DOI: 10.1111/modl.70033

ORIGINAL ARTICLE

The Modern Language Journal

Developing L2 semantic knowledge of English modality through concept-based language instruction: Do cognitive linguistics materials have an advantage?

Helen Zhao[1] | **James P. Lantolf**[2]

[1]School of Languages and Linguistics, The University of Melbourne, Parkville, Victoria, Australia

[2]Pennsylvania State University, University Park, Pennsylvania, USA

Correspondence
Helen Zhao, School of Languages and Linguistics, The University of Melbourne, Parkville, Victoria, 3010, Australia.
Email: helen.zhao@unimelb.edu.au

Abstract
This study investigates the pedagogical impact of integrating cognitive linguistics (CL) into concept-based language instruction (C-BLI) for teaching English modal verbs. Eighty-nine Mandarin-speaking university learners were assigned to one of three instructional conditions: C-BLI with CL-based conceptualizations (C-BLI-CL), C-BLI with non-CL functional grammar explanations (C-BLI-nonCL), and traditional definition-based instruction (control). All groups completed pre-, post-, and delayed post-tests measuring conceptual languaging and cloze-based modal accuracy. Results showed that both C-BLI groups outperformed the control group, but the C-BLI-CL group demonstrated significantly higher gains in meaningful and systematic languaging, as well as sustained accuracy in modal usage over time. In contrast, the C-BLI-nonCL group showed greater verbal explicability but experienced an accuracy decline at the delayed post-test. The findings support the compatibility of CL and C-BLI in promoting durable L2 development and suggest that instructional design grounded in conceptual coherence facilitates deeper learning. Implications are discussed for advancing meaning-oriented grammar pedagogy through concept-based and cognitively informed approaches.

KEYWORDS
cognitive linguistics, concept-based language instruction, languaging, modal verbs, sociocultural theory

Concept-based language instruction (C-BLI) was proposed by Negueruela and Lantolf (2006) as a pedagogical intervention inspired by the sociocultural theory (SCT) of mind proposed by L. S. Vygotsky (1987).[1] One of the motivations behind C-BLI was the observation that second language (L2) instruction had been predominantly centered on form-focused grammar, that is, directing learners' attention to what are often misleading or incomplete grammatical rules that fall short of fostering a profound conceptual understanding of the underlying meaning that is the foundation of grammar. Many grammatical forms, despite appearing relatively simple in structure (e.g., modals, articles, and particles), pose difficulties in acquisition due to the complex semantic meanings associated with the forms. C-BLI addresses this situation by cultivating L2 learners' systematic understanding of the meanings that underlie grammatical forms, thereby intentionally facilitating the process of L2 development.

SCT, as a psychological theory of development, incorporates disciplinary-specific scientific knowledge into its pedagogical framework. In the realm of language learning, SCT is not aligned with a specific linguistic theory. However, cognitive linguistics (CL), a theory of language that privileges conceptual meaning over form, including in the domain of grammar (Langacker, 2008; Talmy, 1988), has been strongly advocated for because of its compatibility with SCT and C-BLI, a psychological theory in which semiotic concepts play a central role in organizing higher forms of thinking (Lantolf, 2011; Lantolf & Poehner, 2014; Masuda et al., 2015; Negueruela & Lantolf, 2006; White & Masuda, 2023).[2] In addition, CL, similar to C-BLI, relies on schematic diagrams to represent conceptual information (Langacker, 2008). C-BLI also uses haptic portrayals of concepts when appropriate (see Zhang & Lantolf, 2015), and CL also acknowledges the importance of embodied cognition (Talmy, 1988). The alignment of theoretical assumptions and methodological approaches between SCT, including C-BLI, and CL justifies their potential for integration in L2 pedagogy.

Although CL does seem to align well with C-BLI, there are other meaning-based models of language that could potentially be more effective in promoting learner understanding of the meaning of particular linguistic features in an L2. Consequently, the present study was conducted in order to assess the impact of two different theories of English modal verbs and the relative impact of pedagogical explanations based on the two theories. The CL approach relies on the concept of force dynamics as proposed by Sweetser (1990) and Tyler (2012), and the other is based on the explanation of English modals formulated by the well-known semanticist Frank Palmer (1990, 2001). In the study, two intact classes were taught using principles of C-BLI (explained in the next section), with the only difference being the conceptual explanation and visual representation provided to learners. In what follows, we will first present an overview of the fundamental principles of SCT, as proposed by Vygotsky. Then we will explain how these principles are instantiated in C-BLI pedagogy, accompanied by a brief review of some of the pedagogical research that has investigated the effects of C-BLI on L2 grammatical development. Subsequently, we will review the two relevant theoretical approaches to the linguistic analysis of English modality and summarize previous empirical studies on modal verb instruction. The remainder of the manuscript will focus on the study itself, including design, data analysis, findings, and implications.

OVERVIEW OF SCT PRINCIPLES AND CONCEPTS

In this section, we explain the four principles and related concepts that Vygotsky proposed to account for the development of the higher, culturally based, human cognitive system. Each principle was formulated following a detailed investigation of the process through which humans mature from infancy into adult life (Vygotsky, 1987, 1998) as well as cross-cultural comparisons of human societies (Luria & Vygotsky, 1992).

The first principle recognizes that human psychology is comprised of processes that are both biological and cultural in origin. That is, we inherit biologically specified capacities developed over the course of evolution that are shared with other animals with robust cognitive capacities. These include natural memory, perception, attention, and emotions, as well as a rudimentary capacity for communication. At birth, these processes do not form a unified psychological system, but instead operate independently of each other. These biological instincts, which evolved over millennia as involuntary responses to repeated events in our environment (Arievitch, 2017), Vygotsky (1998) referred to as natural, or lower mental processes—lower because they are elements out of which the higher, cultural processes are constructed. As we grow into the life of our communities, and most importantly, as social language develops, the lower processes are organized into a psychological system. In this new higher functional system, the lower processes come under the voluntary control of the individual, which empowers us to deal with the unanticipated events that occur in our world (Arievitch, 2017). Moreover, as each natural process enters into the new system, its way of functioning is transformed. For example, in the new system, we no longer just perceive movement, shapes, hues, substances, and the like; rather, we perceive with meaning. Thus, when we enter a library, we see objects that in English are referred to as "books," regardless of their varying shapes, sizes, and colors. This is the result of the semiotic impact that culture, primarily through language, has on the higher functional system. Memory is also transformed in the new system as it plays a crucial role in voluntary planning of an action, including something as simple as deciding what to prepare for a meal, because, among other things, we need to remember which kinds of foods we prefer to consume and which we do not.

According to the second principle, the source of all higher psychological functions is found not in the brain but in the social relations between people, and as alluded to in the first principle, these relations are largely carried out with the mediation of language. This principle, Vygotsky captured through the concept of "sociogenesis" (Vygotsky, 1998, p. 169). The implication is that each higher function "appears on the stage twice," intermentally between people and then intramentally within our psyche (Vygotsky, 1997b, p. 106). Furthermore, it reveals the dialectical "unity of speech as a means of social behavior and as a means of individual thinking" (Vygotsky, 1998, p. 169). Thus, we use social language for psychological purposes. Vocate (1994) characterizes this social-psychological connection as a shift from an I–You communication in which subject and object are simultaneously carried out by different individuals to an I–Me communication in which the same person fulfills both roles through the process of private/inner speech.

The third principle asserts that each higher function transitions from external social reality to internal psychological reality. The transition happens through the process of internalization. In internalization, nothing is transferred from outside to inside. Rather, internalization is the "gradual construction on the part of the learner of actions equivalent to those manifested in the verbal and other behavior of others and an increasing ability to carry them out independently" (Wells, 1999, p. 117). In this way, social reality has its impact on our psyche making it always "quasi-social" (Vygotsky, 1986, p. 59). However, at the same time, the psyche has its influence on social reality because as individuals or collectives of individuals we have the capacity to change (social) reality. This happens through what Vygotsky (1997a, p. 68) called the "*doubled experience*" (italics in original). When humans act to intentionally influence or change reality, we first act psychologically in our imagination to formulate an action plan. This may entail something as complex as making the architectural plans for a skyscraper to something seemingly as simple as a speech plan to influence the behavior or understanding of an interlocutor. Once the plan is formulated, we instantiate it—act a second time in which our bodies and tools construct the building, or our vocal (or gestural) apparatus implements the speech plan. Thus, the doubled experience, unique to humans, materializes mental activity by transforming reality, and this in turn changes those who come into contact with that reality (see Ilyenkov, 2012 for a detailed discussion).

The fourth principle argues that in the process of internalization, higher mental functions pass through four developmental stages (Vygotsky, 1997a). The first consists of the natural, biologically specified process already mentioned. The second stage entails the regulation or control of mental and physical behavior mediated by others as the natural process begins to be formed into a coherent functional system. In the third stage, individuals appropriate the meanings and behaviors used by others; that is, I-Me speech emerges from I-You speech. At first, I-Me speech, for which Vygotsky borrowed

TABLE 1 Pedagogical phases of concept-based language instruction.

Phase	Description	Pedagogical rationale
1. Pre-understanding	Teacher elicits conceptual knowledge prior to instruction—OBA*	Identify and make visible starting point for instruction
2. Concept presentation	Teacher explains language concept	Comparison between prior and new knowledge
3. Materialization	Teacher concretizes (2) as drawing, graphic, diagram, haptic experience as SCOBA**	Avoid rote memorization through holistic representation of concept. Mediate performance
4. Languaging	Students produce external speech/writing	Reveals understanding. Transforms external to internal knowledge. Shift from SCOBA to reliance on self.
a. Communicated	Students explain concept and its use to others	Speech to begin internalization process. Speech becomes psychological
b. Dialogic	Students explain concept and its use to the self	Transform understanding to private speech
5. Performance	Students engage in goal-directed communicative activity using concept	Ability to embed conceptual knowledge in purposeful practical activity
6. Internalization	Students use concept without external mediation of SCOBA	Concept is generalizable and functionally useful
a. New OBA	Teacher and students compare prior knowledge to new conceptual knowledge	Link conceptual development and communicative performance

Note: The arrow indicates that languaging can spread across more than one phase.
Abbreviations: OBA, orienting basis of action; SCOBA, schema for orienting basis of action.

Piaget's term "egocentric speech" but with a different interpretation, appears social in form, but it is in reality psychological in function (Vygotsky, 1987). Over time, egocentric speech becomes increasingly abbreviated in form and more condensed in meaning. For instance, someone assembling a puzzle may utter, "the red one" in order to direct their attention to the next piece to be inserted (see Wertsch & Hickmann, 1987 for a detailed analysis of the process). Eventually, psychological speech disappears but continues to operate subvocally as the person develops the ability to independently mediate their own thinking and behavior without the need for external mediation. However, when difficulties arise, speech may again become audible as the person re-accesses a previous mediational stage (see Frawley & Lantolf, 1985).

PEDAGOGICAL FRAMEWORK

The four theoretical principles and related concepts provide the foundation for C-BLI. In Table 1 we outline the six phases associated with C-BLI. In the table, each phase is labelled, followed by a description along with its pedagogical rationale.

Vygotsky did not explicate a specific framework for how the theory could be brought into educational practice. Instead, a psychologist and educator, Piotr Gal'perin, whose work was influenced by Vygotsky, formulated a set of specific pedagogical procedures, Systemic Theoretical Instruction (STI), which was designed to intentionally promote development through schooled instruction (see Gal'perin, 1969, 1992; Engeness, 2021; Haenen, 1996; Talyzina, 1981). The phases of C-BLI described in Table 1 are based in large part on Gal'perin's model. However, Gal'perin's view of language, along with others who employed his model (e.g., Carpay, 1974; Kabanova, 1985; Van Parreren, 1975) approached language as a purely formal system rather than as derived from conceptual meaning as in CL. Consequently, contemporary SCT-L2 researchers have maintained the basic structure of

Gal'perin's developmental model but have modified it, including its label, to more appropriately take account of CL theory.

Beginning with the orienting basis of action (OBA), in phase 1, the goal of C-BLI is the development of both conceptual knowledge and the ability to use this knowledge in communicative activities. Conceptual knowledge empowers learners to understand how they can create language-based signs that manifest their own communicative agency rather than adhere to pre-determined rules that give the impression that language is about deploying right and wrong forms.[3] The point of OBA is to allow teachers access to what learners already know or think they know about a language feature. At the same time, it raises learner awareness of what they know by making it visible to them. This is achieved by asking learners, either individually or as a class, to externalize in speech or in writing their knowledge of a relevant language feature. Phase 1 OBAs are revisited in the final phase 6a in order to assess the change in conceptual understanding of the relevant feature. Any change in conceptual knowledge should be accompanied by a change in communicative performance.

In phase 2, the teacher presents a concept-based explanation of the relevant language feature. With the exception of four studies by Negueruela (2003), Yáñez-Prieto (2008, 2014), and Lantolf and Tsai (2018), which drew their meaning-based explanations of grammatical features from other sources, the majority of grammar-based C-BLI studies have formulated their pedagogical explanations on the basis of CL theory. Gánem-Gutiérrez (2016), Kissling and Muthusamy (2022), and Kissling (2023) based their explanations of Spanish tense/aspect on CL, as did Poehner and Infante (2017) and Benhamlaoui and Gánem-Gutiérrez (2022) with regard to English tense and aspect. Lee (2016) adopted the CL explanation of the relationship between literal and metaphorical meaning to provide instruction on English particle verbs. Jacobsen (2015) developed an explanation for English conditional constructions on the basis of CL, as did Zhao (2026) with regard to English modal verbs. Arnett and Deifel (2015) and Buescher and Strauss (2018) relied on CL for their explanations of prepositions in German and French, respectively. Finally, a good deal of C-BLI research on Asian-language instruction has used CL as the explanatory framework in phase 2: Masuda and Labarca (2015) focused on a set of Japanese particles; Xi (2022) carried out a study of Chinese word order; and as mentioned in endnote 2, Masuda et al. (2025) included instruction on Japanese grammatical features, including passives, causatives, tense/aspect, and benefactives in their book-length publication.

Phase 3, the schema for orienting basis of action (SCOBA) is challenging, as it requires presentation of the concept explained verbally in phase 2 in some kind of visual form, either as a drawing, diagram, figure, or flowchart. In some cases, the knowledge can be presented through teacher gestures (Smotrova & Lantolf, 2013), or even videos (Kissling & Muthusamy, 2022). The SCOBA is intended as a holistic representation of a concept and reduces the likelihood of rote memorization without understanding the explanation in Phase 2. It also mediates learner performance during their initial phase 5 activities.[4] Over time, with increased experience using the concept in communicative activities, learners are expected to decrease their reliance on the SCOBA as an external form of mediation.

Phase 4 introduces the process of languaging—the psychological rather than social function of speaking or writing (Swain, 2006).[5] It reflects the theoretical principle that social, communicatively oriented speech takes on a psychological function (i.e., I–You > I–Me) and is intended to "wean" learners off object mediation provided by the phase 3 SCOBA as well as other mediation provided by teacher prompts, hints and clues, and onto self-mediation. Phase 4 is segmented into two subphases. In phase 4a, learners explain their understanding of the concept to someone else, including the teacher or a peer (I–You, with a psychological function), and in phase 4b, they explain the concept to themselves (I–Me). This can be done either orally or in writing. If it is carried out orally, the languaging should be recorded for analysis by the teacher and the learner. Languaging also figures in phase 5 (indicated by the arrow) as learners are asked to overtly explain to others and to themselves their use of a concept in communicative activities. Negueruela (2003) and Yáñez-Prieto (2008, 2014) provide excellent examples of languaging as learners confront new conceptual knowledge, as well as post-activity languaging in phase 5.

Phase 5 incorporates activities that engage learners in practical communicative activities. These may include scenarios, or mini dramas, as proposed by Di Pietro (1987), and implemented in van Compernolle's (2014) L2 French project. It may include task-based learning activities. It may involve learners in writing activities as in Ferreira and Lantolf (2008) and Yang and Sang (2023), or reading and analysis of texts (Urbanski, 2023).

Phase 6 indicates the attainment of internalized self-mediation in knowledge and use of a concept. It is manifested through communicative performance, which does not rely on external mediation, either as other support, object support from the SCOBA, or self-support in the form of overt private speech. Learners can also recognize the change in their initial phase 1 OBA and their new phase 6 OBA. This is not always easily achieved, especially for learners with a previous history of traditional form-focused instruction, as attested in the work of Negueruela (2003) and Yáñez-Prieto (2008, 2014), who documented the struggle learners exhibited as they attempted to reconcile their original understanding of a grammatical form with their new conceptual knowledge.

LINGUISTIC APPROACHES TO THE ANALYSIS OF ENGLISH MODALITY

To situate the rationale for our instructional design, this section reviews two theoretical approaches to English modality and highlights how each framework conceptualizes modal meaning in ways that differ in their pedagogical usefulness for the present study. Modality in English encompasses the grammatical and lexical means by which speakers express attitude, degree of necessity, possibility, permission, ability, willingness, or likelihood regarding a particular action or situation (Carter & McCarthy, 2006). Understanding modality is crucial for effective communication, as it allows speakers to convey their attitudes, beliefs, and expectations regarding actions and events. The primary means of expressing modality in English is through modal verbs (e.g., *can, could, may, should*), which pose persistent challenges for L2 learners due to their semantic complexity (Celce-Murcia & Larsen-Freeman, 1999). Modal verbs are polysemous and highly context-sensitive (Palmer, 2001), with subtle differences in meaning that are often difficult to systematize (Torres-Martínez, 2019; Tyler, 2012). Modal verbs often lack direct equivalents in the learner's native language. Translating modal meanings from one language to another can be challenging, as not all languages have modal verbs that align precisely with those in English.

The traditional approach to explaining and teaching English modal verbs often frames them in terms of their communicative functions, particularly within the context of speech acts. Modal verbs are commonly grouped according to broad functional categories. For example, "*can, could, may, might*" are grouped under "permission/possibility/ability" in Biber et al. (1999), without further differentiation. Such an approach can easily mislead learners into interpreting modals within the same category as interchangeable or synonymous (Tyler, 2012). While Celce-Murcia and Larsen-Freeman (1999) also adopted a speech-act-based approach (e.g., "*will*, *should*, *may*, and *might*" as the modals for making predictions), they attempted to clarify distinctions in terms of degrees of probability. However, such probabilistic explanations often fail to capture the nuanced semantic differences among modals within each group, leaving learners with an incomplete understanding (Tyler, 2012).

More comprehensive accounts of English modality have moved beyond surface-level speech act categorizations. Among them, Palmer (1990, 2001) offered one of the most in-depth and systematic analyses of modality. He distinguished three types of modalities: deontic modality (influencing social actions), epistemic modality (judgment about the truth of the proposition), and dynamic modality (expressing ability, volition). Within each type, Palmer identified two degrees of modality: possibility and necessity. For example, the distinction between *may* and *must*, in both their deontic and epistemic meanings, can be explained in terms of the difference between possibility and necessity. Deontic possibility as a performative action gives rise to the permission meaning of *may*, whereas deontic necessity implies the speaker's authoritative position to impose a social obligation (*must*). Possibility in epistemic judgments indicates speculation (*may*), whereas epistemic necessity can be interpreted as deduction (*must*: It is necessary to conclude that… usually based on strong evidence).

Palmer (1990, 2001) conceptualized these categories within a two-dimensional matrix: the type of modality mapped along the horizontal axis and the degree of modality along the vertical. The distinction between possibility and necessity within the deontic and epistemic modalities is relatively straightforward, but how this distinction applies to dynamic modality is much less clear (Depraetere, 2015). Palmer (1990) himself acknowledged the limitations of this framework, noting that it could only account for a "rather loosely structured set of relationships" (p. 49).

Cognitive linguists (Johnson, 1987; Sweetser, 1990; Talmy, 1988; Tyler, 2012) offered an alternative account of the semantic system of English modality using CL concepts such as embodiment, schema, polysemy, and metaphor. They explained modality from the perspective of force dynamics (Talmy, 1988), a part of the human conceptual structure related to the exertion of force. Force dynamics encompasses physical force derived from bodily experiences of motion, manipulation, and pressure (Johnson, 1987), as well as psychological and social forces that influence (non-)actions. The interpersonal, psychosocial force underlies the root meaning of modality, and the epistemic meaning is derived from the root meaning through metaphorical transfer (Sweetser, 1990). For instance, the root meaning of *must* is represented as a compelling, irresistible force from an external source, while its epistemic meaning is derived through metaphorical extension from its source domain (psychosocial) to its target domain (reasoning), that is, the compelling evidence forces me to conclude that… (Tyler, 2012).

The CL approach, therefore, aims to explain the full range of modal meaning through embodied experience and metaphorical mapping. While this perspective offers a theoretically coherent and cognitively motivated framework, it is not without limitations. Some researchers have argued that metaphorical analysis struggles to account for certain epistemic uses, particularly those involving modals such as *will*, *would*, *might*, *should*, and *ought to*, which may resist straightforward mappings from psychosocial force schemas[6] (Langacker, 1991; Mulder, 2021).

Taken together, these approaches illuminate complementary aspects of modal meaning, but they differ in their pedagogical affordances. Speech-act classifications are accessible for teaching but risk oversimplification; Palmer's semantic framework provides greater systematicity, yet its categories are descriptive and offer limited guidance for concept development. In contrast, the CL account explains the coherence of modal polysemy through force dynamics and embodied schemas, making it particularly compatible with a meaning-focused, concept-driven pedagogy.

L2 INSTRUCTION OF ENGLISH MODALITY

Several studies have investigated effective methods for teaching English modals to classroom-based L2 learners. Among them, Tyler et al. (2010), also reported in Tyler (2012), conducted a quasi-experimental study that integrated CL concepts into explicit modal verb instruction. Their findings showed that adult ESL learners who received CL-inspired instruction moderately outperformed those taught via a traditional speech act approach. The CL instruction explicitly introduced core concepts (such as embodiment and metaphor) by drawing on learners' bodily experiences (e.g., enacting force dynamics through experiencing physical forces), using schematic diagrams to visualize force interactions among modals, and using diagrams and concrete examples to demonstrate the metaphorical connection between root and epistemic meanings. Although Tyler et al. (2010) did not explicitly situate their instruction within a C-BLI framework, their pedagogical design was closely aligned with several C-BLI principles. Specifically, their instruction was explicit and emphasized semantic understanding. They initiated the process with explicit teaching of CL concepts materialized as schematic diagrams, which served the purpose of a C-BLI's SCOBA that guided learners in a subsequent paired learner discussion task. In comparison, their speech-act instruction largely followed the approach of task-based language teaching (TBLT), which emphasized communicative functions of modals through meaning-based, collaborative tasks without explicit attention to underlying conceptual structures[7].

Suethanapornkul et al. (2014) conducted another study on CL-inspired modal verb instruction. The CL instruction was delivered through a computer-based game that incorporated concrete images and virtual characters. Learners were expected to acquire CL concepts implicitly, as no explicit information on force dynamics and metaphorical extension was provided. Instead of using schematic diagrams, force dynamics was taught by instructing learners to perform actions that manipulate the virtual movements of game characters. However, the findings showed that this CL-based instruction did not lead to greater learning gain compared to a control group receiving speech-act-based instruction. This result suggests that the abstract and complex nature of CL concepts may not lend themselves well to implicit acquisition. Supporting this interpretation, Wang and Zhao (2024a, 2024b) found that schematic diagrams—commonly used in CL-inspired instruction—can interfere with mental imagery in second language lexical decision and sentence comprehension tasks when not accompanied by explicit explanation. Their findings further underscore the importance of making CL concepts pedagogically accessible through explicit instruction, particularly by using schematic diagrams, concrete examples, and guided exercises that clearly illustrate underlying semantic structures.

These previous studies have provided valuable insights into integrating CL concepts in modal verb instruction in classroom settings. However, they also exhibit notable limitations. First, both studies lack a clearly articulated pedagogical framework. Tyler et al. (2010) described their CL instruction as "teacher-fronted" (p. 119), yet the procedural features of their design align with core C-BLI principles—namely, explicit concept presentation, schematic mediation, and guided learner discussion. By contrast, their speech-act condition followed a task-based, meaning-focused approach characteristic of TBLT, emphasising communicative outcomes without explicit conceptual mediation. Suethanapornkul et al. (2014), meanwhile, did not provide sufficient methodological detail about the instructional procedure. When the pedagogical foundation is underspecified, it becomes difficult to discern whether the (in)effectiveness of instruction stems from the quality of the CL materials themselves or from their compatibility with the instructional procedure. Second, instruction effects were solely measured by learners' accuracy in a multiple-choice cloze test before and after instruction, with no delayed posttests assessing learning retention or considering other aspects of learning, such as the development of semantic understanding. Such limitations restrict the conclusions that can be drawn about the long-term and conceptual impact of CL-based instruction.

Qin et al. (2023) represent a recent attempt to apply C-BLI to the teaching of English modals to adult EFL learners. Concepts were taught using SCOBAs (visual diagrams) based on the traditional speech act approach. Learning outcomes were assessed by qualitatively analyzing learner-generated visualizations and the quality of oral and written languaging. Their results showed that after instruction, learners moved away from traditional rule-of-thumb information and relied more on conceptual knowledge. While this new empirical work produced promising findings, it can be improved in several ways. First, as previously discussed, the traditional speech act approach to modality has been widely critiqued for presenting modals as overlapping and context-dependent, without revealing the semantic relationships that structure the modality system (Torres-Martínez, 2019; Tyler, 2012). As a result, learners may view modals as interchangeable and unpredictable, which may reinforce rote memorization rather than conceptual mastery (Zhao, 2025). Instructional materials derived from this approach are thus inherently limited in their capacity to convey the full conceptual complexity of English modality. Second, the SCOBAs used in Qin et al.'s study treated modal verbs in isolation, presenting separate visual representations for each, without offering a unified conceptual framework. Consequently, learner languaging was analyzed on a per-modal basis, which limited insight into whether learners had developed a systematic understanding of the modal verb system—a key aim of C-BLI pedagogy (Negueruela, 2003; Negueruela & Lantolf, 2006). Third, although the researchers administered a modal verb test to the participants, the results were not included in their article, which focused on languaging and SCOBA production. This omission leaves open questions about the extent to which conceptual gains translated into improved performance on more conventional language assessments.

Therefore, the current study addresses the limitations of the previous modal verb projects as it investigates whether the theoretically predicted compatibility between C-BLI and CL (Lantolf, 2011; Lantolf & Poehner, 2014; Negueruela & Lantolf, 2006) yields optimal outcomes for the development of English modals with regard to knowledge and performance. To address the compatibility question, two quasi-experimental conditions were implemented, both of which received English modality instruction through the same C-BLI procedure. The primary difference between the two C-BLI groups was the quality of conceptual knowledge (phase 2) and affiliated SCOBA (phase 3). The C-BLI-CL group was given modal verb explanations and SCOBAs derived from materials created with CL assumptions (Sweetser, 1990; Tyler et al., 2010), whereas the C-BLI-nonCL group received materials grounded in Palmer's (1990, 2001) functional-semantic framework, which systematically distinguishes deontic, epistemic, and dynamic modality based on meaning rather than communicative function. Palmer's framework was selected because it offers a widely recognized, theoretically motivated account of modal meanings that remains compatible with explicit instruction but lacks the embodied, schematic grounding provided by CL. There was an additional control group that received traditional definition-based instruction, which involved dictionary-style explanations and example-based pattern memorization without conceptual or visual mediation. Specifically, the study addresses the following research questions:

1. How do C-BLI-CL and C-BLI-nonCL influence the quality of learner languaging about modal semantics?
2. How do C-BLI-CL and C-BLI-nonCL influence learner performance in modal verb usage (*must*, *should*, *may*, and *will*) across modality functions (epistemic and non-epistemic)?
3. Does traditional definition-based instruction lead to comparable outcomes in learner languaging and modal verb performance relative to the C-BLI instruction groups?

METHODOLOGY

Participants

Eighty-nine Mandarin-speaking learners of English participated in the current study. All participants, aged 17–20 years (mean age = 18.58, SD = 0.58), were second-year undergraduate students from a university in Shenzhen, a city in southeast China. Not all participants completed all the data collection sessions (particularly in the control group); after attrition, 76 participants (46 male and 30 female) completed all the sessions and were included in the final analysis. Their demographic information and language history were collected via selected items from the Language History Questionnaire (LHQ 2.0) (Li, Zhang, Tsai & Puls, 2014). Their onset ages of learning English sub-skills were as follows: speaking (mean = 7.25, SD = 2.48), reading (mean = 7.25, SD = 2.35), and writing (mean = 8.33, SD = 2.57). They had an average of 11.39 years (SD = 2.27) of learning English. They spent a mean length of 1.92 hours (SD = 1.59) of engaging in English-related activities per day. We collected their scores on the College English Test Band 4 (CET-4) (Zheng & Cheng, 2008), a national English test in China that all the participants were required to take, as the index of their general English proficiency. The test, with a total score of 710 points, is composed of listening comprehension, reading comprehension, cloze/error correction and translation, writing, and speaking. Our participants had a mean CET-4 score of 571.16 (SD = 32.10). CET4 has not been formally benchmarked to the CEFR levels (Council of Europe, 2001). Our experience with the cohort suggests that they are generally at the B1 level on the CEFR scale.

Research design and data collection procedure

The study adopted a quasi-experimental design of classroom-based instruction. The participants belonged to three intact parallel classes of College English. One class was for computer science majors, and the other two were for engineering majors. The three classes were randomly assigned to one of the three conditions: C-BLI-CL ($n = 29$), C-BLI-nonCL ($n = 32$), and the control group ($n = 15$). The three groups had a comparable level of English proficiency based on the CET-4 scores ($F(1,59) = 1.084$; $p = 0.302$). They also showed no significant differences in terms of the onset ages of learning English, years of learning English, and the amount of daily English use.

The study adopted a pretest-posttest-delayed posttest design to examine both immediate and retained learning effects. All the groups completed three tests to measure their languaging and usage of four target English modals (*must, should, may, will*): Test I (pretest before instruction), Test II (posttest immediately after instruction), and Test III (delayed posttest three weeks after instruction). These four modals were selected because they represent core force dynamics configurations: *must* and *may* involve external force, *will* reflects internal force, and *should* captures the interaction between internal and external forces. They also span key contrasts in modality type (root vs. epistemic) and degree (necessity vs. possibility). While not exhaustive, this set allows learners to engage with fundamental conceptual distinctions underpinning English modality.

Teaching interventions were conducted during regular class time. The two C-BLI instruction groups received a total of 6 hours of instruction over 2 weeks, taught by the first author of the article. The control group received instruction from their regular College English course instructor, who followed teaching materials prepared by the first author. These materials consisted of dictionary entries for the four target modal verbs (*must*, *should*, *may*, and *will*) taken from the *Oxford Advanced Learner's Dictionary* (Hornby, 2005), accompanied by sample sentences illustrating their typical uses. To ensure comparable instructional duration across groups, the control group also completed additional reading comprehension activities from their regular College English curriculum. These activities were unrelated to the target modals but maintained equivalent class time and engagement across conditions.

The current C-BLI procedure, applied to both the CL and non-CL instructional conditions, included five phases, as described below.

Phase 1: Pre-understanding. The instruction began with a 15-minute introduction of the target modal verbs, designed to orient participants toward the learning goal. Learners interpreted short sentences containing the target modals, explained what each modal expressed, and discussed their interpretations in small groups. This activity revealed variation and uncertainty in how they mapped form to meaning and provided a shared basis for subsequent conceptual instruction.

Phase 2: Concept presentation. Participants were introduced to the overall instructional framework and the full set of SCOBAs for the target modals, but without detailed explanations. The purpose of this phase was to familiarize learners with the overall conceptual landscape and the sequence of instructional activities to follow. The presentation helped them develop an initial global understanding of what they would be learning and how the instructional process would unfold.

Phase 3: Materialization. Participants were then given detailed explanations of the SCOBAs for English modality. These included flowcharts that represent the conceptual structure of the target modals, tailored separately for the CL group (Figure 1) and the non-CL group (Figure 2). It is important to note that both the CL and non-CL conditions targeted the same modal meanings and covered an equivalent semantic scope. Although the CL flowchart appears more streamlined, it integrates root and epistemic senses within a unified force-dynamic schema, whereas the non-CL flowchart presents them separately; thus, the difference reflects conceptual organization rather than content coverage. For the CL group, the SCOBAs also incorporated

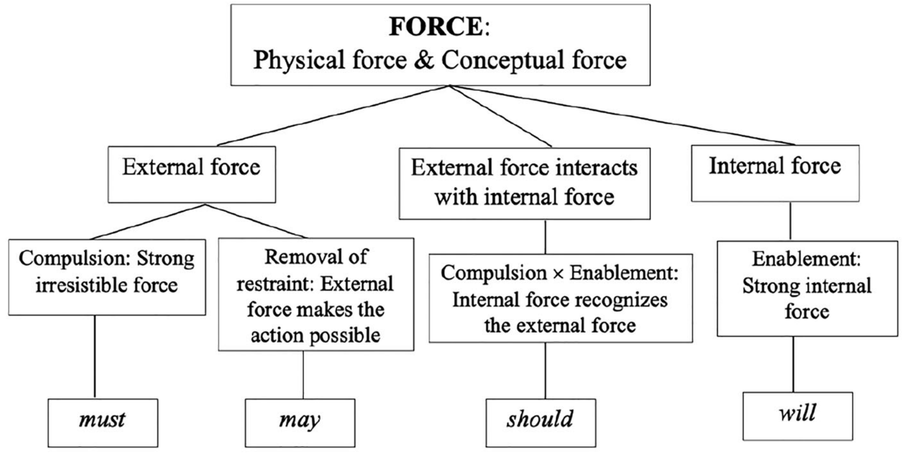

FIGURE 1 Flowchart for the C-BLI-CL condition (based on Sweetser, 1990; Tyler et al., 2010).

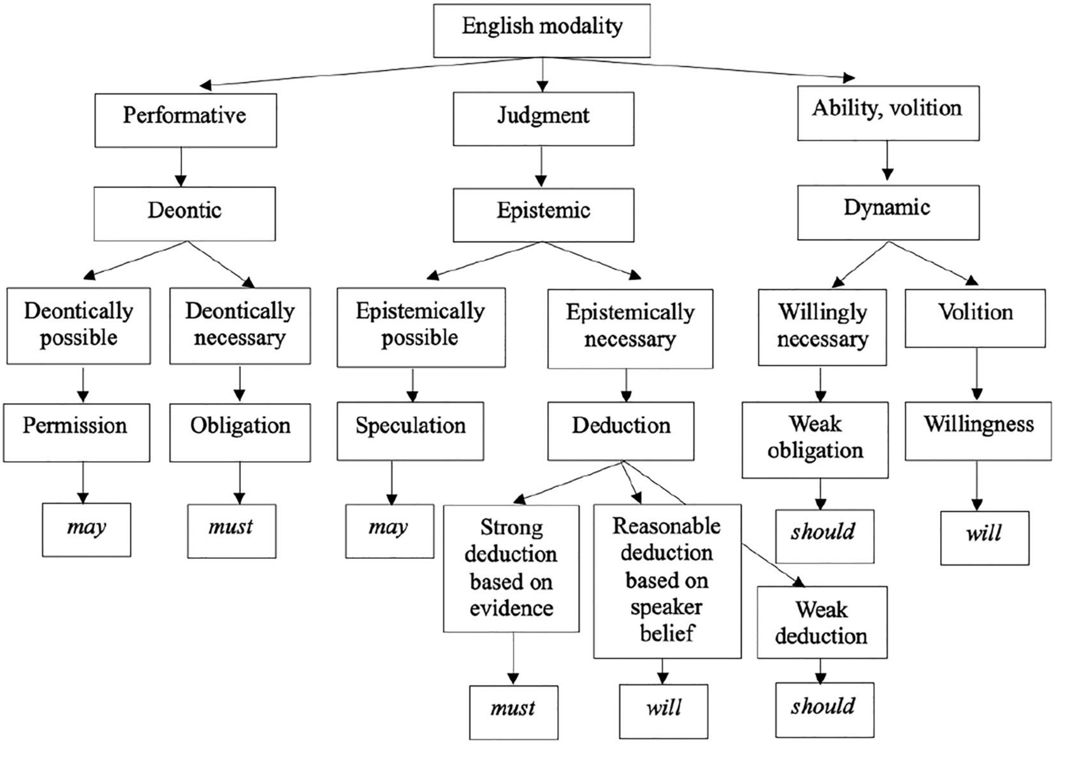

FIGURE 2 Flowchart for the C-BLI-nonCL condition (based on Palmer, 1990, 2001).

schematic diagrams and example sentences that illustrate the semantic meanings of the modals (Appendix A). For the non-CL group, the SCOBAs included example sentences aligned with each semantic function (Appendix B). The instructor elaborated on key concepts (e.g., force dynamics for the CL group; modality type and degree for the non-CL group) by using the flowcharts and diagrams as visual scaffolds.

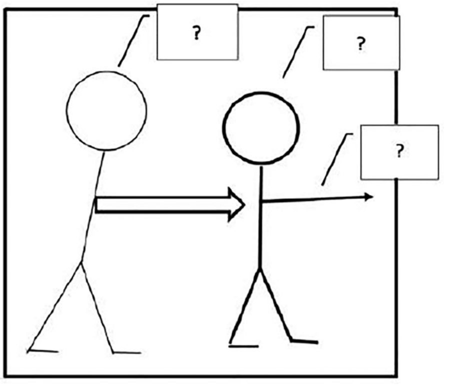

FIGURE 3 Sample material used for oral languaging based on a cognitive linguistics diagram.

Phase 4: Languaging. After the materialization phase, learners engaged in languaging activities designed to reveal and gradually internalize their understanding of the target modals. This phase was implemented in two subphases: (a) *Communicated languaging*: Learners were provided with sample sentences, initially containing the target modals and then without them. They worked in pairs to explain the meaning of the modals to each other using the SCOBAs. Specifically for the CL group, learners were given sample sentences and tasked with identifying FORCE elements from the sentences to match the key features in the schematic diagrams (refer to Figure 3 for an illustration). (b) *Dialogic languaging*: Learners repeated this SCOBAs-guided languaging procedure, this time verbalizing to themselves. They were instructed to audio-record both paired and self-languaging and submit the recordings to the first author immediately after completion. This self-oral languaging was followed by self-written languaging, a paper-and-pencil take-home assignment consisting of cloze sentences (distinct from assessment items). They were asked to insert appropriate modals and explain their meanings, drawing on the SCOBAs for support. CL group participants were additionally encouraged to draw their own diagrams to represent modal meanings. The non-CL group was not asked to produce diagrams, as Palmer's (1990, 2001) framework does not utilize schematic visual representations but instead organizes modal meanings through categorical distinctions and degree scales. A feedback session was held in class following submission of each assignment, allowing learners to refine explanations and consolidate their conceptual understanding.

Phase 5: Performance. Learners engaged in goal-directed oral activities in academic English contexts, applying the mediated conceptual tools for modal meaning in spontaneous, meaningful communication. Short role-plays, such as giving research advice, discussing experimental results, and negotiating project deadlines, were used to elicit target modals in context, enabling learners to apply the conceptual tools from the SCOBAs in academically oriented discourse.

Instruments

Each of the three test sessions (pretest, posttest, and delayed posttest) included two measurements: a written languaging test and a written cloze test. They were designed to assess participants' quality of languaging and their accuracy in modal verb usage, respectively. The languaging test used the same prompt across the test sessions: "*Explain in your own words the relationship between the following modal verbs: must, should, may, and will. And describe when and how each of these modal verbs is used*." Participants could respond in either English or Chinese. The languaging data were coded (see Appendix E) for the number of meaningful units and the number of incorrect units. To approximate the density of conceptually meaningful content in learner languaging, we calculated a *meaningfulness ratio*, defined as the number of meaningful units divided by the sum of meaningful and incorrect units.

Following Negueruela (2003) and García (2017), we also coded languaging for five qualitative features: "*explicability* as the ability to produce a definition; *generality* (based on semantic criteria, and not functional or perceptual features); *abstract* (essential features of the concepts are contained in the definition); *systematic* (coherent and hierarchical taxonomic relations to other concepts are established)" (Negueruela, 2003, p. 271). *Ownership* reflects the learner's agency in consciously and intentionally using the instructed grammatical concepts as mediational tools for language use and development in language activities (García, 2017). This feature was coded only for the posttests. The essential languaging points expected for the C-BLI-CL condition (Appendix C) and the C-BLI-nonCL condition (Appendix D) were aligned with their respective SCOBAs. For explicability, we first coded the C-BLI-CL group's awareness of root and extended meanings separately, and the C-BLI-nonCL group's awareness of deontic, epistemic, and dynamic functions separately; these were then combined into a composite score. Abstractness was first coded for each target modal and subsequently aggregated into composite scores. The full coding scheme, with examples for each languaging feature, is provided in Appendix E. Importantly, the languaging task was not designed to assess learners' ability to reproduce instructional explanations verbatim, but rather to evaluate their ability to reconstruct and apply conceptual understanding using their own linguistic resources. The inclusion of the ownership category (García, 2017) was intended precisely to capture learner agency in appropriating and recontextualizing the mediated conceptual tools introduced during instruction.

Languaging data were coded independently by two raters: the first author and a research assistant (a Linguistics PhD student with training in semantics). Prior to full coding, the research assistant received training using 30% of the dataset pre-coded by the first author. Interrater reliability was then assessed on 25% of the languaging data, yielding a high level of agreement (> 90%) before the research assistant proceeded to code the remaining data. Any discrepancies that arose during this process were resolved through discussion.

The written cloze test (Cronbach's $\alpha = 0.6$) included a complete pool (Appendix F) of 96 target sentences (12 sentences × 2 modality function × 4 modals) and 30 filler sentences featuring non-target modals (*can*, *could*, *would*, and *might*). All sentences were adapted from the academic written section of the Corpus of Contemporary American English (Davies, 2008–), to align with learners' proficiency levels and the genre expectation of the College English curriculum, ensuring syntactic and contextual meaningfulness at the discourse level. The 126 sentences were evenly and randomly distributed across the three tests (pretest, posttest, and delayed posttest), with each version containing 42 unique items (32 target + 10 fillers) and no overlap between test versions. This design ensured balanced representation across stimulus conditions and comparable difficulty across tests, while preventing item repetition or practice effects. None of the test items overlapped with instructional or take-home materials. Participants' responses on the target items were coded for accuracy of response.

Data analysis

While SCT research often employs qualitative, microgenetic methods (e.g., Negueruela, 2003), the present study adopts a quantitative approach to examine developmental trends across groups and timepoints. The quantitative analysis was designed to capture developmental trajectories across the three testing points while controlling for baseline differences and individual variability. For the languaging and cloze data, a mixed-effects modeling approach was used to estimate learning gains across conditions and test phases, with pretest performance included as a covariate. This approach accommodates repeated-measures data, accounts for inter-individual differences, and provides a robust method for examining condition-specific growth over time. Such modeling aligns with the developmental orientation of SCT by quantifying how learners' conceptual performance evolves

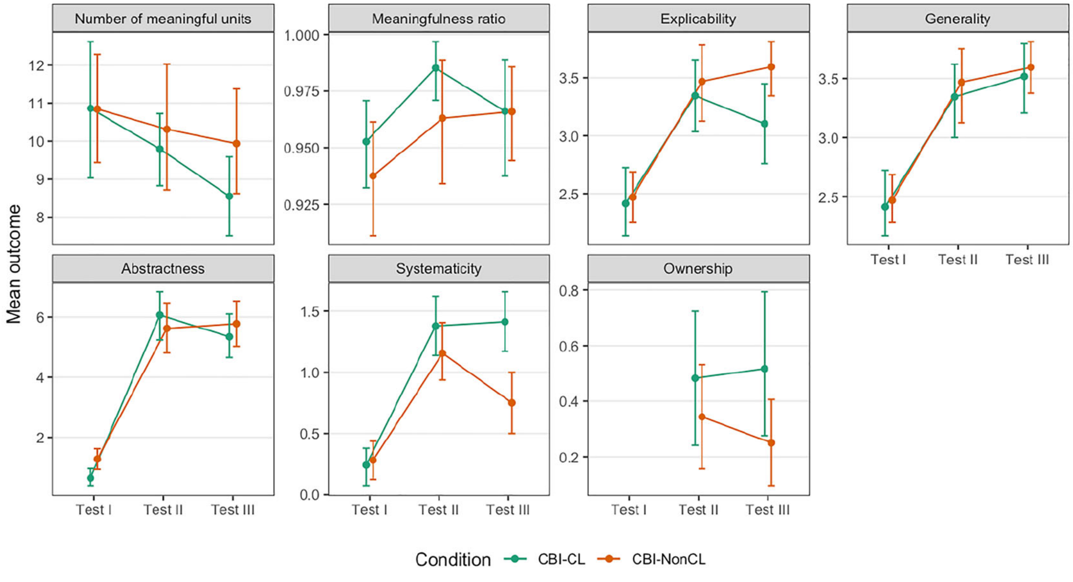

FIGURE 4 Error bars showing means and confidence intervals for languaging features across tests. *Note*: Error bars show individual bootstrap 95% confidence intervals for the mean. [Color figure can be viewed at wileyonlinelibrary.com]

through mediated instruction and complements prior qualitative findings by testing the theoretical compatibility of C-BLI and CL under controlled, replicable conditions.

Data analysis employed R software (version 4.4.0; R Core Team, 2024), utilizing the lme4 package (Bates et al., 2022) for constructing linear mixed-effects models. P values were calculated using the lmerTest package (Kuznetsova et al., 2017). The statistical models built to address the specific research questions and the model formulas are detailed in the Results section. Model convergence and adherence to statistical assumptions were verified. Tukey correction for pairwise comparison was applied using the emmeans package (Lenth et al., 2023). Effect sizes for RTs were reported as Cohen's *d* (1977) and interpreted following the guidelines by Plonsky and Oswald (2014): for within-subjects contrasts, small, medium, and large effect sizes correspond to values of 0.60, 1.00, and 1.40, respectively; for between-subjects contrasts, the thresholds are 0.40, 0.70, and 1.00. Graphics were generated with the ggplot2 package (Wickham, 2016).

RESULTS

Research question 1: Quality of learner languaging (C-BLI-CL vs. C-BLI-nonCL)

To address the first research question, we focused on analyzing the written languaging test data of the two C-BLI groups (C-BLI-CL and C-BLI-nonCL). Figure 4 visualizes the two groups' descriptive means of the analyzed languaging features. Below are the inferential statistics results of the specific languaging features.

Number of meaningful units

A linear mixed effects model[8] was used to assess the effect of condition (C-BLI-CL vs. C-BLI-nonCL) on the number of meaningful units in the languaging posttests. Condition and test number

TABLE 2 Pairwise comparisons of languaging meaningfulness ratio between successive tests.

Condition	Test II/Test I Odds ratio	95% CI	*z*	*p*	Test III/Test I Odds ratio	95% CI	*z*	*p*	Test III/Test II Odds Ratio	95% CI	*z*	*p*
C-BLI-CL	3.5	1.3, 9.6	2.39	0.017	1.7	0.72, 3.9	1.20	0.232	0.48	0.16, 1.5	−1.27	0.205
C-BLI-nonCL	1.8	0.87, 3.7	1.59	0.113	2.1	0.99, 4.5	1.94	0.052	1.2	0.51, 2.7	0.40	0.691

Note: An odds ratio greater than 1 indicates greater languaging meaningfulness ratio on the later test.

were included in the model as fixed effects, with an interaction term. The number of meaningful units at the pretest (Test I) was used as a covariate. Participant was included as a random effect. Examination of the residual plot showed adequate model fit. The results showed no significant difference between the two conditions in the posttest (estimate = −0.5, [−2.4, 1.3], $t = -0.56$, $p = 0.579$) and the delayed posttest (estimate = −1.4, [−3.3, 0.5], $t = -1.48$, $p = 0.143$).

Languaging meaningfulness ratio

A generalized linear mixed model (GLMM)[9] was fitted to the proportion of meaningful units in languaging. A binomial response distribution was used, with the number of meaningful units as the events of interest, and the number of attempts being the total number of units (meaningful and incorrect). Condition and test number were included in the model as fixed effects, with an interaction term. The pretest meaningfulness ratio was used as a covariate, with an empirical logit transform used to ensure it was on the same scale as the model. Participant and test (nested within participant) were included as random effects. Examination of the residual plot showed a good model fit. The results showed no significant difference between the two conditions in the posttest (estimate = 2.0, [0.5, 7.1], $z = 1.04$, $p = 0.298$) and the delayed posttest (estimate = 1.0, [0.3, 3.1], $z = -0.07$, $p = 0.943$).

A GLMM[10] was also used to assess the change in meaningfulness ratio between testing occasions. Condition and test number (nested within participant) were included in the model as fixed effects, with an interaction term. Participant and test were included as random effects. Test number showed a significant main effect ($F = 5.03$, $\chi^2 = 10.06$, $p = .007$). Tukey-adjusted pairwise comparisons were conducted to assess the change in meaningfulness ratio between testing sessions (Table 2). The results showed that C-BLI-CL promoted a significant increase in meaningfulness ratio from the pretest to the posttest (effect size reflected by the odds ratio = 3.5, meaning that the odds of meaningful languaging at the posttest are estimated to be 3.5 times as great compared to that of the pretest in this condition), whereas C-BLI-nonCL did not yield a difference in meaningfulness ratio from the pretest to the posttest. This finding suggests that CL-based conceptual materials more effectively supported learners in producing semantically meaningful language than the non-CL alternative.

Explicability, generality, abstractness, and systematicity of languaging

A linear mixed-effects model[11] was used to assess the effect of condition (C-BLI-CL vs C-BLI-nonCL) on these languaging features. Condition and test number were included in the model as fixed effects, with an interaction term. The pretest score (Test I) was used as a covariate. Participant was included as a random effect. Table 3 summarizes the outputs from pairwise comparisons based on the model. While the two groups scored similarly in most of the comparisons, two comparisons turned out to be statistically significant. First, C-BLI-nonCL outperformed the C-BLI-CL condition in terms of explicability in the delayed posttest with a medium effect size ($d = 0.763$). Second, the C-BLI-CL group greatly outperformed the C-BLI-nonCL group in terms of systematicity in the delayed posttest,

TABLE 3 Comparisons of languaging features between conditions.

	Difference in means (C-BLI-CL—C-BLI-nonCL)[a]					
Test	**Estimate**	**Standard error**	***t***	***df***	***p***	***d***
Explicability						
Test I	−0.055	0.207	−0.265	139	0.792	0.086
Test II	−0.124	0.207	−0.597	139	0.551	0.192
Test III	−0.490	0.207	−2.363	139	0.020	0.763
Generality						
Test I	−0.657	0.808	−0.813	140	0.418	0.261
Test II	0.395	0.808	0.489	140	0.625	0.157
Test III	−0.909	0.808	−1.125	140	0.262	0.362
Abstractness						
Test I	−0.626	0.493	−1.270	148	0.206	0.392
Test II	0.444	0.493	0.901	148	0.369	0.278
Test III	−0.436	0.493	−0.885	148	0.377	0.273
Systematicity						
Test I	−0.040	0.159	−0.251	149	0.802	0.078
Test II	0.223	0.159	1.405	149	0.162	0.434
Test III	0.664	0.159	4.181	149	< 0.001	1.290

[a]A positive difference indicates the test score was higher in the C-BLI-CL condition.

with a large effect size ($d = 1.290$). To examine within-group development, a similar model[12] excluding the pretest baseline covariate showed significant gains for both groups from the pretest to the posttest and from the pretest to the delayed posttest across all four languaging features (all $p < 0.001$). These results indicate that C-BLI led to sustained improvement in learners' ability to articulate modal semantics over time.

To illustrate these group differences qualitatively, delayed posttest languaging data from two representative learners are presented below (their full pre-, post-, and delayed posttest languaging data are available in Supplementary Materials A). Participant B12 (C-BLI-nonCL) produced clear, functional-semantic definitions of each modal, which reflected the degree-of-necessity hierarchy specified in the non-CL SCOBAs. The definitions he provided covered the key concepts of necessity and possibility across deontic, epistemic, and dynamic modality, demonstrating high explicability (essential features of the grammatical concept are present). However, the explanations were discrete rather than integrative, with no overarching conceptual connection among modals, reflecting lower systematicity.

C-BLI-nonCL participant B12's delayed posttest languaging:

1. *Must* is used when there is a strong requirement or when you have sufficient evidence to determine that a certain situation is necessary.
2. *Should* is weaker than *must* in terms of the degree of requirement (obligation), or when there is a slight necessity of a certain situation occurring.
3. *May* is a way of expressing permission, such as in the question "May I watch TV after finishing my homework," or it is also used when there are many possibilities for something to happen.
4. *Will* is used in situations where there is a voluntary (autonomous) decision, or when there is considerable confidence in predicting that something will happen.

Participant A9 (C-BLI-CL) demonstrated strong systematicity but relatively lower explicability. The learner consistently framed all four modals within a unified force-dynamic system (external vs.

TABLE 4 ANOVA output for the cloze test results.

Model term	df1	df2	F-ratio	p-value
Condition	1	58.96	0.445	0.507
Modal	3	841.75	13.200	<0.001
Test_Number	1	59	3.855	0.054
Pretest	1	899.89	1.755	0.186
Condition × Modal	3	841	1.885	0.130
Condition × Test_Number	1	59	1.275	0.263
Modal × Test_Number	3	841	4.262	0.005
Condition × Modal × Test_Number	3	841	1.061	0.365

internal force), showing coherent conceptual organization. However, the individual modal explanations were brief, omitting epistemic use for most modals and containing one inaccurate statement (*should: "external force causes internal force"*). These omissions resulted in lower explicability despite high conceptual integration.

C-BLI-CL A9 participant's delayed posttest languaging:

They (modals) all express force.
must: *There is a strong external force. It means someone outside forces someone to do something.*
should: *There is an obligation involved. It means external force causes internal force.*
may: *An external authority allows actions. There is no evidence to indicate whether to do something.*
will: *This is an internal force, originating from within oneself.*

Ownership

A linear mixed model[13] was fitted to the ownership measure to assess the difference between conditions and the change between the three tests. As there was no baseline measure of ownership, the same model was used for both sets of tests. Condition and test number were included as fixed effects, along with the interaction term. Participant was included as a random effect. The results revealed no group difference in the posttest (estimate = 0.14, [−0.18, 0.46], $t = 0.87$, $p = 0.385$) and the delayed posttest (estimate = 0.27, [−0.05, 0.58], $t = 1.67$, $p = 0.097$). Both groups maintained the same level of ownership with respect to languaging from the posttest to the delayed posttest: C-BLI-CL (estimate = 0.03, [−0.25, 0.32], $t = 0.24$, $p = 0.812$) and C-BLI-nonCL (estimate = -0.09, [−0.37, 0.18], $t = -0.68$, $p = 0.498$).

Research question 2: Quality of learner performance (C-BLI-CL vs. C-BLI-nonCL)

To address the second research question, we analyzed the cloze test performance of the two instruction groups (C-BLI-CL and C-BLI-nonCL). Figure 5 presents the descriptive means and confidence intervals for cloze test accuracy across variables. A linear mixed-effects model[14] was used to assess the effect of instructional condition on cloze test scores. Condition, modal verb, and test number were included as fixed effects, along with all interaction terms[15]. The pretest score was included as a covariate to control for baseline differences. Participant and test number (nested within participant) were included as random effects. The full ANOVA table for this model is presented below (Table 4).

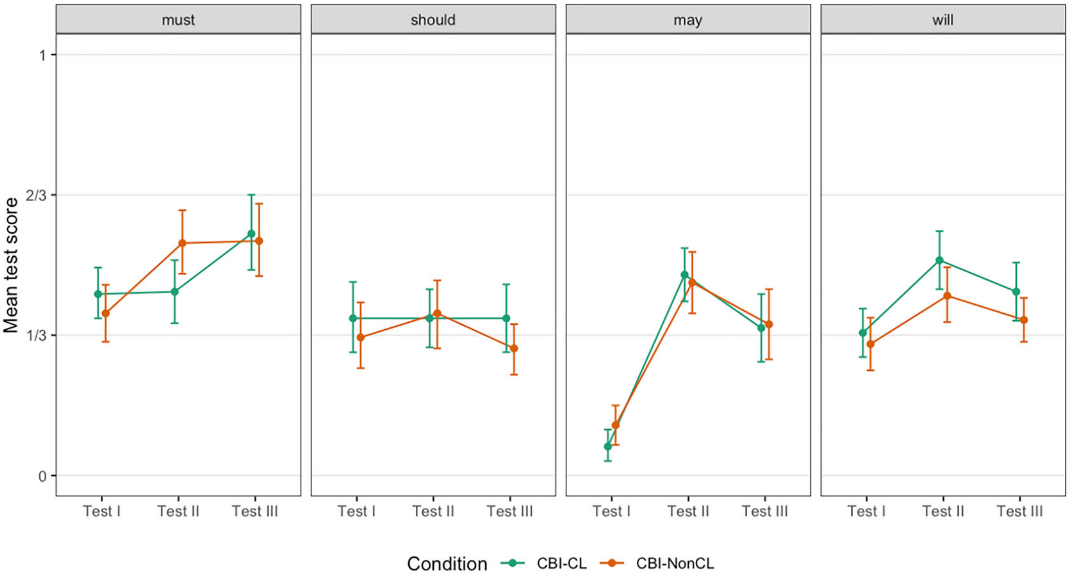

FIGURE 5 Means and confidence intervals for the cloze test scores. *Note*: Error bars show individual bootstrap 95% confidence intervals for the mean. [Color figure can be viewed at wileyonlinelibrary.com]

The pretest scores did not significantly predict the cloze test outcomes at later time points. This may suggest that the instructional interventions had a stronger influence on learning outcomes than initial ability, or that pretest scores were relatively homogeneous across participants. To assess differences between successive tests more directly, a similar model[16] was used without the pretest score as a covariate. The pattern of the results remained the same as the original model. Accordingly, the following post hoc analyses are based on the model without the pretest covariate to facilitate interpretation across all test sessions.

The ANOVA results showed significant fixed effects of modals and test number on cloze test accuracy, but no significant effect of instructional condition. Regarding modals, participants demonstrated the following accuracy hierarchy: *must* > *may* ≈ *will* > *should* (averaged over levels of condition and test number). *Must* emerged as the best-acquired modal, significantly outperforming all others (all small effect sizes: d ranged from 0.333 to 0.630). *Should* was the weakest, scoring significantly below the other modals (all small effect sizes: d ranged from 0.257 to 0.630). The modals *may* and *will* showed no significant difference (estimate = −0.011, $SE = 0.026$, $t = -0.413$, $p = 0.976$).

A significant interaction was found between the modal verb and test number (Figure 6), averaged across instructional conditions. Accuracy improvements across test sessions were observed for *must*, *may*, and *will*, but not for *should*. Specifically, *must* showed significant gains from pretest to posttest (Test II-I: estimate = 0.086, $t = 2.587$, $p = 0.027$, $d = 0.334$) and from pretest to delayed posttest (Test III-I: estimate = 0.158, $t = 4.734$, $p < 0.001$, $d = 0.613$). *May* demonstrated the largest improvement (Test II-I: estimate = 0.373, $t = 11.198$, $p < 0.001$, $d = 1.449$; Test III-I: estimate = 0.261, $t = 7.816$, $p < 0.001$, $d = 1.011$). *Will* also improved significantly from pretest to posttest (Test II-I: estimate = 0.144, $t = 4.309$, $p = .0001$, $d = 0.557$), with a marginal gain from pretest to delayed posttest (Test III-I: estimate = 0.078, $t = 2.325$, $p = 0.052$, $d = 0.301$). In contrast, *should* showed no significant improvement across tests (Test II-I: estimate = 0.029, $t = 0.859$, $p = 0.666$, $d = 0.111$; Test III-I: estimate = −0.013, $t = -0.391$, $p = 0.919$, $d = 0.051$).

Instruction condition did not yield an overall main effect. But a closer look at the instruction effects across time through a pairwise comparison (Table 5) showed that both types of C-BLI instruction yielded significant gains in accuracy from the pretest and the posttest (C-BLI-CL: $d = 0.568$;

TABLE 5 Pairwise comparisons of cloze results between successive tests.

	Test II—Test I					Test III—Test I					Test III—Test II				
Condition	**Estimate**	**SE**	**95% CI**	***t***	***p***	**Estimate**	**SE**	**95% CI**	***t***	***p***	**Estimate**	**SE**	**95% CI**	***t***	***p***
C-BLI-CL	0.15	0.02	0.10, 0.20	5.91	<0.001	0.13	0.02	0.08, 0.18	5.27	<0.001	−0.02	0.02	−0.06, 0.03	−0.64	0.525
C-BLI-nonCL	0.17	0.02	0.12, 0.22	7.17	<0.001	0.11	0.02	0.06, 0.16	4.69	<0.001	−0.06	0.02	−0.11, −0.01	−2.49	0.014

Note: A positive difference indicates the cloze test score was higher on the later test.
Results are averaged over function and modal verb.

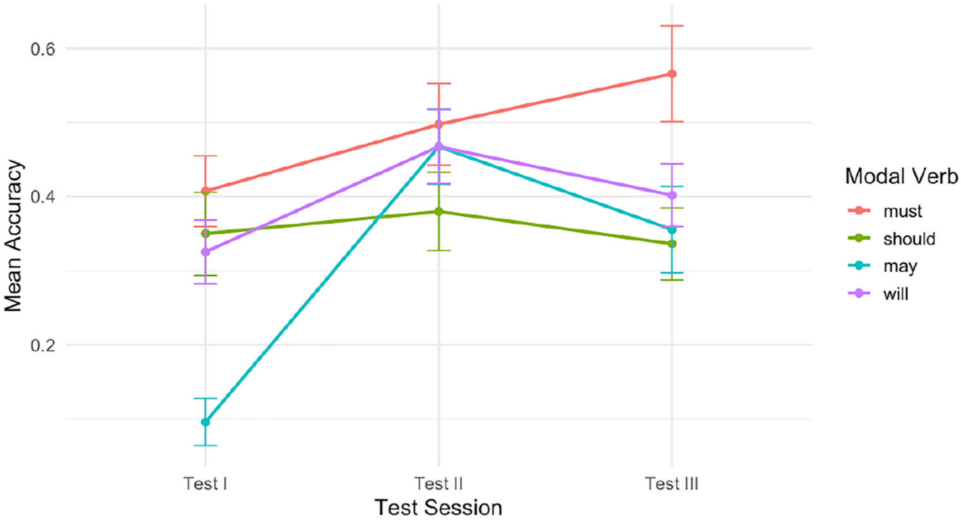

FIGURE 6 Interaction between modal verb and test session on cloze test accuracy. *Note*: The figure displays mean accuracy scores for each modal verb (*must, should, may, will*) across the three test sessions (Test I pretest, Test II posttest, Test III delayed posttest), with 95% confidence intervals. [Color figure can be viewed at wileyonlinelibrary.com]

C-BLI-nonCL: $d = 0.657$) and from the pretest and the delayed posttest (C-BLI-CL: $d = 0.507$; C-BLI-nonCL: $d = 0.430$). The only group difference we observed was that the C-BLI-CL group maintained its performance between the posttest and delayed posttest, whereas the C-BLI-nonCL group exhibited a significant decline from the posttest to the delayed posttest ($d = 0.228$).

Research question 3: C-BLI groups vs. control group

To address the third research question, we analyzed the test performance of the control group. As expected, the control group did not exhibit the same degree of improvement in languaging or cloze test accuracy as the C-BLI groups. In terms of meaningful languaging units, the control group started at a comparable level to the C-BLI groups on the pretest (estimate = 0.319, $t = 0.269$, $p = 0.788$). They showed a near-significant increase in the number of meaningful units at the posttest (estimate = 0.667, $t = 2.449$, $p = 0.053$, $d = 0.894$). But their verbalized meaningful units at the delayed posttest dropped to below the pretest level (estimate = −2.333, $t = -8.573$, $p < .001$, $d = 3.130$). With respect to languaging meaningfulness ratio, the control group also scored similarly as the C-BLI groups at the pretest (estimate = 0.027, $t = 1.428$, $p = 0.158$), showed no improvement at the posttest (estimate = 0.008, $t = 0.640$, $p = 0.799$), and declined below baseline at the delayed posttest (estimate = −0.036, $t = -2.891$, $p = 0.020$, $d = 1.056$).

The control group also showed no significant gains in languaging quality features. At the pretest, their scores on explicability (estimate = 0.109, $t = 0.527$, $p = 0.6$), generality (estimate = −0.200, $t = -0.692$, $p = 0.491$), abstractness (estimate = 0.050, $t = 0.179$, $p = 0.858$), and systematicity (estimate = −0.137, $t = -1.047$, $p = 0.298$) were comparable to those of the C-BLI groups. However, while the other features remained stable across test sessions, the control group's explicability declined at the delayed posttest (estimate = −0.733, $t = -3.487$, $p = 0.005$, $d = 1.273$). Finally, cloze test results mirrored this pattern. The control group began with similar accuracy to the C-BLI groups at the pretest (estimate = −0.023, $t = -1.720$, $p = 0.216$) and did not demonstrate any significant improvement over time.

DISCUSSION

Summary of results

The current findings show that both C-BLI-treated groups demonstrated significant instructional gains, with improvements in both the quality of languaging features and the accuracy of modal verb usage in the cloze. In contrast, traditionally instructed learners, who were provided with modal definitions and examples, failed to achieve effective learning outcomes. When comparing the two C-BLI conditions, the C-BLI-CL group significantly outperformed the C-BLI-nonCL group in increasing the proportion of meaningful languaging units from pretest to posttest and in enhancing the systematicity of their languaging at the delayed posttest. Conversely, the C-BLI-nonCL group showed higher explicability in their languaging at the delayed posttest. The two groups did not differ significantly on the remaining languaging measures. Regarding modal usage in the cloze test, the C-BLI-CL group maintained their accuracy gains over a three-week interval, whereas the C-BLI-nonCL group experienced a significant drop in accuracy at the delayed posttest.

The role of cognitive linguistics in enhancing concept-based language instruction

The superior performance of the C-BLI-CL group in terms of promoting meaningful and systematic languaging in the delayed posttest suggests that CL-based conceptual explanations provided a more effective framework for learners to internalize the meaning of the four target modals. This finding aligns with systemic theoretical instruction (STI) (Gal'perin, 1969, 1992), the theoretical foundation of C-BLI, which posits that learner development must be guided by explicit conceptual understanding before automatic performance can emerge. It also resonates with a growing body of C-BLI research in which CL-based conceptualizations have supported durable L2 development across a range of grammatical domains and languages, including aspect (Gánem-Gutiérrez, 2016; Kissling & Muthusamy, 2022; Poehner & Infante, 2017), conditional constructions (Jacobsen, 2015), modal verbs (Zhao, 2026), verbal particles (Lee, 2016), and prepositions (Arnett & Deifel, 2015; Buescher & Strauss, 2018), as well as studies in Japanese and Chinese (Masuda & Labarca, 2015; Masuda et al., 2025; Xi, 2022). Taken together, these findings indicate that CL-based conceptual explanations provide a robust foundation for mediating the development of form-meaning connections in L2 learning.

In contrast, the control group failed to show meaningful improvement in either languaging quality or accuracy of modal usage. In some cases, their performance at the delayed posttest even declined below baseline levels. This pattern is not unexpected. The modal definitions presented to this group were already familiar from prior classroom instruction and offered little new conceptual input. Moreover, the repeated testing without corresponding cognitive engagement may have diminished participants' motivation over time, possibly contributing to the observed performance decline.

Cognitive linguistics contributes to this pedagogical vision by offering a meaning-based account of grammar grounded in embodiment, force dynamics, construal, and metaphor. Through the lens of force dynamics (Sweetser, 1990; Tyler et al., 2010), CL provides learners with an experientially motivated conceptual system that renders modal meanings coherent and structured. The schematic representations (SCOBAs) used in CL-inspired instruction likely enhanced deep and durable conceptualization by directing learners' attention to fundamental semantic distinctions (e.g., external vs. internal force), thereby supporting retention beyond surface-level memorisation.

This theoretical and instructional alignment between CL and SCT-based C-BLI is not coincidental. Both frameworks are inherently meaning-oriented: CL offers rich conceptual content for grammar instruction, while SCT provides a developmental methodology for internalizing these concepts. As

Langacker (2008) and MacWhinney (2008) argued, meaning in language is not merely referential but shaped by construal: the speaker's perspective on a scene or event. C-BLI operationalizes this insight by making construals explicit through guided instruction, languaging, and practice. In this way, language functions as a mediational tool for conceptual development, consistent with the SCT principle of mediation.

Although previous studies have explored the pedagogical value of integrating CL and SCT (see review in the Pedagogical Framework section), most have relied on qualitative analyses of learner cases. While these studies documented gains in conceptual understanding and, in some cases, usage, they lacked larger-scale quantitative evidence. The current study helps to fill this gap by providing controlled, replicable, group-level evidence on developmental trends over time. Although the instructional conditions were systematically controlled to distinguish the pedagogical approaches, the intervention was implemented in intact College English classes during regular class hours, ensuring that findings reflect authentic classroom learning rather than a laboratory setting and as such reflect Vygotsky's (1997b, p. 94) contention that the hand of the researcher must be effectively replaced by the hand of practitioners in real life because for him practice ultimately "serves as the supreme judge of theory" (Vygotsky, 1997a, p. 306). The results show that CL-enhanced C-BLI not only improves conceptual understanding but also leads to measurable, durable gains in language performance. At the same time, as is typical in quantitative research, not all learners within a group followed identical trajectories. Mixed-effects modeling accommodates such variability statistically, but future research will benefit from complementary microgenetic analyses that trace how conceptual understanding stabilizes or fails to stabilize at the individual level.

At the same time, the finding that the C-BLI-nonCL group outperformed the C-BLI-CL group in explicability at the delayed posttest warrants further consideration. One possible explanation lies in the structure of the C-BLI-nonCL SCOBA, which was based on Palmer's (1990, 2001) three-way classification of modality (deontic, epistemic, dynamic), each subdivided by possibility and necessity. This framework encouraged learners to articulate multiple discrete points, contributing to higher explicability scores. However, our analysis of the number of meaningful languaging units showed no significant group difference, suggesting that the higher scores may reflect a broader but more segmented form of explanation, rather than greater elaboration overall. In contrast, several learners in the C-BLI-CL group at the delayed posttest focused their explanations primarily on root meanings and the force-dynamics relationships among modals, without elaborating on epistemic interpretations. These responses were coded as showing awareness of root meaning only, which contributed to lower explicability scores despite the presence of conceptually rich content. The C-BLI-CL SCOBA promoted a more streamlined and integrated conceptual structure, which may explain the group's significantly higher systematicity and their ability to relate modal concepts within a unified explanatory framework. This contrast illustrates a pedagogical trade-off between explicability (breadth) and systematicity (coherence): while functional grammar explanations based on Palmer's taxonomy may elicit a wider array of explicitly stated points, they may not support the same level of conceptual integration as the CL-based approach.

Indeed, although the C-BLI-nonCL group initially matched or exceeded the CL group in certain languaging measures, their performance in modal usage accuracy declined by the delayed posttest. In contrast, the C-BLI-CL group maintained their gains, suggesting that conceptual grounding, rather than mere verbal articulation, may be more predictive of long-term performance. These findings confirm that CL and C-BLI are not only theoretically compatible but mutually reinforcing (Lantolf, 2011; Lantolf & Poehner, 2014; Negueruela & Lantolf, 2006; White & Masuda, 2023). CL provides the conceptual depth and explanatory richness needed to reframe grammatical features as systems of meaning, while C-BLI provides the pedagogical pathway to guide learners in internalizing these systems through mediated instruction. Together, they offer a promising framework for advancing meaning-oriented, developmentally grounded L2 pedagogy.

Development of modal verbs and functions

Participants' accuracy in using modal verbs revealed a clear hierarchy: must > may ≈ will > should. *Must* emerged as the most accurately used modal across all tests, while *should* consistently score lowest. Significant improvements in accuracy were observed across test sessions for *must*, *may*, and *will*, but not for *should*. The observed hierarchy in modal verb acquisition reflects both semantic complexity and conceptual accessibility. The superior performance on *must* may stem from its consistent force-dynamic profile (Sweetser, 1990; Tyler et al., 2010) and relatively unambiguous mapping to necessity (Palmer, 1990).

By contrast, *should* is the least well-acquired modal and did not show improvements after instruction. From a CL perspective, *should* is conceptually more complex than *must*, *may*, or *will*. While these latter modals each reflect a single force source (external force for *must* and *may*, internal force for *will*), *should* involves an interaction between external and internal forces (external compulsion combined with internal alignment) (Tyler et al., 2010). This dual-source force dynamic is more abstract and less readily imageable, potentially contributing to learners' confusion. From a functional grammar viewpoint, *should* is classified under dynamic modality as expressing "willing necessity" (Palmer, 1990, 2001), but this categorization is problematic. Its core meaning, often associated with weak obligation, does not neatly align with either volition (as in *will*) or necessity (as in *must*), which are more prototypical of dynamic and deontic modality, respectively. This complexity was also reflected in the SCOBAs used in instruction. Prior functional analyses (e.g., Hoye, 1997) have also noted that *should* frequently functions with mitigated force, which may reduce its salience and communicative clarity. Consequently, *should* occupies a less clear-cut position within both conceptual and functional modality systems, which may have hindered its acquisition. A further possibility is that the difficulty with *should* stems from its discourse-pragmatic behavior in contemporary English. Corpus-based research shows that *should* frequently functions as a stance marker expressing soft directives, expectations, or advisability, especially in formal and academic registers (Biber, 2006; Biber et al., 1999). These nuanced interpersonal uses are often underrepresented in pedagogical materials, creating a form-function mismatch that may further contribute to learners' persistent difficulty with *should*. Taken together, these factors help explain why *should* lagged behind the other modals in learners' performance accuracy and showed no significant improvement across tests.

One avenue for refinement concerns how *should* is taught within a CL-based pedagogy. In a force-dynamic account, *should* encodes a negotiable external force that seeks internal alignment, making it less imageable than the single-source patterns of *must*, *may*, and *will*. Two CL-consistent adjustments may better support internalization. First, an embodied mediating activity (Zhao, 2026) could heighten the sense of "negotiable force": learners experience partial resistance and partial compliance (e.g., light push/pull with a resistance band) and then map that sensation onto *should* using the existing SCOBA. Second, targeted contrast tasks with minimal pairs can sharpen category boundaries: *must-should* (non-negotiable vs. negotiable compulsion) and *may-should* (barrier removed vs. gentle pressure toward action). Such refinements aim to make the dual-source configuration of *should* more experientially salient and learnable.

While this study offers valuable insights into the role of C-BLI and CL in L2 modal learning, several limitations should be acknowledged. First, although we collected additional data during the training phase, including participants' oral languaging episodes and, for the C-BLI-CL group, take-home assignments involving their own diagrammatic representations of modal meanings, this study focused solely on the quantitative analysis of written languaging and cloze test performance at the test sessions. A qualitative analysis of these rich training-phase data could yield deeper insights into learners' conceptual development and their interaction with different types of SCOBAs, as shown in previous studies (e.g., Negueruela, 2003; Qin et al., 2023; Yáñez-Prieto, 2008, 2014). Second, we only targeted four modal verbs in this study, which is not a full representation of the modality system. Additionally, some complexities of modal usage (e.g., present vs. past tense modals, negation, and questions) were not explicitly taught. Future research should examine whether the conceptual knowledge gained from instructed modals transfers to uninstructed modals, which would provide stronger evidence for the effectiveness of CL-based C-BLI instruction.[17] Third, the ultimate test of pedagogical effectiveness lies in the results produced in authentic classrooms. Earlier, Lysinger (2010) demonstrated that a cognitive grammar-based approach to teaching German and Russian cases led to improved learning outcomes over traditional form-focused instruction, even in beginning language classes taught by novice graduate student instructors. The present study provided encouraging evidence that a CL-based conceptual explanation of a grammatical feature can be more effectively integrated into C-BLI pedagogy in a basic language program than a meaning-based explanation derived from an alternative theory. Future research should examine whether similar results can be achieved for other language features.

CONCLUSION

The present study contributes to the growing body of research on C-BLI by demonstrating the pedagogical potential of integrating CL into L2 grammar instruction. The findings highlight that explicit conceptual explanations, particularly those grounded in CL, can enhance learners' ability to internalize modal verb meanings and retain knowledge over time. Moving forward, replicating these findings in authentic classroom environments, extending instruction to a broader range of modals, and examining the potential for transfer to uninstructed forms will be crucial for refining C-BLI as a robust pedagogical framework. By continuing to explore the interface between CL, SCT, and L2 instruction, future research can further illuminate how conceptual instruction enhances long-term language development.

AUTHOR CONTRIBUTIONS

Helen Zhao: Conceptualisation (lead); methodology (lead); investigation (lead); project administration (lead); resources (lead); data curation (lead); formal analysis (lead); validation (lead); visualisation (lead); writing—original draft preparation (lead); writing—review and editing (lead). **James P. Lantolf:** Conceptualisation (supporting); writing - the SCT/C-BLI theoretical framework (lead); writing - original draft preparation (supporting); writing—review and editing (supporting).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ENDNOTES

[1] In fact, Negueruela and Lantolf (2006) originally referred to the approach as concept-based Instruction (CBI). However, to avoid confusion with content-based instruction, also frequently abbreviated as CBI, the approach was eventually referred to as Concept-based Language Instruction, abbreviated as C-BLI.

[2] van Compernolle (2014), for instance, published a book-length study on the teaching of pragmatic competence in L2 French using C-BLI; Urbanski (2023) used the approach to provide instruction in L2 reading; Kim and Lantolf (2018) used C-BLI to provide instruction on the concept of sarcasm in English; Stam et al. (2023) provided instruction on motion events within Slobin's thinking for speaking framework, and a recent book-length publication by Masuda et al. (2025) includes studies focused on pragmatic and grammatical features of Japanese.

[3] Voloshinov (1973, p. 68) makes an important distinction between signs and signals. Only the former can be understood, while the latter can only be recognized. Accordingly, "a signal is an internally fixed singular thing that does not in fact stand for anything else … but is simply a technical means for indicating this or that object or this or that action … it is "stable and always self-equivalent"" (p. 68). Signs, however, are "always changeable and adaptable" and therefore capable of manifesting "the speaker's point of view." Effective communication then, is about creating and comprehending a signs novelty, not producing and recognizing signality (p. 69). Thus, if instruction and learning of a new language focuses on signal production and recognition, which, in our view, is what happens when language is approached as rule-governed behavior rather than concept-based signs, to paraphrase Voloshinov (1973, p. 69), fails to fully become a language. An example of what we have in mind here is provided from the commentary of one of Yáñez-Prieto's (2008) students, who, at the conclusion of his C-BLI experience, stated the following: "It's kind of funny how you can have a grammar st…the gram…grammatical structure

actually tell a story. I'd not really noticed that or seen that before. I mean, the words are telling the story and the grammar is telling the story, which is kind of weird. Yeah, I'd never seen that before. Interesting."

[4] A special issue of *Language and Sociocultural Theory* addressed the process through which SCOBAs are generated. It also offers some recommendations that teachers should keep in mind when formulating and using SCOBAs in C-BLI pedagogy (see White et al., 2021).

[5] We use the term languaging rather than verbalization used in Gal'perin's STI, because the former term is intended to refer uniquely to the use of language in its psychological, self-regulatory function, while the latter term is ambiguous, given that it can also refer to social, other-directed communicative speech.

[6] For example, in *He will be in his office now*, *will* expresses a confident inference based on expectation or typical behavior. Under Sweetser's metaphor, the epistemic meaning should derive from a force schema along the lines of *some force compels the event to happen*. But in this epistemic *will*, there is no agent exerting force and no compulsion applied to the subject; the statement refers instead to a default inference/expectation about how the world normally unfolds. There is no obvious "force" that compels the conclusion. What the speaker expresses is closer to habitual expectation than to dynamic coercion. As scholars such as Langacker (1991) argue, such instances reveal that epistemic modality often concerns reasoning and evidential assessment rather than force-dynamic interaction.

[7] A similar pattern of findings has emerged cross-linguistically. For instance, in the context of German modal verbs, Nikolova (2015) demonstrated that a CL-based approach grounded in force dynamics led to greater gains in learners' receptive understanding compared to traditional instruction. Although not situated within a C-BLI framework, this study reinforces the broader potential of force-dynamic explanations for teaching modal meaning.

[8] Model formula: Number_of_meaningful_units ∼ Condition × Test_Number + Test_I_Number_of_meaningful_units + (1|No)

[9] Model formula: cbind(Number_of_meaningful_units, Number_of_incorrect_units) ∼ Condition × Test_Number + logit_Test_I_meaningfulness_ratio + (1|No/Test_Number)

[10] Model formula: cbind(Number_of_meaningful_units, Number_of_incorrect_units) ∼ Condition × Test_Number + (1|No/Test_Number)

[11] Model formula: Post_Test_languaging ∼ Condition × Test_Number + Pre_Test_languaging + (1|No)

[12] Model formula: languaging ∼ Condition × Test_Number + (1|No)

[13] Model formula: ownership ∼ Condition × Test_Number + (1|No)

[14] Model formula: Post_Test ∼ Condition × Modal × Test_Number + Pre_Test + (1|No/Test_Number)

[15] Modal function was not included as a fixed effect in the model because the two instructional conditions were based on different functional categorisations of modality. The CL-based instruction distinguished between root and epistemic functions, whereas the Palmer-based instruction followed a three-way classification into deontic, epistemic, and dynamic modalities. Including function as a shared predictor would therefore obscure these theoretical differences and conflate distinct pedagogical frameworks.

[16] Model formula: Post_Test ∼ Condition × Modal × Test_Number + (1|No/Test_Number)

[17] Lee's (2016) study showed that learners were able to transfer knowledge gained from instructed English particle verbs to uninstructed particle verbs.

REFERENCES

Arievitch, I. M. (2017). *Beyond the brain. An agentive activity perspective on mind, development, and learning*. Sense Publishers.

Arnett, C., & Deifel, K. (2015). A cognitive linguistic approach to two-way prepositions in L2 German. In K. Masuda, C. Arnett, & A. Labarca (Eds.), *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching* (pp. 183–202). De Gruyter Mouton.

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., Krivitsky, P. N., Tanaka, E., Jagan, M., Boylan, R. D., & Ly, A. (2022). lme4: Linear Mixed-Effects Models using 'Eigen' and S4. https://cran.r-project.org/web/packages/lme4/lme4.pdf

Benhamlaoui, M., & Gánem-Gutiérrez, G. A. (2022). A classroom application of concept-based language instruction (C-BLI). *Language and Sociocultural Theory*, *9*(1), 1–33.

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, *5*(2), 97–116. https://doi.org/10.1016/j.jeap.2006.05.001

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.

Buescher, K., & Strauss, S. (2018). Conceptual frameworks and L2 pedagogy: The case of French prepositions. In A. Tyler, L. Ortega, M. Uno, & H. I. Park (Eds.), *Usage-inspired L2 instruction: Researched pedagogy* (pp. 95–115). John Benjamins.

Carpay, J. A. M. (1974). Foreign language teaching and meaningful learning. A Soviet Russian point of view. *I.T.L. Review of Applied Linguistics*, *25*(6), 161–187.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide*. Cambridge University Press.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course* (2nd ed.). Newbury House.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (Rev. ed.)*. Academic Press.

Davies, M. (2008–). *The Corpus of Contemporary American English (COCA)*. https://www.english-corpora.org/coca/

Depraetere, I. (2015). Categorization principles of modal meaning categories: A critical assessment. *Anglophonia*. Presses Universitaires du Midi. http://anglophonia.revues.org/453; https://doi.org/10.4000/anglophonia.476

Di Pietro, R. (1987). *Strategic interaction*. Cambridge University Press.

Engeness, I. (2021). *P. Y. Gal'perin's development of human mental activity*. Springer.

Englund, D. L. (2015). A blended cognitive, linguistic, and Vygotskian approach for teaching and learning the prepositions *in*, *on*, and *of* in the advanced ESL classroom [Unpublished doctoral dissertation]. University of Kansas.

Council of Europe. (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Ferreira, M. M., & Lantolf, J. P. (2008). A concept-based approach to teaching writing through genre analysis. In J. P. Lantolf, & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 285–320). University of Toronto Press.

Frawley, W., & Lantolf, J. P. (1985). L2 discourse: A Vygotskian perspective. *Applied Linguistics*, *6*(1), 19–44. https://doi.org/10.1093/applin/6.1.19

Gal'perin, P. (1992). Stage-by-stage formation as a method of psychological investigation. *Journal of Russian and East European Psychology*, *30*(4), 60–80.

Gánem-Gutiérrez, G. A. (2016). Enhancing metalinguistic knowledge: Preterite and imperfect in L2 Spanish. *Language and Sociocultural Theory*, *3*(1), 27–54.

Gal'perin, P. Y. (1969). Stages in the development of mental acts. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology* (pp. 249–273). Basic Books.

García, P. N. (2017). A sociocultural approach to analyzing L2 development in the Spanish L2 classroom. *Vigo International Journal of Applied Linguistics*, *14*, 99–124.

Haenen, J. (1996). *Piotr Gal'perin. Psychologist in Vygotsky's footsteps*. Nova Science Publishers.

Hornby, A. S. (2005). *Oxford advanced learner's dictionary of current English* (7th ed.; S. Wehmeier, Ed.). Oxford University Press.

Hoye, L. F. (1997). *Adverbs and modality in English*. Longman.

Ilyenkov, E. (2012). Dialectics of the ideal. *Historical Materialism*, *20*(2), 149–193.

Jacobsen, N. D. (2015). A cognitive linguistic analysis of English conditionals in English for Academic Purposes (EAP) instruction: Implications from Sociocultural Theory. In K. Masuda, C. Arnett, & A. Labarca (Eds.), *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching* (pp. 103–125). De Gruyter Mouton.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination and reason*. Chicago University Press.

Kabanova, O. Y. (1985). The teaching of foreign languages. *Instructional Science*, *14*, 1–47.

Kim, J., & Lantolf, J. P. (2018). Developing conceptual understanding of sarcasm in L2 English through explicit instruction. *Language Teaching Research*, *22*(2), 208–229. https://doi.org/10.1177/1362168816675521

Kissling, E. M. (2023). Can concept-based language instruction change beginning learners' aspectual development?: Preliminary experimental evidence that novice learners taught boundedness are less influenced by lexical aspect. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*. https://doi.org/10.22049/jalda.2023.27948.1447

Kissling, E. M., & Muthusamy, T. (2022). Exploring boundedness for concept-based instruction of Aspect: Evidence from learning the Spanish preterite and imperfect. *Modern Language Journal*, *106*(2), 371–392.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Langacker, R. W. (1991). *Foundations of cognitive grammar*. (Vol. 2: Descriptive application). Stanford University Press.

Langacker, R. W. (2008). *Cognitive grammar*. Oxford University Press.

Lantolf, J. P. (2011). Integrating sociocultural theory and cognitive linguistics in the second language classroom. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 303–318). Routledge.

Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge.

Lantolf, J. P., & Tsai, M. H. (2018). L2 developmental education and systemic theoretical instruction: The case of English verb+noun collocations. In A. Tyler, et al. (Eds.), *Usage-inspired L2 instruction*. (pp. 29–53). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.49.02lan

Lee, H. (2016). Concept-based instruction: Imagistic and metaphorical understanding of phrasal verbs. *English Teaching*, *71*(4), 167–191. https://doi.org/10.15858/engtea.71.4.201612.167

Lenth, R. V., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Piaskowski, J., Riebl, H., & Singmann, H. (2023). emmeans: Estimated marginal means, aka least-squares means. Comprehensive R Archive Network. https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language History Questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/S1366728913000606

Luria, A. R., & Vygotsky, L. S. (1992). *Ape, primitive man, and child. Essays in the history of behavior*. CRC Press.

Lysinger, D. (2010). *Teaching cases to L2 German and Russian learners: A cognitive approach* [Unpublished doctoral dissertation]. University of California.

MacWhinney, B. (2008). How mental models encode embodied linguistic perspectives. In R. Klatzky, B. MacWhinney, & M. Behrmann (Eds.), *Embodiment, ego-space, and action* (pp. 369–410). Lawrence Erlbaum.

Masuda, K., Arnett, C., & Labarca, A. (2015). *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching*. De Gruyter.

Masuda, K., & Labarca, A. (2015). Schematic diagram use and languaging quality in learning Japanese polysemous particles ni and de. In K. Masuda, C. Arnett, & A. Labarca (Eds.), *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching* (pp. 203–232). De Gruyter Mouton.

Masuda, K., Ohta, A. S., & Tsujihara, R. (2025). *Concept-based language instruction: Usage-based linguistics and sociocultural theory in teaching Japanese*. Routledge.

Mulder, W. D. (2021). Force dynamics. In X. Wen, & J. R. Taylor (Eds.), *The Routledge Handbook of Cognitive Linguistics* (pp. 228–241). Routledge.

Negueruela, E. (2003). A sociocultural approach to teaching and researching second language: Systemic-theoretical instruction and second language development [Unpublished doctoral dissertation]. Pennsylvania State University.

Negueruela, E., & Lantolf, J. P. (2006). Concept-based instruction and the acquisition of L2 Spanish. In R. Salaberry, & B. A. Lafford (Eds.), *The art of teaching Spanish: Second language acquisition from research to praxis* (pp. 79–102). Georgetown University Press.

Nikolova, M. M. (2015). *The effect of a Cognitive Linguistics approach to teaching the semantic meanings of German modal verbs* [Unpublished doctoral dissertation]. University of Illinois at Chicago.

Palmer, F. (1990). *Modality and the English modals* (2nd ed.). Routledge.

Palmer, F. (2001). *Mood and modality*. Cambridge University Press.

Plonsky, L., & Oswald, F. L. (2014). How big Is “big”? Interpreting effect sizes in L2 research: Effect sizes in L2 research. *Language Learning*, *64*, 878–912. https://doi.org/10.1111/lang.12079

Poehner, M. E., & Infante, P. (2017). Mediated development: A Vygotskian approach to transforming second language learner abilities. *TESOL Quarterly*, *51*(2), 332–357.

Qin, L., Ouyang, X., & Ren, W. (2023). Decoding EFL learners’ internalization of English modal verbs in concept-based language instruction: Integrating SCOBAs with languaging. *The Modern Language Journal*, *107*, 24–45. https://doi.org/10.1111/modl.12821

R Core Team. (2024). R: A language and environment for statistical computing (Version 4.4.0) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Smotrova, T., & Lantolf, J. P. (2013). The function of gesture in lexically focused L2 instructional conversation. *The Modern Language Journal*, *97*(2), 397–416. https://doi.org/10.1111/j.1540-4781.2013.12008.x

Stam, G., Urbanski, K., Lantolf, J. P., & Smotrova, T. (2023). How concept-based instruction works in teaching thinking for speaking in an L2. *IRAL*, *61*(1), 111–153. https://doi-org.ezaccess.libraries.psu.edu/10.1515/iral-2022-0073

Suethanapornkul, S. (2014). Embodied experiences in second language learning of English modal verbs. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. Marcos Miguel, A. M. Tseng, A. Tuninetti, & D. Walter (Eds.), *Selected proceedings of the 2012 Second Language Research Forum* (pp. 181–192). Cascadilla Proceedings Project.

Swain, M. (2006). Languaging, agency and collaboration in advanced language proficiency. In H. Byrnes (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). Georgetown University Press.

Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49–100. https://doi.org/10.1207/s15516709cog1201_2

Talyzina, N. (1981). *The psychology of learning*. Progress Press.

Torres-Martínez, S. (2019). Taming English modals. *English Today*, *138*(35), 50–57. https://doi.org/10.1017/S0266078418000081

Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Routledge.

Tyler, A., Mueller, C., & Ho, V. (2010). Applying cognitive linguistics to instructed L2 learning: The English modals. *AILA Review*, *23*, 30–49. https://doi.org/10.1075/aila.23.03tyl

Urbanski, K. B. (2023). *Second language literacy pedagogy. A sociocultural theory perspective*. Multilingual Matters.

van Compernolle, R. M. (2014). *Sociocultural theory and L2 instructional pragmatics*. Multilingual Matters.

van Parreren, J. (1975). Grammatical knowledge and grammatical skill. In T. Heron-Muller, J. P. Menting, & A. J. van Essen, (Eds.), *The context of foreign language learning* (pp. 117–131). Van Gorcum.

Vocate, D. R. (1994). Self-talk and inner speech. Understanding the uniquely human aspects of intrapersonal communication. In D. R. Vocate (Ed.), *Intraperesonal communication. Different voices, different minds* (pp. 3–32). Lawrence Erlbaum.

Voloshinov, V. N. (1973). *Marxism and the philosophy of language*. Seminar Press.

Vygotsky, L. S. (1986). Concrete human psychology. *Psikhologiya* 1, 51–64.

Vygotsky, L. S. (1987). *The collected works of L. S. Vygotsky*: *Volume 1. Problems of general psychology, including the volume Thinking and speech*. Plenum.

Vygotsky, L. S. (1997a). *The collected works of L. S. Vygotsky: Volume 3. Problems of the theory and history of psychology*. Plenum.

Vygotsky, L. S. (1997b). *The collected works of L. S. Vygotsky: Volume 4. The history of the development of higher mental functions*. Plenum.

Vygotsky, L. S. (1998). *The collected works of L. S. Vygotsky: Volume 5. Child psychology*. Plenum.

Wang, M., & Zhao, H. (2024b). Perceptual representations in L1 and L2 spatial and abstract language processing: Applying an innovative sentence-diagram verification paradigm. *Frontiers in Human Neuroscience*, *18*(1425576), 1–18. https://doi.org/10.3389/fnhum.2024.1425576

Wang, M. H., & Zhao, H. (2024a). Are schematic diagrams valid visual representations of concepts? Evidence from mental imagery in online processing of English prepositions. *Language and Cognition*, *16*(4), 1564–1587. https://doi.org/10.1017/langcog.2024.18

Wells, G. (1999). *Dialogic inquiry. Toward a sociocultural practice and theory of education*. Cambridge University Press.

White, B., Gánem-Gutiérrez, G. A., & Schulze, M. (2021). SCOBAs. *Language and Sociocultural Theory*, *8*(1).

White, B., & Masuda, K. (2023). Integration and compatibility of sociocultural theory and cognitive linguistics for second language lexicogrammar instruction. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances* , https://doi.org/10.22049/jalda.2024.28052.1476

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. https://ggplot2.tidyverse.org

Xi, J. (2022). Cultural-historical psychology and cognitive linguistics dialectically unified in second language pedagogy [Unpublished doctoral dissertation]. Xi'an Jiaotong University.

Yáñez-Prieto, C. (2008). On literature and the secret art of invisible words: Teaching literature through language [Unpublished doctoral dissertation]. The Pennsylvania State University.

Yáñez-Prieto, C. (2014). *Sense* and subjectivity: Teaching literature from a sociocultural perspective. *Language and Sociocultural Theory*, *1*(2), 179–203.

Yang, R. Y., & Sang, Y. R. (2023). Concept-based language instruction and the teaching of citation in English academic writing. *Modern Language Journal*, *107*(S1), 46–47.

Zhang, X., & Lantolf, J. P. (2015). Natural or artificial: Is the route of second language development teachable? *Language Learning*, *65*(1), 152–180. https://doi.org/10.1111/lang.12094

Zhao, H. (2026). Concept-based language instruction of English modal verbs: A Cognitive Linguistics-inspired approach. In B. White, & N. Dolvoga (Eds.), *Applying cognitive linguistics in the English language classroom*. Bloomsbury Press.

Zheng, Y., & Cheng, L. (2008). College English Test (CET) in China. *Language Testing*, *25*(3), 408–417. http://doi.org/10.1177/026553220809243

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Zhao, H., & Lantolf, J. P. (2026). Developing L2 semantic knowledge of English modality through concept-based language instruction: Do cognitive linguistics materials have an advantage?. *Modern Language Journal*, 1–28. https://doi.org/10.1111/modl.70033