

ORIGINAL ARTICLE

Sensitivity to syntactic violation and semantic ambiguity in English modal verbs: A self-paced reading study

Nadia Mifka-Profozic^{1*} , David O'Reilly¹  and Juan Guo²

¹University of York and ²Central University of Finance and Economics, Beijing

*Corresponding author. E-mail: nadia.mifka-profozic@york.ac.uk

(Received 12 April 2019; revised 20 April 2020; accepted 20 April 2020; first published online 03 September 2020)

Abstract

The present study is, to our knowledge, the first self-paced reading experiment to investigate the effects of syntactic violation and semantic ambiguity on processing English modal auxiliaries. Forty undergraduate students, native speakers of English, took part in the study and read 36 target sentences, each containing a modal verb in context. Two of the most frequent English modals, *can* and *may*, were used in three distinct categories of modal expression: agent-oriented/ability, epistemic possibility, and speaker-oriented/permission. The two modal auxiliaries were manipulated such that they were either congruent or incongruent with the context, or in the case of permission felicitous or infelicitous relative to the context. We found that incongruent modal use in an agent-oriented context resulted in a reading penalty that was observed in a spillover on the segments following the modal and the lexical verb. Incongruent modal use to express epistemic possibility significantly affected reading times immediately after the modal auxiliary, and also spilt over to the following segments. Reading times in sentences expressing speaker-oriented modality were not affected by inconsistency in the use of the modal verb unmarked for formality in a formal context. The substantive and methodological implications of findings are discussed.

Keywords: processing; self-paced reading; modal auxiliaries; epistemic; agent-oriented

Introduction

Human ability to express modal meaning plays a fundamental role in enabling people to comprehend and produce meanings that belong to “parallel worlds” (Traugott & Dasher, 2002) contrasting with pure facts and actuality. Such are the worlds of belief, knowledge, imagination, possibility, necessity, and inference. Being able to communicate the notions of possibility, probability, necessity, and inferred certainty, or being able to give and ask for permission, make requests, and so forth, makes communication among people considerably sophisticated and complex. Modality is a quality shared by all languages around the world (Bybee et al., 1994) but each language instantiates it in a different way. In

© The Author(s), 2020. Published by Cambridge University Press.

English, for example, modality is most frequently expressed by the use of modal auxiliaries (e.g., *can, may, could, might, shall, should, will, would, must*), or quasi-auxiliaries (e.g., *have to, ought to, need to, dare to, supposed to, had better*), but there are also modal adverbs (e.g., *perhaps, maybe, certainly*), adverbials, and modal clauses such as conditional clauses (linked by *if*), conditional concessive (linked by *even if*) and concessive clauses (linked by *although*).

Modality, and in particular modal auxiliaries in English, have been an area of intensive investigation over the past decades. Most research has been carried out in logic and in formal and functional linguistics (e.g., Huddleston, 1976; Lyons, 1977; Palmer, 1986, 2003), as well as in cognitive linguistics (Bybee, 1985; Bybee *et al.*, 1994; Portner, 2009; Talmy, 1988). However, even after years of fruitful research, modals still remain “one of the biggest problems for grammatical analysis” (Palmer, 2003, p. 1). By this statement it is acknowledged that modality is difficult to define because there are no firm syntactic and semantic criteria that could be used as defining criteria. Modals differ from other verbs in the English language on a number of characteristic features. For example, the acronym NICE, referring to the properties of *negation, inversion, code, and emphasis* (Huddleston, 1976, p. 333), points to the following distinctions: the *negation* of modal verbs is formed without the auxiliary *do*; the interrogative form of modal verbs uses *inversion* with the subject; the *code* refers to the modal behavior similar to that of auxiliaries, such as in “*She can swim and so can he.*” Also, similar to auxiliaries, modal verbs can have an emphatic form such as in “*You can do that,*” where the modal is pronounced with an *emphasis*. In addition, specific to the English language are three further characteristics that distinguish English modals from other English verbs: there is no “s” form for the third-person singular, there are no nonfinite forms (i.e., no infinitive, and neither present or past participle), and there is no co-occurrence of two modal verbs in a sentence.

Thus, it is not surprising that the peculiar nature of modal auxiliaries has stimulated a wide and perpetual interest in the research on modality. Most of the research conducted over the years, stretching from empirical and diachronic studies to theoretical description, has focused on language production (Coates, 2014; Facchinetti *et al.*, 2003; Palmer, 1986, 1990, 2003). Comprehension has rarely been studied, but this strand of investigation has resulted in remarkable and very important studies related to child acquisition of modality (e.g., Noveck 2001; Ozturk & Papafragou 2015; Papafragou & Ozturk, 2006). Furthermore, if the scope of modality is extended beyond the restrictions of syntactic and lexical categorization, then a series of self-paced reading (SPR) studies conducted by Stewart *et al.* (2009a), which examined online processing of conditional sentences, can be considered valuable contributions to the examination of online processing of modality. However, modal auxiliary verbs, as specific and in many respects unique representations of modal meaning, have so far received only minimal attention (e.g., Huette *et al.*, 2010) in research into online processing. The current study aims to address this gap by focusing on real-time processing of the modal auxiliaries *can* and *may*, when used in a piece of discourse.

The goal of this article is to investigate whether the choice of a modal verb within a sentence (*can* vs. *may*) affects the processing of that sentence. We are interested to see how this variation in the use of two verbs with respect to the context influences

online sentence processing. By addressing this issue, we hope to contribute the findings from online sentence processing using SPR to research on modality in the English language. More specifically, the question we ask is whether native speakers of the English language are sensitive to syntactic and semantic anomalies occurring in the use of modal auxiliaries when they express ability (*can*) and epistemic possibility (*may*). This is compared with the level of sensitivity to the use of the same modal verbs when exerting their performative function by expressing permission in a more or less formal context. We are interested in seeing whether the subjects show the same or dissimilar online behavior in these three conditions, that is, when encountering syntactic violation, semantic ambiguity, or infelicitous use of the verb in a particular context. We focus on two of the most frequent and occasionally interchangeable used English modal auxiliary verbs (*can* and *may*) because they present a dichotomy that allows us to detect anomalies if used in a mismatching context so that they can be interpreted as either syntactically or semantically ambiguous, as well as pragmatically infelicitous. We ask the following research questions:

1. How does a match and a mismatch between the context and the target agent-oriented modal verb expressing ability (*can*) affect reading times for target sentences?
2. How does a match and a mismatch in the use of the target modal verb expressing epistemic possibility (*may*) affect reading times for target sentences?
3. How does a match and a mismatch at the level of formality between the context and the target speaker-oriented modal verb expressing permission (*may*) affect reading times for target sentences?

We chose an SPR paradigm to examine the preceding questions because of its capacity to tap into cognitive processes (Keating & Jegerski, 2015; Marsden et al., 2018). SPR is used to elicit behavioral data, similarly to eye-tracking, but an advantage of SPR is that there is no option for readers to skip any word (Just et al., 1982). Moreover, SPR has already been successfully used in discourse studies (Stewart et al., 2009b; Haigh et al., 2013).

Within the wider aim of making further SPR studies on modality more transparent and robust, the current study had several methodological goals. First, in constructing our stimuli, we sought to estimate instrument reliability, a step often omitted with SPR stimuli (particularly with second language studies, Marsden et al., 2018). This allowed us to gauge how measurement consistency varied across instrument features such as modal verb type (Plonsky & Derrick, 2016). In estimating instrument reliability, we aimed to apply recent recommendations from the psychometrics literature, and gave careful consideration to the characteristics of our data when selecting a coefficient, in our case Revelle's Omega, which offered a more appropriate and robust alternative to the popular (but often misapplied) Cronbach's alpha. (For an overview of the restrictive assumptions of Cronbach's alpha, its misuse, and several recommended alternative coefficients, see McNeish, 2018; cf. Raykov & Marcoulides, 2019). For the main analysis, we again aimed to apply recent guidance on approaches to multivariate analysis (Plonsky & Oswald, 2017) and used mixed-effects regression. This method, in comparison to analysis of variance, allowed us to account for random subject and item variance

in a single analysis, avoided the need to average out observations for each participant, and offered an approach robust to violations of homoscedasticity and sphericity (Cunnings & Finlayson, 2015).

In what follows, we first present the types and sources of modality, then explain how we use the modal verbs *can* and *may* in three different contexts. Next, we refer to the psycholinguistic studies on modality, as well as to the corpus studies, and finally, to studies using SPR in investigating modality. This is followed by the methodology section and the presentation of the findings from a SPR task in which 40 English native speakers took part. We discuss our findings in the context of other psycholinguistic studies including those that investigated online processing of modality.

Background

Researchers approaching modality from different perspectives have used various terminology to classify modal verbs, but most often the differentiation has been along the lines of epistemic versus nonepistemic. This distinction can be most clearly explained by reference to the degree of certainty in one's knowledge (epistemic) on the one hand, and on the other, to the performative use or acts that one exercises to achieve a certain goal, or to point to someone's agentive quality (nonepistemic, root, or deontic). In other words, epistemic modality refers to the speaker's knowledge, beliefs, or opinion, whereas deontic/root modality concerns the situations that involve obligation or compulsion, using "language as action" (Palmer, 1986, p. 121), or indicates the absence of constraints on events and states (Traugott & Dasher, 2002). By this it is meant that there are no obstacles for an action to be completed, such as, for example, in (1).

The following two examples with *may* illustrate the difference between its epistemic and deontic meaning:

- (1) *If you wish to talk to the manager you **may** do so* [= you are allowed to talk to the manager]—permission, deontic meaning
- (2) *She **may** have already left* [= it is possible that she has already left]—epistemic meaning

Types and sources of modality

Research so far has confirmed that modality, both diachronically and in individuals, has the development from root/deontic to epistemic but not vice versa (Traugott & Dasher, 2002). The developmental tendencies indicate that modal meanings shift from externally described situations toward the meanings based on the internally (evaluative, perceptual, cognitive) described situations (Traugott, 1989). Such tendencies are important to consider in understanding the differences between the root/deontic and epistemic modal meanings.

Depending on the source of modality, Bybee *et al.* (1994) classified English modal verbs into four types: those that are agent-oriented, speaker-oriented, epistemic, and subordinate. Agent-oriented modality deals with the "internal and external

conditions on an agent with respect to the completion of the action expressed in the main predicate” (Bybee et al., 1994, p. 177). This type of modality includes obligation (*must*), necessity (*need*), ability (*can*), and desire (*would*). Speaker-oriented modality is not about the existence of conditions on the agent “but rather allow[s] the speaker to impose such conditions on the addressee” (Bybee et al., 1994, p. 179). Speaker-oriented modal verbs are used in all directives such as commands, demands, requests, entreaties, warnings, exhortations, and recommendations. In Bybee et al.’s (1994) terms, these can be imperatives, prohibitives, optatives, hortatives, admonitives, and permissive. The third type, epistemic modality, refers to those assertions that indicate the “extent to which the speaker is committed to the truth of the proposition” (p. 179). The commonly expressed epistemic modalities are possibility, probability, and inferred certainty. The fourth type, termed as “subordinating moods” involves the same forms that are used to express the speaker-oriented and epistemic modalities, to mark the verbs in certain types of subordinate clauses.

The current study, limiting its scope to the investigation of modals *can* and *may*, has adopted this classification by selecting one meaning for each of the first three categories and leaving out the last one—subordinate—as this type appears only in subordinate clauses, which are dependent on the main clause; therefore subordinate modals are always dependent on the truth of the proposition in the main clause and they are always the secondary source of modality.

We focus on ability as an agent-oriented type of modality, which “reports the existence of internal enabling conditions in the agent with respect to the predicate action” (Bybee et al., 1994, p. 177). An example is provided in (3).

(3) *She can type fast although she is a beginner.*

Epistemic modality is the second type we focus on. One of its meanings, which is expressed by the use of the verb *may*, is epistemic possibility, indicating that the proposition may possibly be true. An example is provided in (4)

(4) *He may have forgotten my address.*

Among speaker-oriented modalities, which allow the speaker to impose some conditions on the addressee, we focus on permissives by which the speaker grants permission to the addressee, as in (5).

(5) *You may use only the main entrance to the building.*

Pairing may and can

English modal verbs are usually considered in pairs with regard to their meaning (e.g., *can/could*, *may/might*, *shall/should*, *will/would*). The pair *can* and *may*, even though not linked by meaning, occur in a wide range of contexts, from their treatment in theoretical linguistics to textbooks for learners of English (Coates, 1980; Palmer, 1986). These two verbs are also linked by the negative notion of “no

obstacle” (Palmer, 1986), which can explain why they are sometimes interchangeable. If they are interchangeable, this can happen only in the root/deontic sense, but never in the epistemic sense (Coates, 2014; Palmer 1990).

Important for the current study is the fact that the modal *can* never encodes epistemic meaning, apart from being employed in statements expressing negation, such as in (6).

- (6) *He is a bachelor, so he can't be married.* [= *it is not possible* that he is married]

Can is used to express root possibility, meaning “it is possible for.” The difference between the two meanings, epistemic and root, when referring to possibility is that the use of *may* suggests it is the speaker’s belief, or their lower level of certainty about the truth of the proposition, as in (7), while the use of *can* does not involve the speaker’s belief or attitude and presents a more objective proposition, as in (8).

- (7) *This dog may be dangerous.* [= *it is possible* that this dog is dangerous/
perhaps it is dangerous]
(8) *Lightening can be dangerous.* [= *it is possible for* lightening to be dangerous/
perhaps cannot be used with sentences of root possibility]

Therefore, linguists have argued that the use of *can* in a context that suggests epistemicity will be perceived as ungrammatical (Palmer, 1986, 1990), as the example (9) shows.

- (9) *She *can have already left.*

Reciprocally, when *can* is used to express ability, conveying the meaning of one’s capability or capacity, the use of *may* is not perceived as acceptable. Due to the fact that both ability and permission are associated with animate subjects and agentivity, the context will condition the interpretation of the utterance containing the modal auxiliary. For example, the interpretation of the statement “*She can play the piano*” in (10) and (11) will depend on the context provided. Only in example (10) could the modal *can* be replaced with *may*.

- (10) *She can play the piano in this room*—permission (human authority/rules and regulations allow her to play in this room)
(11) *She can play the piano but she can't sing*—ability (inherent properties, i.e., her ability and talent allow her to play the piano)

When indicating permission, *may* and *can* are sometimes interchangeable, suggesting the use of *can* which is unmarked for formality in less formal environments, as in (12).

- (12) *You can stay here longer if you wish.*

The preceding examples demonstrate how the context shapes the inferences drawn from modal verbs and why it is crucial to interpreting modality. Indeed, many scholars (e.g., Bybee et al., 1994; Coates, 1980; Papafragou, 1998; Traugott & Dasher, 2002) consider modality an inherently pragmatic phenomenon and suggest that modals should be examined in discourse or “situation context.” This is what our study set out to do.

Usage and frequencies of can and may

Using two English language corpora from her own survey and from the Lancaster corpus, Coates (2014) found that the most common usage of *may* was to express epistemic possibility meaning “it is possible that,” while the secondary usage of *may* was to ask for or give permission, particularly in formal contexts. The use of *can* was found to be most frequently used to express root possibility, followed by ability and then permission.

In corpus research, several studies have identified a trend of changes in the use of modals, not only over several centuries, but over a few decades. For example, Leech (2003) compared the corpora of written and spoken British English (Lancaster–Oslo–Bergen corpus) collected in the 1960s with the same corpus from the 1990s, and pointed to a sharp decline of the use of *may* in its speaker-oriented modal meaning indicating permission, while the epistemic sense of *may* still held a strong position. At the same time, the use of *can* had become more pronounced, which Leech attributed to two broad trends arising from societal influences, namely, Americanization and colloquialization. The trend of colloquialization was also supported in investigations of historical corpora by Biber and Finegan (1989).

To obtain an idea of how frequently modal auxiliaries *can* and *may* are used, we point to the Longman Grammar of Spoken and Written English (Biber et al., 1999), based on the Longman Spoken and Written English corpus, in which the overall frequency of *can* and *may* is, respectively, 2,500 and 1,000 occurrences per million words. While *can* is relatively common in all registers, *may* is less common in spoken conversation, but extremely common in academic prose. For example, in academic prose, *can* has less than 500 occurrences per million words for expressing permission, slightly less than 1,500 occurrences for root possibility, and slightly more than 1,500 for ability. At the same time, the frequency of *may* is less than 500 occurrences for permission but very high (almost 3,000 occurrences per million words) when expressing epistemic possibility.

Psycholinguistic studies on modality

Literature on sentence processing distinguishes two basic, competing theories of sentence comprehension as reflected in the restricted and unrestricted accounts of parsing (Crocker et al., 2006; Garrod & Pickering, 1999). The two positions differ in that the restricted or syntax-first accounts assume that parsing is a “two-stage” process with syntactic information employed initially, whereas the unrestricted or constraint-based accounts reduce the distinction between syntax and semantics, assuming that all relevant sources of information (syntactic, semantic, or discourse) affect initial parsing decisions.

Psycholinguistic studies on modality are extremely rare. Among them, we first point to Huette *et al.*'s (2010) eye-tracking study that investigated whether the use of the speaker-oriented modal *should* and a stronger *must* would reveal differences in processing unambiguous propositions. Although reaction times did not show differences in processing, there was a shift in fixations to the target word for *should*. It was therefore concluded that two mental models were activated at the same time, including both agreement and disagreement with the statement in question. Suggesting that pragmatic constraints influence one's decisions more than syntactic constraints, the authors considered phrases with modal verbs as statements that are gradually resolved, having continuous access to semantic and pragmatic constraints. This way they provide evidence for unrestricted or constraint-based theories of sentence processing. Another study that aimed to contribute to the sentence processing accounts is a study by Giskes (2018) that used event-related potentials (ERPs) to investigate the online processing of Norwegian modal verbs *believe* and *doubt*. The author hypothesized that if these verbs are processed fully and immediately, the relations created by them will influence plausibility and comprehenders' expectations when processing the verb complements containing semantically implausible words. The findings suggested that at least some part of the verb lexical semantics is processed at the moment of encountering the critical words.

Stewart *et al.* (2009a, b) conducted a series of SPR experiments to examine how readers process counterfactual conditionals. They reported a significant reading-time penalty for the critical region when there was a mismatch between the previous information and the incoming input. Similar results, showing that semantic ambiguities caused by anomalies at discourse level trigger rapid increased reaction times in the same sentence in which the semantic mismatch occurred, were also provided in SPR discourse processing studies by Myers and O'Brien (1998) and O'Brien *et al.* (1998). Stewart *et al.* (2009b) pointed out that in discourse and sentence processing it is often not possible to isolate only one word that determined the degree of consistency (on which the increase in reading time occurs) but rather that there is a gradual accumulation of semantic information over the span of several words, which helps disambiguation.

Contributing to a better understanding of online sentence processing of modality, these studies have provided some evidence toward the psycholinguistic theoretical models that propose rapid semantic activation and discourse interpretation at the speed that is guided by the ability to interpret the context as the text unfolds.

Self-paced reading studies and online processing

The current study used a SPR task because the SPR paradigm has been extensively exploited in research measuring native and nonnative sensitivity to morpho-syntactic violations, syntactic or semantic ambiguity, and pragmatic plausibility (e.g., Hopp 2006, 2016; Jegerski, 2012, 2016; Marinis, 2003; Mifka-Profozic, 2017; Pliatsikas & Marinis, 2013; Roberts & Felser, 2011; Roberts & Liszka 2013, Stewart *et al.*, 2009a; Tokowicz & Warren, 2010). Research on monolingual sentence processing has provided evidence that native speakers vary their reading times on a word-by-word basis and adjust their reading depending on the properties of each

word such as the length, frequency, and/or word complexity (Just et al. 1982; Keating & Jegerski, 2015). An increase in reading time is shown either on the target anomalous words or on the immediately following words, indicating problems in assigning the meaning to form when either grammatical violations occur, or logical/semantic meaning of the sentence is made ambiguous.

In SPR techniques, two assumptions have been accepted. The first is the eye-mind assumption, which posits that our eyes remain fixated on a word as long as it is being processed, so gaze duration or time spent on reading a word directly indicates the time taken to process a newly seen word involving the use of information from a preceding context. The second assumption is related to the immediacy of reaction, suggesting that the interpretations at all levels of processing occur immediately after the problem has been encountered. Immediate longer reaction times, either on the target word or on the following word (spillover), have been recorded in studies that tested the processing of grammatical violations (e.g., Hopp, 2006; Jegerski, 2016; Pliatsikas & Marinis, 2013; Roberts & Liszka, 2013; Tokowicz & Warren, 2010).

To date, it has been widely accepted and supported by theories of comprehension that sentence processing is incremental, and that syntactic analysis is computed immediately on each word before the next word is encountered (Jegerski, 2012, 2014; Keating & Jegerski, 2015). As mentioned earlier, theoretical accounts differ on whether it is assumed that syntactic analysis is performed independently, or it is affected by semantic and thematic information (Garrod & Pickering, 1999). Research into garden-path effects, for example, provides evidence that semantic analysis also takes place incrementally such that the word meaning is rapidly integrated with the preceding context (Crocker et al., 2006). In order for semantic roles to be satisfied, a context is required, which readers or listeners use to obtain contextual cues to facilitate the processing of a sentence. The use of context is not restricted to a single sentence but extends to the information from preceding sentences too (Stewart et al., 2009b).

SPR studies have provided evidence for both restricted and constraint-based models of sentence processing. For example, in a study on processing English regular and irregular past tense morphology at the sentence level, Pliatsikas and Marinis (2013) found a difference that they attributed to the morphological distinction between the two verb types—regular being explained by the activation of automatic regular rule, and irregular as discrete items existing in the mental lexicon. Native speakers showed a clear distinction between processing regular and irregular past tense forms, but shorter reaction times when encountering anomalous regularized verblike forms compared to irregularized forms. These differences were explained by the modular model of processing. Jegerski (2012) investigated subject-object ambiguities in a study including Spanish native speakers and near-native speakers with English as an L1. Her results also aligned with a modular model of parsing where the earliest stage occurs on the basis of syntactic information, and additional sources of linguistic information, semantic and thematic, come into play only later.

In contrast, Hopp (2016), who argues that native and nonnative comprehension processes are fundamentally similar, found that lexical processing of less frequent words systematically affected syntactic processing of L2 speakers, with linear

decreases in speed of building syntactic structures as verb frequencies decreased. Native processing showed similar tendencies but for much lower frequency verbs. By demonstrating that lexical access of the verb precedes the structural, syntactic processing, and that lexical information can guide syntactic processing, Hopp's findings support the constraint-based models of sentence processing.

Evidence coming from ERPs is also inconclusive: Some studies have suggested that syntactic and semantic processing are distinguishable and syntactic parsing independent of semantic influences (Brown & Hagoort, 2006; Hagoort, 2003; Hagoort *et al.*, 1993; Kuperberg *et al.*, 2003) whereas others have found evidence of integration in which syntactic analysis is affected by semantic information, that is, semantic and pragmatic knowledge is modulating the parsing operations (Kim & Osterhout, 2005; Molinaro *et al.*, 2011). Kim and Osterhout (2005), for example, observed that, at least under certain conditions, semantics (and not syntax) has a dominant role in online sentence processing. Kuperberg (2007) suggested that normal language comprehension takes place involving both semantic mechanisms and syntactic (combinatorial) mechanisms that assign structure to a sentence on the basis of morpho-syntactic rules and semantic-thematic constraints.

The present study

Because modal auxiliaries are characterized by indeterminacy (Coates, 2014) and obvious difficulties in assigning them grammatical roles, the manipulation of *can* and *may*, as indicated previously, can result in either syntactic violation or semantic ambiguities, as well as in pragmatic infelicities. We hoped that such differentiations would become visible in a SPR task where the distinctive use of *can* and *may*, depending on either their epistemic or nonepistemic (agent-oriented or speaker-oriented) modal meaning, can be closely monitored and measured.

As already indicated, we chose one modal verb for each category as formulated by Bybee *et al.* (1994): '*can*' with agent-oriented modal meaning indicating *ability*, '*may*' bearing modal meaning of *epistemic possibility*, and '*may*' with speaker-oriented modal meaning indicating *permission* in formal contexts. Table 1 shows an example of a sentence for each category, in its congruent (semantically appropriate, grammatical, or marked for formality) and incongruent (semantically ambiguous, ungrammatical, and unmarked for formality) usage in context.

As can be seen in Table 1, the target sentence always starts with a noun phrase (subject) that is usually a personal pronoun linked to a verb phrase that starts with the modal (*can* or *may*), which is then followed by either a lexical verb (in *ability*, *permission*, and some *epistemic* examples), or can be followed by an auxiliary (*be* or *have*) and a lexical verb in the rest of the epistemic examples. The rest of the verb phrase usually contains an object or an adverb with their attachments. The noun phrase at the beginning of each sentence is considered segment 0, the modal verb is always segment 1, and a lexical verb is either segment 2 or segment 3 (with one exception in the epistemic condition where the lexical verb falls on segment 4). The rest (segments 5 and 6), in some cases followed by a word or two, are the remainder of the verb phrase.

Table 1. Self-paced reading task: Examples of items in each modal category

	Congruent	Incongruent
Ability	<i>Molly has recently enrolled on a computer course and is learning how to use the keyboard. She can already type fast and accurately although she is a beginner. Molly is a good learner.</i>	<i>Molly has recently enrolled on a computer course and is learning how to use the keyboard. She *may already type fast and accurately although she is a beginner. Molly is a good learner.</i>
Epistemic	<i>Simon has not seen his neighbors for quite a long time. They may have left for South America to visit their daughter. Simon has never met their daughter.</i>	<i>Simon has not seen his neighbors for quite a long time. They *can have left for South America to visit their daughter. Simon has never met their daughter.</i>
	Marked for Formality	Unmarked for Formality
Permission	<i>Cathy would like to discuss some job-related matters with her company director. “The director is not available but you may talk to his assistant,” says the secretary. The secretary will make an appointment for Cathy.</i>	<i>Cathy would like to discuss some job-related matters with her company director. “The director is not available but you can talk to his assistant,” says the secretary. The secretary will make an appointment for Cathy.</i>

Table 2. Distribution of segments in target verb phrases

Modality type	Segment						
	0	1	2	3	4	5	6
Ability^a	<i>He</i>	<i>can/may</i>	<i>read</i>	<i>books</i>	<i>in</i>	<i>six</i>	<i>languages</i>
Epistemic	<i>They</i>	<i>may/can</i>	<i>have</i>	<i>left</i>	<i>for</i>	<i>South</i>	<i>America</i>
Permission^b	<i>You</i>	<i>may/can</i>	<i>talk</i>	<i>to</i>	<i>his</i>	<i>assistant,</i>	<i>says . . .</i>

^aAgent-oriented.
^bSpeaker-oriented.

Table 2 shows the distribution of segments in the target verb phrases.

Our objective in the current study was to test whether the mismatch or incongruence of the modal verb and the context would result in reading penalties in terms of longer reaction times, and if yes, whether these occur immediately (i.e., in the target sentence). With that aim, we looked at reading times of seven words starting from the beginning of the sentence, as indicated in the previous paragraph.

We made the following predictions:

Hypothesis 1 (H1): With regard to research question (hereafter RQ) 1, a mismatch or the incongruent use of the modal verb would result in semantic ambiguity and increased reading times following the segments containing the modal auxiliary and the lexical verb, when the context allows for disambiguation.

Hypothesis 2 (H2): With regard to RQ2, a mismatch or the incongruent use of the modal verb would result in syntactic violation and rapidly increased reading times on the segment containing the lexical verb and possibly spillover on the following segments;

Hypothesis 3 (H3): With regard to RQ3, a mismatch in the level of formality between the context and the modal verb unmarked for formality would not cause significant changes in the reading times.

Our hypotheses are based on empirical findings to date that show that language processing is incremental, and that any violation, either syntactic or semantic, causes an increase in reading time in native speakers.

Method

All research materials, analytic methods, and data files associated with this study are available in the Open Science Framework using the following link: https://osf.io/mg4x2/?view_only=58c1db6fa9a94f1c8b31ef96af968faa

Participants

Forty participants volunteered in this study. The participants were undergraduate students from the Education, Biology and Economics Department at a UK university. Their mean age was 20.95 ($SD = 2.17$; range 18–35). They were paid £6 for participation in the study.

Materials

The study used 36 target items with the modals *can* and *may* manipulated so that each of the two modal verbs was used either as matching the context (congruent) or in a mismatching context (i.e., incongruent, where the counterpart modal should have been used). Altogether each participant had to read 56 items (36 experimental target items and 20 fillers) and 20 comprehension questions (10 corresponding to target items and 10 to fillers) to check that participants were processing the sentences rather than skipping through without properly reading. The experimental target items and fillers were presented to the participants in a pseudorandomized manner (see “Task and Procedure” section). The 36 target items included in the task were divided according to their meaning into *agent-oriented* (expressing *ability*), items expressing *epistemic possibility*, and *speaker-oriented* (expressing *permission*).

An important feature of this study, different from a number of other SPR studies in which target sentences were presented in isolation, was the context in which target sentences containing a modal verb appeared. The role of a context is especially important for semantic interpretation (Altmann & Steedman, 1988) because each ensuing word of a sentence is processed incrementally and checked against the previous context to facilitate interpretation and possible lexical ambiguity resolution. Thus, each item in the present study consisted of three sentences where the first sentence was an introduction providing the context, the second sentence was the target sentence containing the modal verb, and the third sentence wrapped up the whole event. Similarly, fillers contained three sentences, with the second sentence containing matched/mismatched items that were not modal verbs (e.g., old/oldest, making/make, like/likes).¹

Task and procedure

The general order of presentation of the 56 items was two experimental target items followed by a filler (a target item-to-filler rate of 2:1 for 32 target items and 1:1 for two target items), with 20 randomly appearing comprehension questions, 10 following the 36 target items, and 10 following the 20 fillers.

All effort was made to design items that were as comparable as possible in terms of the number of syllables (Jegerski, 2014). The first word in the sentence (segment 0) was always either a personal pronoun or a name, the second word (segment 1) was the modal (can/may); the third word (segment 2) was a one syllable word in 34 out of 36 sentences (two words had two syllables); the fourth word (segment 3) was a one syllable word in 29 of the target sentences and the remaining 7 words had two syllables. All these words were among the most frequent words used in everyday life (see Open Science Framework). The experimental target items and fillers were counterbalanced to account for possible order effects (see also “Data Analysis” section). Thus, half of the participants read one “version” and the other half read another “version” of the task. Both versions comprised 18 target items containing the congruent (consistent with the context) modal auxiliary use, 18 items containing incongruent modal auxiliary use, 10 grammatical fillers, and 10 ungrammatical fillers. The versions were identical except for two key differences: (1) if an item was acceptable (i.e., congruent/grammatical/marked for formality) in one version, it was unacceptable (i.e., incongruent/grammatically incorrect/unmarked for formality) in the other version, and vice versa; (2) the order of presentation of target items and fillers was mirrored so that half the participants encountered items in one order (target items 1–36, fillers 1–20), and the other half in the reverse order (target items 36–1, fillers 20–1). As already said, in each target item, target sentences were embedded between two other sentences. Each filler item also consisted of three sentences.

Before starting to read the sentences, the participants read three items for practice—these had a similar structure to the experimental items but were unrelated to the use of modal auxiliaries and were designed only to help the participants familiarize themselves with the SPR task.

The experiment used the PsychoPy software (Peirce, 2007, 2009) that is freely available on the Internet and has so far been widely used in eye-tracking and SPR studies. The method used in the present SPR task was the “stationary window” where the whole text of each experimental item (and of each filler) appeared on the screen word by word until the end of the entire task. Thus, only one word was visible on the screen at a time, meaning participants could not return to re-read earlier segments or sentences. The words of the sentences were white on the black screen and were positioned in the middle of the screen. Participants received the instructions both orally and in written form before the commencement of the task. They were instructed to press the space bar to proceed, and as they pressed the space bar the word on the screen would disappear and the next word would appear. After each set of sentences an instruction appeared on the screen to remind participants what they were required to do. The instruction read: *Please read each sentence carefully. Press the space bar to proceed.* At the end of the whole task there was a sentence saying that the task was finished. Lastly, participants were thanked for their contribution to the experiment.

The task was administered individually so that each participant was tested separately, in a quiet classroom. As noted previously, among the 56 items containing the 36 experimental target items and the 20 fillers, 20 (55.5%) random comprehension yes/no questions were inserted to make sure that the participants were paying attention to the meaning of the sentences they were reading. Specifically, comprehension was tested on 10/36 target items (27% of the time) and 10/20 fillers (50% of the time). The questions were unrelated to the use of modals not to interfere with the processing of the sentences in which modals were used, as explained in Roberts and Liszka (2013).

Study design

We ran three mixed-effects regression analyses; one for each modal verb type (ability, epistemic, permission). All three analyses had one continuous dependent (or outcome, predicted) variable and three independent (or predictor, explanatory) variables. In all three analyses, the dependent variable was log-transformed reading times (see “Data Analysis” section). In the analysis of ability and epistemic modal verbs, the categorical independent variables were congruency (two levels: congruent, incongruent), segment (seven levels: segments 0, 1, 2, 3, 4, 5, 6), and target stimuli order (two levels: version 1 order, version 2 order). Congruency and segment were measured within participants because all participants encountered equal numbers of items from all levels of both independent variables. Stimuli order was measured between participants because half the participants encountered the stimuli in one order, and the other half in the reverse order (see “Task and Procedure” section). In the analysis of permission modal verbs, the independent variables were formality (two levels: marked, unmarked), segment (as mentioned previously), and stimuli order (as mentioned previously).

Data analysis

We first collected all reading times and responses to the comprehension questions corresponding to target and filler items using PsychoPy. Reading times were calculated in milliseconds starting from segment 0, that is, the word preceding the modal verb, to segments (i.e., words) 1, 2, 3, 4, 5, and 6. Accuracy for the 20 comprehension questions was high ($M = 19.25$, $SD = 0.78$). Participants were removed before further analysis if they scored less than 90% comprehension accuracy (i.e., gave incorrect answers to more than two comprehension questions). A total of 18 participants got all questions correct (100% accuracy), 14 gave an incorrect answer to one question (95% accuracy), eight gave incorrect answers to two questions (90% accuracy), while four gave incorrect answers to more than two questions ($< 90\%$ accuracy). These latter four participants were therefore removed before further analyses. Because no participants guessed the aim of the experiment, there were no removals for this reason, and the final number of participants retained was 40.

In line with Keating and Jegerski’s (2015) recommendations (see also McManus & Marsden, 2018), we then removed values outside 150–2,000 ms. This resulted in the removal of 26 data points (0.26% of the data). We then checked for any remaining extreme cases by identifying whether any participant’s segment reading time exceeded their mean reading time for that sentence plus two standard deviations,

but none were found. The distribution of the data were then checked, and reading times log-transformed to reduce positive skew.

In line with recent calls for greater methodological transparency, particularly in the related field of second language research involving SPR tests (Marsden et al., 2018), we report the full list of instrument reliability estimates of our stimuli (see Open Science Framework). To this end, we hope to allow for better understanding of the amount of error in the data, and to aid future instrument development by advancing understanding on the general psychometric properties of SPR tests. Instrument reliability for each modal verb type (ability, epistemic, permission) was calculated as the internal consistency of reading times for each of segment, per “set” of six congruent/marked or incongruent/unmarked sentences. This fine-grained approach provided data on how the reliability of the SPR test stimuli varied according to instrument features (Plonsky & Derrick, 2016), in our case, by the congruency of items, modal verbs, test versions, sentences, and segments (see Open Science Framework).

The 84 estimates, obtained using the UserFriendlyScience (Peters et al., 2018) package in R (R Core Team, 2019), showed, on average, high internal consistency ($Mdn = .94$, $IQR = .06$, range .65–.99)² by Revelle’s (2018) Omega, which, unlike Cronbach’s alpha, allowed us to take into account differences in the degree to which individual sentence segments tapped the construct in question (McNeish, 2018). Looking at the various instrument features, internal consistency of items was high for both congruent and incongruent items (42 estimates each, respectively $Mdn = .94$, $.94$, $IQR = .04$, $.08$), and across the different types of modality (28 estimates per verb type, ability modal verbs: $Mdn = .95$, $IQR = .05$; epistemic: $Mdn = .94$, $SD = .03$; permission: $Mdn = .94$, $SD = .10$). While both stimuli versions 1 and 2 (both 42 estimates) also had high internal consistency, instrument reliability was slightly higher, on average, for stimuli version 1 ($Mdn = .97$, $IQR = .03$) than version 2 ($Mdn = .91$, $IQR = .06$). Similarly, reliability of measurement was consistently high across all segments, but somewhat higher for segments 0 and 1 (respectively, $M = .96$, $.95$, $SD = .03$, $.04$) compared with segments 2 to 6 (which ranged from $Mdn = .92$ to $.94$, and $IQR = .03$ to $.09$).

To analyze the interactive effect of congruency and segment on log-transformed reading times (log RTs), we conducted mixed-effects linear regression in R (R Core Team, 2019) using lme4 (Bates et al., 2015), running separate analyses for each modal verb type (ability, epistemic, permission). Mixed-effects models offered a more flexible and robust approach to analysis of variance, for example avoiding violation of the assumption of independence (Cunnings & Finlayson, 2015). The first explanatory variable, congruency/formality, was sum-coded to yield main (rather than simple) effects. The two levels of this categorical predictor were “congruent” (for permission, “marked”) coded as +1, and “incongruent” (for permission, “unmarked”) as -1, while the reference level was the mid-point (mean) of mean congruent/marked and incongruent/unmarked log RTs. The second explanatory variable, segment, was a categorical predictor with seven levels, which was treatment-coded; thus, at the intercept, all segments were coded as 0, for segments 1 to 6 each segment/effect in question was coded as 1 and all other segments as 0. The third explanatory variable, stimuli order, was included in the analysis to account for any variance explained by the order in which participants encountered

sentences (see “Task and Procedure” section). Stimuli order was a categorical predictor, sum-coded to aid interpretability (see previous text), and with two levels, “version 1” coded as -1 and “version 2” as $+1$. The reference level of stimuli order was the mid-point (mean) of the mean version 1 and version 2 log RTs.

Congruency/formality, segment, and stimuli order were entered as fixed effects, with an interaction specified between congruency/formality and segment, while participant and sentence were entered as random effects. Specifically, the fixed effect of congruency/formality was conditioned on participant, meaning the model allowed participants to vary both in terms of their overall speed of response (random intercepts) and the extent to which congruency/formality affected their log RTs (random slopes). Segment was conditioned on sentence, meaning sentences were allowed to vary in terms of overall log RTs (random intercepts) and to exhibit by-sentence variation with regard to the strength of the effect of segment on log RTs (random slopes). Importantly, because the random effects linked segment to sentences in this way, the effects of segment differences that we were not interested in (e.g., word length) on log RTs were accounted for in the model. Finally, stimuli order was conditioned on participant to account for by-participant variation with regard to both overall log RTs (random intercepts) and the strength of any stimuli order effect on log RTs (random slopes).

Because our predictors were theoretically motivated and we had clear hypotheses about their relative effect on log RT for the different types of modality, we took a confirmatory approach and entered them consistently across all models rather than building optimal models in a stepwise fashion (Winter, 2020).

For all models, we first used chi-square to check for a statistically significant overall interaction between congruency/formality and segment (a null model with no interaction vs. a model with interaction terms), and then looked at the various fixed effect interaction parameters to identify which segments showed delayed log RTs due to disambiguation, and the magnitude of slow down. We adjusted the alpha value from .05 to .0083 to account for the increased chance of a Type I error from the six repeated tests (in each model, any of the six interaction parameters could support the hypothesis in question). We report effect estimates, corresponding p-values and 95% confidence intervals (CIs), standard errors, degrees of freedom, Wald t -values, and R^2 values that show the proportion of variance explained by the fixed effects (marginal R^2) and fixed and random effects combined (conditional R^2), calculated using *MuMIn* package (Bartoń, 2018) in R (R Core Team, 2019). In line with Plonsky and Ghanbar (2018), we interpreted R^2 values as small (.18), medium (.32), or large (.51) with regard to the amount of variance in log RTs explained.

Results

The trimmed reading times, log transformed, were modeled for each of the different types of modality (ability, epistemic, permission) to investigate the magnitude and statistical significance of any interaction between congruency/formality and segment as predictors of log RT (RQs 1–3), and any effect of stimuli order, which we were seeking to rule out.

Tables 3–5 show the effects of the fixed factors including the degree of interaction between independent variables. The results are visualized in Figures 1–3.

Table 3. Ability modality: Summary of fixed effects (outcome = log RT)

Parameter	Estimate	95% CIs for Estimate	SE	Df	t value	p
(Intercept)	6.02	[5.95,6.10]	0.04	50.02	151.27	<.001***
Congruency1 × Seg1	0.00	[-0.02,0.03]	0.01	3227.00	0.30	.764
Congruency1 × Seg2	0.01	[-0.02,0.03]	0.01	3227.00	0.48	.633
Congruency1 × Seg3	0.00	[-0.03,0.02]	0.01	3227.00	-0.18	.855
Congruency1 × Seg4	-0.11	[-0.14,-0.08]	0.01	3228.00	-8.14	<.001***
Congruency1 × Seg5	-0.04	[-0.06,-0.01]	0.01	3227.00	-2.80	.005*
Congruency1 × Seg6	-0.02	[-0.05,0.00]	0.01	3227.00	-1.67	.094

Notes: Alpha value adjusted to .0083 to correct for multiple (i.e., six) tests. Seg = Segment.
 *** $p < .001$. ** $p < .005$. * $p < .0083$.

Table 4. Epistemic modality: Summary of fixed effects (outcome = log RT)

Parameter	Estimate	95% CIs for Estimate	SE	Df	t value	p
(Intercept)	6.01	[5.93,6.10]	0.04	40.86	138.39	<.001***
Congruency1 × Seg1	-0.01	[-0.04,0.02]	0.02	3217.00	-0.58	.563
Congruency1 × Seg2	-0.03	[-0.06,-0.00]	0.02	3217.00	-2.01	.044
Congruency1 × Seg3	-0.05	[-0.08,-0.02]	0.02	3217.00	-3.09	.002**
Congruency1 × Seg4	-0.09	[-0.12,-0.06]	0.02	3219.00	-5.62	<.001***
Congruency1 × Seg5	-0.04	[-0.07,-0.01]	0.02	3217.00	-2.63	.008
Congruency1 × Seg6	-0.03	[-0.06,0.00]	0.02	3217.00	-1.66	.096

Notes: Alpha value adjusted to .0083 to correct for multiple (i.e., six) tests. Seg = Segment.
 *** $p < .001$. ** $p < .005$. * $p < .0083$.

Table 5. Permission modality: Summary of fixed effects (outcome = log RT)

Parameter	Estimate	95% CIs for Estimate	SE	Df	t value	p
(Intercept)	5.97	[5.91,6.03]	0.03	41.95	186.68	<.001***
Formality1 × Seg1	-0.01	[-0.03,0.02]	0.01	3232.00	-0.67	.504
Formality1 × Seg2	-0.01	[-0.04,0.01]	0.01	3232.00	-0.89	.374
Formality1 × Seg3	-0.02	[-0.04,0.01]	0.01	3232.00	-1.39	.163
Formality1 × Seg4	-0.01	[-0.04,0.01]	0.01	3232.00	-1.10	.272
Formality1 × Seg5	-0.02	[-0.04,0.01]	0.01	3232.00	-1.46	.145
Formality1 × Seg6	-0.02	[-0.04,0.01]	0.01	3232.00	-1.35	.177

Notes: Alpha value adjusted to .0083 to correct for multiple (i.e., six) tests. Seg = Segment.
 *** $p < .001$. ** $p < .005$. * $p < .0083$.

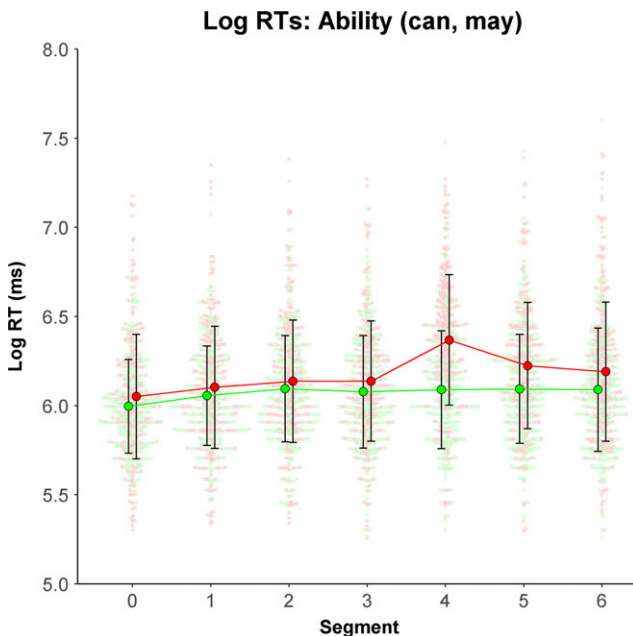


Figure 1. (color online) Ability modal verb log RTs (small dots = individual log RTs, larger points = mean log RTs connected by lines, vertical black bars = $\pm 1 \times$ standard deviation, green = congruent, red = incongruent, Y-axis truncated at 5.0).

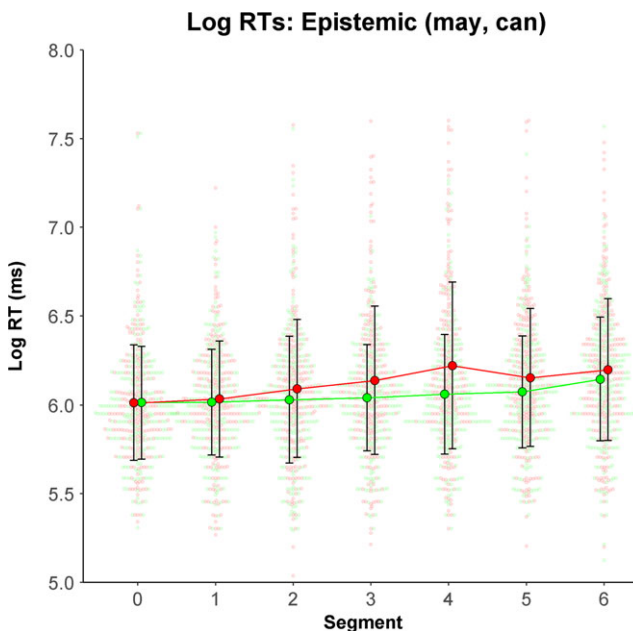


Figure 2. (color online) Epistemic modal verb log RTs (small dots = individual log RTs, larger points = mean log RTs connected by lines, vertical black bars = $\pm 1 \times$ standard deviation, green = congruent, red = incongruent, Y-axis truncated at 5.0).

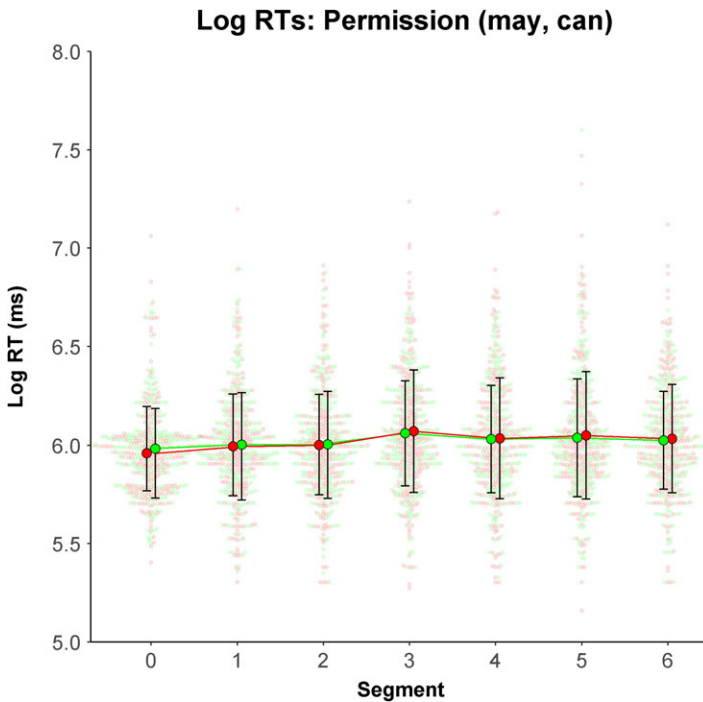


Figure 3. (color online) Permission modal verb log RTs (small dots = individual log RTs, larger points = mean log RTs connected by lines, vertical black bars = $\pm 1 \times$ standard deviation, green = marked for formality, red = unmarked for formality, Y-axis truncated at 5.0).

(To allow for comparison with other studies and inclusion of these results in future meta-analyses, we provide the full descriptive statistics of the reading times, before log-transformation, in the Open Science Framework at the previously mentioned link).

For all models, the fixed effects explained a relatively small proportion of variance—marginal $R^2 = .09$, $.06$, and $.03$ for ability, epistemic, and permission, respectively, a magnitude of explained variance comparable with similar studies using SPR (e.g., McManus & Marsden, 2018)—and the fixed and random effects in combination explained a large amount of variance (conditional $R^2 = .66$, $.64$, and $.60$, for the preceding models, respectively).

We first provide a detailed explanation of the results of the ability model, focusing on the interactions rather than additive effects (provided in supplementary materials), which simply show how quickly each segment was read in comparison with segment 0, ignoring congruency and stimuli order. With the key elements clarified, we move to a more concise report for the epistemic and permission models.

Table 3 shows the effects of congruency and segment on ability log RTs (RQ1). Recall that the reference level for congruency was the mid-point (mean) of mean congruent and incongruent log RTs, and for segment, level 0 (see “Data Analysis” section). Thus, the intercept estimate (log RT = 6.02 ms) denotes the

mid-point of the mean congruent and mean incongruent sentence log RTs at segment 0. The key information with regard to H1 concerns the congruency \times segment interaction estimates. The overall interaction between congruency and segment on log RT was significant as shown by comparing a null model with additive predictors only to the full ability model, which had the additional six interaction parameters ($\chi^2(6) = 111.22, p < .001$, not shown in Table 3).

The Congruency1 \times Seg1 (2, 3, 4 etc.) estimates (Table 3) show the different predictive effects on log RT for congruent and incongruent sentences, at the segment in question, as a change from the reference levels. The statistically significant, negative estimate for Congruency1 \times Seg4, $b = -0.11$, 95% CIs $[-0.14, -0.08]$, $t(3228.00) = -8.14, p < .001$; and Congruency1 \times Seg5, $b = -0.04$, 95% CIs $[-0.06, -0.01]$, $t(3227.00) = -2.80, p = .005$, with CIs not passing through zero, show that segments 4 and 5 had a significantly different predictive effect on log RT for congruent and incongruent sentences (see also Figure 1, ability log RTs). In other words, participants' sensitivity to the semantic ambiguity associated with the incongruent verb (*may*) was detectible in segments 4 and 5 of the sentence. Stimuli order (not shown in Table 3) had a nonsignificant effect on log RT $b = -0.06$, 95% CIs $[-0.12, 0.01]$, $t(41.42) = -1.67, p = .103$.

To summarize then, for ability modal verbs, log RT was influenced by congruency and segment in combination, with a spillover effect observable at segments 4 and 5, where the semantic ambiguity caused a significant delay. Thus, our hypothesis (H1) that the incongruent use of the modal verb (i.e., semantic ambiguity) would result in increased reading times following the lexical verb but within the verb phrase was supported.

Table 4 shows both the effects of congruency and segment for epistemic log RTs (RQ2). Comparison of a null additive model (no interaction specified) to the full epistemic model, which had the additional six interaction parameters, showed an overall significant interaction between congruency and segment ($\chi^2(6) = 41.41, p < .001$). Table 4 shows that participants displayed sensitivity to the incongruency, with slower log RTs for incongruent sentences, from segment 3, Congruent1 \times Seg3, $b = -0.05$, 95% CIs $[-0.08, -0.02]$, $t(3217.00) = -3.09, p = .002$; to segment 4, at which the effect was strongest, Congruent1 \times Seg4, $b = -0.09$, 95% CIs $[-0.12, -0.06]$, $t(3219.00) = -5.62, p < .001$. The magnitude of effects for the other segments show that participants *might* have been sensitive to the incongruency even earlier, at segment 2, and continuing until segment 5 because confidence intervals for these two parameters do not pass through zero. However, this could not be reliably established because the effects were nonsignificant at the adjusted alpha level of .0083 (see also Figure 2, epistemic log RTs). Stimuli order (not shown in Table 4) had a nonsignificant effect on log RT $b = -0.07$, 95% CIs $[-0.15, 0.00]$, $t(30.04) = -1.94, p = .062$.

To summarize for epistemic modal verbs, log RT was influenced by congruency and segment in combination, where the delay caused by the grammatical violation (*can*) was first observable at segment 3 and with a spillover to segment 4. Thus, our hypothesis (H2) that a mismatch or the incongruent use of the modal verb would result in grammatical violation and rapidly increased reading times on the immediate constituents of the verb phrase and possibly spill over on the following

segments was supported. Thus, a reading time delay for epistemic modal verbs started one segment earlier than in sentences with ability modal verbs.

Finally, for permission modal verbs (RQ3), comparison of an additive null model (no interaction) with the full permission model, containing the additional six interaction parameters, showed no overall significant interaction between formality and segment ($\chi^2(6) = 3.23, p = .779$). This is reflected in the interaction estimates in Table 5 (see also Figure 3, permission log RTs), which are all nonsignificant at the .0083 level, with CIs passing through zero. While estimates are consistently negative, suggesting all segments of unmarked for formality sentence (*can*) were read more slowly than those of marked for formality sentences (*may*), the significance of this effect could not be reliably established (see preceding text). Stimuli order (not shown in Table 5) also had a nonsignificant effect on log RT $b = -0.04, 95\%CIs [-0.09, 0.01], t(30.50) = -1.54, p = .134$.

To summarize for permission modal verbs, log RT was not influenced by formality and segment in combination. Put simply, with permission modality, participants neither slowed down nor sped up significantly according to whether they read the modal marked for formality (*may*) or the modal unmarked for formality (*can*) in a formal/informal context. Thus, our hypothesis (H3) that a mismatch in the level of formality between the context and the unmarked for formality modal verb would not cause significant changes in the reading times, was supported.

Discussion

We start our discussion with the reference to our research questions, which asked whether a match or a mismatch between the context and the target (RQ1) agent-oriented modal verb expressing ability, (RQ2) the modal verb expressing epistemic possibility, and (RQ3) the speaker-oriented modal verb expressing permission would affect the reading times in the target sentences. We hypothesized that there would be an effect for incongruency in relation to ability and epistemic possibility, but there would not be an effect for the modal used when asking or giving permission. Following an elaboration of each substantive finding, we then reflect on the methodological dimensions of our article, which provide a significant contribution to the quality of published SPR studies.

The main contribution of our study is that it clearly differentiated the use of modal auxiliaries *can* and *may*, demonstrating that they are mutually exclusive when employed as an agent-oriented modal to indicate ability (*can*), and as a modal expressing epistemic possibility (*may*). However, in speaker-oriented meaning indicating permission their mutual exclusivity is not supported. To date, various descriptions of modal auxiliaries have been offered, for example semantic (Coates, Lyons, Palmer), formal (Huddleston), functional (Bybee), and so forth. However, none has been grounded in empirical (experimental) research into the processing of modal auxiliaries. Furthermore, no account thus far has provided experimentally tested support for the fine line of distinction between syntactic, semantic, and pragmatic aspects of meaning in each individual modal. We believe that the current study makes one step in that direction and thus contributes to better understanding of the English modal auxiliary verbs.

The analyses and the results presented in the preceding text demonstrate that our three hypotheses were all supported: For modal expression of ability, there was a significant interaction between congruency and segment, with a detectable spillover starting at segment 4 and continuing on segment 5. Recall that the use of *may* instead of *can* made the modal auxiliary incongruent in the given context, which led to semantic ambiguity. As shown in examples (10) and (11), it is possible to use modal auxiliary *can* both as agent-oriented and speaker-oriented expressions of modality, but the context will signal which use is appropriate in a given situation. Here we deal with a problem of semantic ambiguity that cannot be resolved at the point where the modal followed by a lexical verb is encountered, but only a step further where more context is provided and an attempt at integration fails. Clearly, there is a cost associated with maintaining two interpretations. This cost is incurred at the point of disambiguation and therefore, we see the spillover on segments 4 and 5. The penalty in reading time suggests that readers are trying to lexically disambiguate the problem. Similar effects have been observed in Roberts and Liszka (2013) and Stewart *et al.* (2009a). The spillover effect still supports the immediacy assumption because the effect is observable on segments immediately following the lexical verb, and still within the verb phrase. The reason for this delay can be found in the fact that the context to disambiguate the meaning comes only after the modal. If our items had been constructed differently, with the context prior to the modal verb precluding the speaker-oriented interpretation of *can*, we might have been able to see an earlier effect caused by semantic ambiguity.

However, incongruency, that is, ungrammaticality in the use of epistemic *may*, has slightly different consequences: as can be seen in the “Results” section, there was a significant interaction between congruency and segment with the start of reading slowdown on segment 3, which then continued on segment 4, with an effect at segment 5, just short of significance using the adjusted alpha value. This is in line with SPR studies that have investigated the impact of syntactic violations and ungrammaticality on readers’ reaction times in milliseconds (e.g., Hopp, 2006, 2016; Jegerski, 2012, 2016; Pliatsikas & Marinis, 2013). Support for the immediacy assumption is obtained here even more convincingly. A slowdown in most sentences with incongruent/ungrammatical use of the modal verb in the context suggesting epistemic possibility, takes place within the immediate constituents of the verb phrase, most frequently on the lexical verb. This seems logical because a reaction to ungrammatical use of the modal auxiliary can be detected immediately, as soon as the incongruent modal has been encountered. Recall also that the segments immediately following the modal auxiliary in the target verb phrase are all one-syllable words, that is, either the auxiliary *be* or *have*, or a one-syllable lexical verb, so syntactic parsing occurs very quickly in an incremental manner.

Although the main effect of grammatical anomaly is observed immediately, on segment 3, there is also a spillover observable on segment 4 with a tendency to continue to segment 5. This is both similar and different from the condition that dealt with semantic ambiguity. It is similar in the sense that the spillover on segment 4 occurs in both conditions; and different because syntactic violation is first observed one segment earlier. There are several explanations for this: one is that syntactic distinctions are categorical, that is, a structure is either grammatical or ungrammatical, while semantic distinctions are more gradient, and they are highly dependent

on the context (Hagoort, 2003). Semantic anomalies always require a context to be detected and disambiguated (Altmann & Steedman, 1988). The second explanation can be found in the construction of items because most items did not provide sufficient information prior to the modal verb to block a possibility of grammatically acceptable but semantically unacceptable use of *may* in agent-oriented modality condition. For that reason, the results in relation to timing differences in reading penalty caused by either semantic or syntactic anomalies should be interpreted with caution.

In this case, due to the slightly delayed increase in time caused by semantic ambiguity, our results do not entirely align with Giskes (2018) or Hopp (2016), who suggest that lexical and semantic information is processed prior to syntactic information and that it influences syntactic processing. Because of the incomparability of the two conditions in our study (i.e., semantic and syntactic), it is also not possible to accept the interpretation that assumes the independent role of syntax as being unaffected by nonsyntactic information (Hagoort, 2003; Jegerski, 2012). Therefore, our results cannot be interpreted as supporting either of the two processing accounts (restricted and constraint-based). An investigation in this direction remains a task for future research on the processing of modal verbs.

As predicted, for modals of permission there was no significant interaction between formality (marked/unmarked) and segment. Participants neither slowed down nor sped up significantly across segments as a function of marking of the match or mismatch in sentences. This indicates that the participants in our study, young university students, do not exhibit significant changes in reading times when presented with a pragmatically less felicitous speaker-oriented use of the modal auxiliary expressing permission. One may argue that the experimental items could have been more controlled for the context formality, or that a different type of test could have been used that would clearly elicit preferences (e.g., a forced-choice task), so this also remains a task for future research. It is worth noting, however, that pragmatic preference is a qualitatively different type of distinction when compared with distinctions stemming from syntactic or semantic anomalies.

The prevalent use of the modal auxiliary *can* in speaker-oriented permission condition is not surprising in light of written and spoken corpus data collected in the 1990s, showing that in British English the use of *can* was on the rise, while *may* was used significantly less frequently as an expression of permission (Leech, 2003). To explain the linguistic behavior of young university students, it helps to remember that semantic change in the use of modals was already recorded several decades ago, showing a trend of language colloquialization, which Biber and Finegan (1989) named a “stylistic drift.” In today’s language use, *may* still holds a strong position as an expression of epistemic possibility. Our findings demonstrate that native speakers clearly show sensitivity to grammatical violation in such contexts. However, the use of *may* to formally request or give permission is undoubtedly in decline. The present study provides evidence for both.

In addition to the substantive results that this study has produced, it is also important to underline its methodological aspects, which aim to contribute to further improvements, and more scrutiny and transparency in conducting SPR studies. Our study is likely to be (1) one of few to estimate SPR instrument reliability (e.g., Marsden et al.’s [2018] methodological synthesis of SPR in second language

research found only two studies that did this), and (2) among the first in the modality literature to respond to recent calls to implement superior coefficients to Cronbach's alpha for instrument reliability estimation (McNeish, 2018; cf. Raykov & Marcoulides, 2019). We intend to encourage others to do the same.

In our study, the fine-grained set of estimates showed that instrument reliability was consistently high across the following instrument features: congruent/incongruent (or formally marked/unmarked) sentences, types of modality (ability, epistemic, permission), test versions, and segments, giving increased weight to the substantive findings. Where measurement consistency seems to have varied (albeit slightly), this appears to have been with the second test version, and from the second segment onward. While the two versions were sufficiently reliable for present purposes (e.g., their reliability far exceeded .74, a recently recommended general [but not absolute] threshold for minimum acceptable reliability in second language research; Plonsky & Derrick, 2016), future studies might explore, more comprehensively, how the consistency of measurement is affected by moderator variables such as instrument version and segments of the stimuli, how this may potentially affect substantive findings, and how SPR instrumentation can be made more reliable.

As is increasingly recognized in our field, the mixed model offered a more robust approach than analysis of variance, and allowed us to generalize above and beyond the idiosyncrasies of the participants and sentences used in this particular study because we modeled these as random effects.

Finally, we showed that it is possible to provide a highly nuanced analysis in which we were able to identify the segments (i.e., words) on which the reading slowdown was clearly observable. We acknowledge that such analyses may not be possible in studies in which longer discourse fragments are included, but in sentence-level processing it seems important to identify and localize the effects of either syntactic or semantic anomalies influencing the reading reaction times.

Notes

1. Importantly, because each item contained three sentences with only one as the target sentence containing the target verb, participants encountered 168 sentences altogether (108 in the target items and 60 in the filler items), among which the 36 target sentences appeared. Due to the comparatively large total number of sentences, we decided not to include a filler for every target item.
2. Where instrument reliability estimates were not normally distributed by Shapiro–Wilk test ($p < .05$), we report median and interquartile range, where estimates were normally distributed, we report the mean and standard deviation.

References

- Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.
- Bartoń, K. (2018). MuMIn: Multi-Model Inference. R package version 1.42.1. <https://cran.r-project.org/web/packages/MuMIn/index.html>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
- Biber, D., & Finegan, E. (1989). Drift and the evolution of English style: A history of three genres. *Language*, *65*(38), 487–517.

- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education Ltd.
- Brown, C., & Hagoort, P. (2006). On the electrophysiology of language comprehension: Implications for the human language system. In M. W. Crocker, M. J. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 213–237). Cambridge: Cambridge University Press.
- Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Bybee, J., Perkins, R., and Pagliuca, W. (1994). *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Coates, J. (1980). On the non-equivalence of *may* and *can*. *Lingua*, *50*(3), 209–220. [https://doi.org/10.1016/0024-3841\(80\)90026-1](https://doi.org/10.1016/0024-3841(80)90026-1)
- Coates, J. (2014). *The semantics of modal auxiliaries*. Routledge.
- Crocker, M. W., Pickering, M. J., & Clifton, C. (2006). *Architectures and mechanisms for language processing*. Cambridge: Cambridge University Press.
- Cunnings, I., & Finlayson, I. (2015). Mixed effects modeling and longitudinal data analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 159–181). New York: Routledge.
- Facchinetti, R., Krug, M. G., & Palmer, F. R. (2003). *Modality in contemporary English*. Berlin: Mouton de Gruyter.
- Garrod, S., & Pickering, M. (1999). *Language processing*. Hove: Psychology Press, Taylor and Francis.
- Giskes, A. (2018). *The online processing of epistemic modal verbs during language comprehension*. An ERP study (Unpublished MA thesis).
- Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic semantic violations. *Journal of Cognitive Neuroscience*, *15*(6), 883–899.
- Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP-measure of syntactic processing. *Language and Cognitive Processes*, *8*, 439–483.
- Hahne, A., & Frederici, A.D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, *13*(3), 339–356.
- Haigh, M., Stewart, A., & Connell, L. (2013). Reasoning as we read: Establishing the probability of causal conditionals. *Memory and Cognition*, *41*(1), 152–158.
- Hopp H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22*, 369–397.
- Hopp, H. (2016). The timing of lexical and syntactic processes in second language sentence comprehension. *Applied Psycholinguistics*, *37*(5), 1253–1280.
- Huddleston, R. D. (1976). Some theoretical issues in the description of the English verb. *Lingua*, *40*, 331–383.
- Huette, S., Matlock, T., & Spivey, M. J. (2010). The online processing of modal verbs: Parallel activation of competing mental models. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *32*(32), 1154–1159.
- Jegerski, J. (2012). The processing of subject–object ambiguities in native and near-native Mexican Spanish. *Bilingualism: Language and Cognition*, *15*(4), 721–735.
- Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York: Routledge.
- Jegerski, J. (2016). Number attraction effects in near-native Spanish sentence comprehension. *Studies in Second Language Acquisition*, *38*(1), 5–33.
- Just, M. A., Carpenter, P., & Woolley, J. D. (1982). Paradigms and processes in reading Comprehension. *Journal of Experimental Psychology, General*, *111*(2), 228– 238.
- Keating, G., & Jegerski, J. (2015). Experimental designs in sentence processing research. *Studies in Second Language Acquisition*, *37*(1), 1–32. <https://doi.org/10.1017/S0272263114000187>
- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*(2), 205–225.
- Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49. <https://doi.org/10.1016/j.brainres.2006.12.063>
- Kuperberg, G. R., Holcomb, P., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*(2), 272–293.

- Leech, G. (2003). Modality on the move: The English modal auxiliaries 1961–1992. In R. Facchinetti, M. G. Krug & F. R. Palmer (eds.), *Modality in contemporary English* (pp. 223–240). Berlin: Mouton de Gruyter.
- Lyons, J. (1977). *Semantics*, 2. Cambridge: Cambridge University Press.
- Marinis, T. (2003). Psycholinguistic techniques in second language acquisition research. *Second Language Research*, 19(2), 144–161.
- Marsden, E., Thompson, S., & Plonsky, L. (2018). A methodological synthesis of self-paced reading in second language research. *Applied Psycholinguistics*, 39(5), 861–904. <https://doi:10.1017/S0142716418000036>
- McManus, K., & Marsden, E. (2018). Signatures of automaticity during practice: Explicit instruction about L1 processing routines can improve L2 grammatical processing. *Applied Psycholinguistics*, 40(1), 205–234. <https://doi:10.1017/S0142716418000553>
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433. <https://doi:10.1037/met0000144>
- Mifka-Profozic, N. (2017). Processing epistemic modality in a second language: A self-paced reading study. *IRAL International Review of Applied Linguistics in Language Teaching*, 55(3), 245–264. <https://doi.org/10.1515/iral-2017-0107>
- Molinaro, N., Barber, H. A., & Carreiras, M. (2011). Grammatical agreement processing in reading: ERP findings and future directions. *Cortex*, 47(8), 908–930.
- Myers, J. L., & O'Brien, E. J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, 26, 131–157.
- Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, 78(2), 165–188.
- O'Brien, E. J., Rizella, M. L., Albrecht, J. E., & Halleran, J. G. (1998). Updating a situation model: A memory-based text processing. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 24(5), 1200–1210.
- Ozturk, O., & Papafragou, A. (2015). The acquisition of epistemic modality: From semantic meaning to pragmatic interpretation. *Language Learning and Development*, 11(3), 191–214. <https://doi:10.1080/15475441.2014.905169>
- Palmer, F. R. (1986). *Mood and modality*. Cambridge: Cambridge University Press.
- Palmer, F. R. (1990). *Modality and the English modals*. 2nd ed. London: Longman.
- Palmer, F. (2003). Modality in English: Theoretical, descriptive and typological issues. In R. Facchinetti, M. Krug, & F. Palmer (Eds.), *Modality in contemporary English* (pp. 1–17). Berlin: Mouton de Gruyter.
- Papafragou, A. (1998). The acquisition of modality: Implications for theories of semantic representation. *Mind and Language*, 13, 370–399.
- Papafragou, A., & Ozturk, O. (2006). Children's acquisition of epistemic modality. Cascadilla Proceedings Project.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2(10). doi: 10.3389/neuro.11.010.2008
- Peters, G.-J., Verbook, P., & Green, J. (2018). UserFriendlyScience: Quantitative analysis made accessible. R package version 0.7.2. <https://cran.r-project.org/web/packages/userfriendlyscience/index.html>
- Pliatskias, C., & Marinis, T. (2013). Processing of regular and irregular past tense morphology in highly proficient second language learners of English: A self-paced reading study. *Applied Psycholinguistics*, 34, 943–970.
- Plonsky, L., & Derrick, D. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal*, 100(2), 538–553.
- Plonsky, L., & Ghanbar, H. (2018). Multiple regression in L2 research: A methodological synthesis and guide to interpreting R2 values. *The Modern Language Journal*, 102(4), 713–731.
- Plonsky, L., & Oswald, F. L. (2017). Multiple regression as a flexible alternative to ANOVA in L2 research. *Studies in Second Language Acquisition*, 39(3), 579–592.
- Portner, P. (2009). *Modality*. Oxford: Oxford University Press.
- Raykov, T., & G. A. Marcoulides. (2019). Thanks coefficient alpha, we still need you! *Educational and Psychological Measurement*, 79(1), 200–210.

- R Core Team.** (2019). *R: A language and environment for statistical computing* (R Foundation for Statistical Computing). <https://www.R-project.org/>
- Revelle, W.** (2018). Using R and the psych package to find ω . <http://personality-project.org/r/psych/HowTo/omega.pdf>
- Roberts, L., & Felser, C.** (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics*, *32*(2), 299–331.
- Roberts, L., & Liszka, S.** (2013). Processing tense/aspect agreement violations online in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research*, 413–439.
- Stewart, A. J., Haigh, M. Kidd, E.** (2009a). An investigation into the online processing of counterfactual and indicative conditionals. *Quarterly Journal of Experimental Psychology* *62*(11), 2113–2125. <https://doi.org/10.1080/17470210902973106>
- Stewart, A. J., Kidd, E. and Haigh, M.** (2009b). Early sensitivity to discourse-level anomalies: Evidence from self-paced reading. *Discourse Processes*, *46*(1), 46–69. <https://doi.org/10.1080/01638530802629091>
- Talmy, L.** (1988). Force dynamics in language and cognition. *Cognitive Science*, *2*, 49–100.
- Traugott, E.** (1989). On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language*, *65*(1), 31–55.
- Traugott, E. C., & Dasher, R. B.** (2002). *Regularity in semantic change*. Cambridge: Cambridge University Press.
- Tokowicz, N., & Warren, T.** (2010). Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study. *European Journal of Cognitive Psychology*, *22*(7), 1092–1106. <https://doi.org/10.1080/09541440903325178>
- Winter, B.** (2020). *Statistics for linguistics: An introduction using R*. Abingdon: Routledge.

Cite this article: Mifka-Profozic, N., O'Reilly, D., and Guo, J. (2020). Sensitivity to syntactic violation and semantic ambiguity in English modal verbs: A self-paced reading study. *Applied Psycholinguistics* *41*, 1017–1043. <https://doi.org/10.1017/S0142716420000338>