‘yes’ or ‘no’. A ‘yes’ would be followed by action, and a ‘no’ would be followed by a reason why the hearer cannot perform *x*. Conversely, it would appear odd and/or rude to have someone say, for example, ‘I can help you’ (‘commissive’) and then walk away, leaving the purpose of the utterance to only state that he/she has the ‘ability’, ‘possibility’, or even ‘permission’ to help. The ‘directive’ and ‘commissive’ uses thus stand on their own in that the expected response is an action, or justification for a lack of action, and they should be recognised as such.

### Phrase

When the data for CAN and COULD are combined, the category ‘phrase’ is used more frequently than the ‘directive’ or ‘commissive’ and ‘permission’ categories. In my weighted data, I reported 3.3% of instances of CAN used as a ‘phrase’, with COULD at 5.6%. In my review of the literature and previous corpus-based studies,

I did not find any other linguist who used ‘phrase’ as a separate category in relation to the modal auxiliaries in this study, or any modal auxiliaries, nor was this type of utterance singled out. For example, Coates (1983: 90) includes ‘*can’t face*’ and ‘*can’t stand*’ in her ‘ability’ meaning category, and Hermerén (1978: 103) takes a similar approach in his analysis and includes an instance of ‘*could bear*’ in the category of ‘ability’, when the present study would have included these in the ‘phrase’ category. The relevance of assigning ‘phrases’ their own usage category becomes clearer when examining the difference between instance pairs (a)/(b).

- (a) But he *can’t* wait much longer. If you don’t accept by the end of the month then he’ll advertise. (BNC) (modal auxiliary)
- (b) I *can’t* wait to hear them. (BNC) (‘phrase’ = I am excited to)

### Volition

For the ‘volition’ usage frequency percentages, in the weighted data for CAN, 1% of instances were reported, and for COULD, 1.1%. Though percentages are low, I found it difficult to try to incorporate these uses into other categories, as *want to/don’t want to* is a very different meaning to possibility, and one that I felt should be acknowledged as such. From a communicative perspective, distinguishing the difference between a modal use and ‘volition’ helps us distinguish when someone really can’t do *x* and when someone does not want to do *x*. The ‘volition’ use is connected to social politeness and how speakers turn an offer down, as opposed to the direct truth of *I don’t want to*, which would most likely be perceived as rude.

### Conclusion

One of the most noticeable differences between the present study and previous studies on CAN and COULD is the difference in number of categories of use. By applying a semantic and pragmatic lens to these usage categories, this study distinguishes more categories. As demonstrated above, the categories of ‘directive’ or ‘commissive’, ‘phrase’ and ‘volition’ warrant their own categories as they represent distinct uses. This study was based on an analysis which included identifying clear criteria for each usage category and employing expanded context in the classification of tokens. Context plays a central role in distinguishing these uses of modal auxiliaries and this study suggests that it would be beneficial for anyone writing about modal auxiliaries to fully account for context when modal uses are being examined. A study which does not base findings on these considerations may not give the complete picture on how these modal auxiliaries are used. I am hopeful that the study reported here will assist in moving the study of modality forward by making category criteria as explicit as possible and highlighting the need for context to understand use. The study also contributes to corpus linguistics in advocating for linguists to show greater accountability through making their analyses more transparent.

**Acknowledgements** I am grateful to two anonymous reviewers for their helpful and insightful comments on an earlier draft of this paper. I would also like to express a special thanks to Dr. Elaine Vine, Victoria University of Wellington, for her work reviewing an even earlier version of this paper.

## Compliance with Ethical Standards

**Conflict of interest** The corresponding author states that there is no conflict of interest.

## References

- Bald, W.-D. (1990). Modal auxiliaries: Form and function in texts. *Anglistentag, Proceedings*, 348–361.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course, Second Edition* (2nd ed.). Boston, MA: Heinle & Heinle.
- Coates, J. (1983). *The semantics of the modal auxiliaries*. London: Croom Helm.
- Coates, J. (1995). The expression of root and epistemic possibility in English. In J. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse: Typological studies in language* (pp. 55–66). Amsterdam: John Benjamins.
- Collins, P. (2009). *Modals and quasi-modals in English*. Amsterdam: Rodopi.
- Davies, M. (2004). *BYU-BNC. (Based on the British National Corpus from Oxford University Press)*. Retrieved from <http://corpus.byu.edu/bnc/>.
- Ehrman, M. E. (1966). *The meanings of the modals in present-day American English*. The Hague: Mouton & Co.
- Facchinetti, R. (2002). Can and could in contemporary British English: a study of the ICE-GB corpus. In P. Peters, P. Collins, & A. Smith (Eds.), *New frontiers of corpus research: Papers from the twenty first international conference on English language research on computerized corpora Sydney 2000* (pp. 229–246). Amsterdam: Rodopi.
- Halliday, M. A. K., & Hasan, R. (2014). *Cohesion in English*. London: Routledge.
- Hermén, L. (1978). *On modality in English: A study of the semantics of the modals*. Lund: GWK Gleerup.
- Holmes, J. (1983). Speaking English with the appropriate degree of conviction. In C. Brumfit (Ed.), *Learning and teaching languages for communication: Applied linguistics perspectives* (pp. 100–113). London: Centre for Information on Language Teaching and Research.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Joos, M. (1964). *The English verb: Form and meaning*. Madison, WI: University of Wisconsin Press.
- Kennedy, G. (2003). *Structure and meaning in English: a guide for teachers*. Harlow: Pearson Education.
- Klinge, A. (1993). The English modal auxiliaries: From lexical semantics to utterance interpretation. *Journal of Linguistics*, 29(2), 315–357.
- Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston, MA: Heinle.
- Leech, G. (2004). *Meaning and the English verb* (3rd ed.). Harlow: Pearson Longman.
- Leech, G., & Coates, J. (1980). Semantic indeterminacy and the modals. In S. Greenbaum, G. Leech, & J. Svartvik (Eds.), *Studies in English linguistics for Randolph Quirk* (pp. 79–90). London: Longman.
- Leech, G., Cruickshank, B., & Ivanic, R. (2001a). *An A–Z of English grammar & usage* (2nd ed.). Harlow: Longman.
- Leech, G., Rayson, P., & Wilson, A. (2001b). *Word frequencies in written and spoken English*. Harlow: Pearson Education.
- Mindt, D. (1995). *An empirical grammar of the English verb: Modal verbs*. Berlin: Cornelsen.
- Palmer, F. R. (1990). *Modality and the English modals* (2nd ed.). London: Longman.
- Papafragou, A. (2000a). *Modality: Issues in the semantics-pragmatics interface*. Amsterdam: Elsevier.
- Papafragou, A. (2000b). On speech-act modality. *Journal of Pragmatics*, 32, 519–538.
- Perkins, M. R. (1983). *Modal expressions in English*. Noorwood, NJ: Ablex.

- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive grammar of the English language*. New York, NY: Longman.
- Römer, U. (2004). A corpus-driven approach to modal auxiliaries and their didactics. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 187–199). Amsterdam: John Benjamins.
- Scott, M. (2012). *WordSmith Tools 6.0* [Lexical Analysis Software]. Stroud.
- Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.
- Sinclair, J. (2006). *Collins Cobuild advanced learner's English dictionary* (5th ed.). Glasgow: HarperCollins.
- Sweetser, E. (1982). Root and epistemic modals: Causality in two worlds. In *Proceedings of the eighth annual meeting of the Berkeley Linguistics Society*, pp. 484–507.
- The British National Corpus, version 3 (BNC XML Edition). (2007). Retrieved from <http://www.natcorp.ox.ac.uk/>.
- Westney, P. (1995). *Modals and periphrastics in English*. Max Niemeyer Verlag GmbH & Co KG.
- Yin, Z. (2014). *Linking adverbials in English*. (Unpublished doctoral dissertation). Victoria University of Wellington.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.