

UNCERTAINTY MARKERS IN SPOKEN LEARNER DISCOURSE

Petra Huschová

Abstract

The paper discusses linguistic forms conveying epistemic stance in L2 spoken production, focusing on expressing uncertainty. The source of stance markers is a subset of the Corpus of Czech Students' Spoken English, particularly informal student-student discussions on a given topic. The main aim of the corpus analysis is to identify a variety of grammatical devices indicating the speaker's uncertainty about the truth-value of a proposition and to examine their distributional patterns and positional preferences. Additionally, the paper explores roles of the construction *I think*, the most frequent stance marker in the discussions analysed. The findings indicate that Czech learners of English tend to employ a restricted set of items recurrently and use *I think* not only to express uncertainty but also to mitigate potential disagreement and signal turn-taking.

Keywords

epistemic stance, uncertainty markers, L2 spoken discourse, Czech learners of English

1 Introduction

Reflecting the speaker's degree of commitment to the veracity of information and helping to manage discourse, epistemic stance markers play an important role in natural spoken interaction. Therefore, advanced L2 speakers of English should acquire the pragmatic skill of stance-taking to be able to evaluate the content of their utterances as well as "to take their parts in co-constructing the interaction in a manner appropriate to the social context, interlocutor roles, and purpose of the exchange" (Gablasova et al., 2017, p. 613) (see also Xu & Cao, 2023). The Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2020) lists discourse management and propositional precision, which includes qualifying opinions and expressing degrees of commitment, among abilities of L2 speakers' pragmatic competence (see Castello, 2023). Understanding how Czech learners of English employ epistemic stance markers in face-to-face interaction can thus shed light on some pragmatic aspects of their spoken production.

The current paper investigates how advanced Czech learners of English exploit English resources of epistemic stance to convey their attitude to a proposition and an addressee. The main aim is to identify and discuss epistemic stance markers which signal the speaker's uncertainty, focusing on overt grammatical devices,

that is, epistemic verbs, adverbs, adjectives, and nouns. The objectives are to examine a variety of linguistic forms expressing uncertainty, to explore their contexts of occurrence and distributional patterns, and to discuss functions of the marker *I think* in the interaction process. In order to investigate the usage of uncertainty markers by Czech learners of English, a subset of the Corpus of Czech Students' Spoken English (CCSSE) (Ježková, 2015) is analysed, particularly 114 student-student face-to-face interactions in which the interlocutors discuss a given topic and are expected to co-construct the interaction, express their stance and justify their viewpoint.

2 Epistemic stance

Stance-taking can be defined as “the positions we take towards our propositions and audience” (Hyland & Zou, 2021, p. 1), implying that it is a part of social interaction (Coates 1987, p. 120; Gablasova et al., 2024, p. 187). Du Bois (2007) emphasizes: “No stance stands alone. It is the stance utterance with *its dialogic context* that is the relevant unit for stance interpretation” (p. 158). Similarly, Kärkkäinen (2003) highlights that stance-taking is “an interactional practice engaged in by coparticipants in conversation, rather than an isolated mental position of an individual speaker” (p. 16).

Generally, stance refers to how speakers express their feelings, degrees of certainty and assessments of their own as well as others' propositions (Biber, 2006; Wu & Paltridge, 2021). This broad definition of stance thus includes a variety of concepts discussed in the scholarly literature, for example, ‘attitude’ (Halliday, 2004), ‘evaluation’ (Hunston & Thompson, 2000), or ‘affect’ and ‘evidentiality’ (Hyland, 1999). Hyland and Zou (2021, p. 3) explain that affect indicates the speaker’s feelings and affective attitudes, whereas evidentiality concerns the reliability of the knowledge contained in propositions. Holmes (1990, p. 201) labels the two main functions of stance markers ‘affective’ (related to the speaker’s attitude to the addressee) and ‘epistemic’ (related to the speaker’s degree of certainty).

Epistemic stance, corresponding to Hyland’s (1999) conception of evidentiality, is defined as the speaker’s position to the status of information in a proposition and assessment of its reliability, which implies that it concerns the speaker’s degree of certainty (Biber et al., 2021, p. 964; Kärkkäinen, 2003, p. 1). As Hyland and Zou (2021) state, epistemic stance markers “convey the speaker’s judgements of the reliability that might be invested in a statement, either casting doubt or asserting certainty” (p. 5). Based on the degree of certainty, they distinguish between hedges, that is, devices that withhold the speaker’s complete commitment to the truth-value of a proposition, and boosters, that is, devices that allow to present one’s claims with assurance.

Uncertainty markers, the focus of the current paper, can be viewed as a basic and dominant category of hedges (Vold, 2006, p. 64) because hedges are characterized as means that qualify utterances to mitigate their strength in terms of the author's lack of knowledge to the truth of the proposition (Crompton, 1997, p. 273; Hyland, 1995, p. 34). Additionally, hedges explicitly offer a personal assessment of the speaker's claims to addressees who can provide their alternative views (Hyland & Zou, 2021, p. 5). Similarly, Brezina (2009) emphasizes "the process of negotiating knowledge" in which uncertainty markers indicate not only the degree of likelihood of the truth value in a proposition but also justify the speaker's beliefs in a particular social setting (p. 43). Accordingly, in this paper, uncertainty markers are understood as multifunctional interactional devices whose interpretation is reliant on the immediate discourse context (Coates, 1987; Gablasova et al., 2024; Holmes, 1990; Kärkkäinen, 2003; White, 2008).

2.1 Categories of epistemic stance markers

Epistemic stance can be expressed and encoded through an array of devices (paralinguistic, grammatical, lexical); however, this paper focuses only on linguistic forms that overtly convey the speaker's position to a proposition. Therefore, based on the theoretical framework of epistemic stance proposed by Biber et al. (2021) and Biber (2006), the following grammatical devices have been included in the investigation: adverbs, modal verbs, and lexical verbs/adjectives/nouns followed by a complement clause (Biber, 2006, pp. 101–102; Biber et al., 2021, pp. 964–966). Biber et al. (2021, p. 962) argue that epistemic adverbs and complement clause constructions are the clearest cases of grammatical marking of stance because they comprise the component expressing a stance that frames the component presenting a proposition. Other studies (e.g., Brezina, 2009; Gablasova et al., 2024; Kärkkäinen, 2003) also report the above-mentioned word classes as the major categories of epistemic stance markers commonly appearing in spoken discourse.

Since the scope of this study is narrowed to uncertainty markers, Table 1 includes only the forms that can convey the speaker's uncertainty. Based on synthesizing the previously reported uncertainty markers (Biber et al., 2021; Biber, 2006; Brezina, 2009; Gablasova et al., 2024; Hoye, 2013; Xu & Cao, 2023), Table 1 summarizes all the expressions that were searched for in the corpus discussions. Drawing upon Xu and Cao (2023, p. 5), the uncertain position is perceived as a continuum ranging 'from not knowing whether something is true to believing that something is true'. Therefore, the category of adverbs includes also disjuncts that in terms of likelihood allow for some doubt (Hoye, 2013, p. 194).

Category	Linguistic form
Modal verb	<i>could, can, may, might, would, should</i>
Lexical verb	<i>think, don't know, suppose, guess, believe, doubt, assume, seem, appear</i>
Adverb	<i>maybe, perhaps, possibly, probably, apparently, arguably, conceivably, likely, presumably</i>
Adjective	<i>(un)likely, not sure/certain, possible, probable, doubtful</i>
Noun	<i>possibility, probability, likelihood, assumption</i>

Table 1: Grammatical devices conveying the speaker's uncertainty

2.2 Multifunctionality of epistemic stance markers: Communicative functions of *I think*

As mentioned above, epistemic stance markers are multifunctional items that genuinely convey the speaker's position to both a proposition and an addressee. Du Bois (2007) explains their multifunctional nature through the notion of the stance triangle as "I evaluate something, and thereby position myself, and thereby align with you" (p. 163). In other words, the speaker simultaneously evaluates a proposition, positions himself/herself to its truth value, and expresses agreement with other interlocutors. White (2008) argues that speakers employ epistemic stance markers primarily to position themselves with respect to other speakers (to make space for alternative perspectives), whereas their standpoint with respect to the truth value of a proposition is of secondary importance.

Regarding specific social functions of epistemic stance markers in spoken interaction, Kärkkäinen (2003, p. 26) lists, besides displaying the degree of certainty, a politeness/face-saving function, achieving conversational actions (such as avoiding agreement or disagreement), constructing discourse statuses of participants, and regulating aspects of interaction (like topic transition). As this paper is concerned with functions of the marker *I think* in spoken production by Czech learners of English, communicative functions of this epistemic phrase in spoken interaction are briefly introduced here.

Firstly, *I think* primarily functions as a device of epistemic stance conveying the speaker's subjective perspective. As Aijmer (1997), Baumgarten and House (2010), Caprario (2023), Coates (1987), Holmes (1990) or Kärkkäinen (2003) observe, *I think* can be used tentatively to indicate the speaker's uncertainty or deliberately to express certainty and deliberative opinion.

Secondly, *I think* is commonly employed as a device of discourse management and organization. Kärkkäinen (2003) argues that this epistemic phrase often "simply marks a boundary of some kind in the talk" (p. 121), meaning that it functions just as a routine organizational device. Baumgarten and House (2010)

agree that L1 speakers often use *I think* as “the pragmaticalized verbal routine” to organize their discourse, but they point out that L2 speakers employ it mainly to evaluate their propositions in terms of degrees of certainty (p. 1197). This functional diversity of *I think* is also noticed in spoken trial discourse investigated by Szczyrbak (2020), who concludes that this construction functions as a hedge when it is found with other uncertainty markers or as a discourse marker when co-occurring with other discourse-organizing devices. Caprario (2023), analysing classroom discourse, identifies an additional function of deciding in real time.

Thirdly, *I think* acts as a face-saving device signalling that a proposition can be questioned and further discussed, thus helping to manage relations between interlocutors (Baumgartner & House, 2010; Caprario, 2003; Coates, 1987; White, 2008). Kärkkäinen (2003, p. 105) adds that speakers can use *I think* to introduce a different perspective and to signal that they disagree with the prior turn. Myers (2010) observes that, in the discourse of blogs, *I think* mainly signals a relation to another person, rather than marking epistemic uncertainty.

In summary, the epistemic marker *I think* can perform different functions in spoken discourse, often simultaneously, depending on the context in which it occurs. Besides signalling the speaker’s perspective, it is commonly utilized as a device of discourse management and a face-saving device.

3 Methodology

3.1 Data for corpus analysis

This paper exploits the Corpus of Czech Students’ Spoken English (CCSSE) (Ježková, 2015) as the source of epistemic stance markers in question. The corpus was compiled at the University of Pardubice (Department of English and American Studies) and comprises informal spoken English (153,295 words) recorded in 228 monologues, 228 dialogues and 114 discussions. The participants were 228 native Czech learners of English (aged 19 to 22 years), 71% females and 29% males. All the participants were first-year university students (from three Czech universities) enrolled in a Teaching English as a Foreign Language programme. Their average level of proficiency in English was B2 (according to the CEFR), which is the officially declared level of secondary school leavers in the Czech Republic.

For the purpose of the current study, only one subset of CCSSE was analysed. The selected categories of uncertainty markers were excerpted from 114 student-student discussions on a given topic (the total number of words in the sub-corpus examined was 39,101; the average number of words per discussion was 343). The discussions covered eleven topics, mostly related to the students’ English

language learning experience or their university studies: the teaching profession, university tuition fees, the internet as a source of information, working while studying, studying abroad, professional agencies offering ready-made seminar papers, social networks in one's life, the school leaving exam, healthy eating, studying in one's hometown, living in university halls of residence. The discussions were designed to be interactive tasks and the participants, who were social equals and had known each other for some time, were expected to express their stance and justify their viewpoint, as illustrated in this sample excerpt from the corpus:

<Discussion no 4: working while studying>

<Speaker A, female, 20 years old, 11 years of learning English, 1 week in English speaking country>

<Speaker B, female, 19 years old, 9 years of learning English, 0 weeks in English speaking country>

<A> *Well I think that (er) it's good to work while you're studying because you need some money for your life.*

 Yes, it's true, but I think that you don't have a much time to your activities for example, when you do some sports then you can't do this because you are at work.

<A> *Sure and maybe you've got, you don't have a time for studying too much but, (er) when you've got some good work when you you have to work maybe only one day in week, so I think it's good.*

 Yeah it's good and you have more money for some parties and and so on. And you are independent person I think. That you don't need the money from your parents which is also good.

3.2 Corpus analysis

In the first stage of analysis, I searched the discussions for the linguistic forms listed in Table 1. In the second stage, I manually analysed all the occurrences collected because the uncertainty reading can be just one of their potential readings. Therefore, I examined them in their linguistic and situational context to ensure that they convey the speaker's uncertainty.

The manual analysis uncovered a large number of instances which did not convey uncertainty in the category of verbs, particularly modal verbs and the lexical verb *think* proved to be ambiguous in their nature. Interpretation of these verbs is highly context sensitive and, as highlighted by Holmes (1990), their reading "cannot be identified in a social and textual vacuum" (p. 186), which is reflected in (1) and (2).

- (1) a. <A> *So we both live in a in a hall of residence of University.*
 Yes. So the thing it's better than renting a flat [...]
 <A> *Why do you think?*
 *Well I'm not sure because I haven't tried yet, but well it **could** have some advantages. For example (er) you have more room (er) more, more space and (er) [...] you don't have to share bathroom with so many people.*
- b. *We are needed.*
 <A> *Yeah. And **could** you name any benefits?*
 Benefits? Well, there was a time when being teacher meant something and we are were resp= respected but I don't really think that's that's the case today.

The two exchanges in (1) demonstrate the indeterminate nature of modal verbs, meaning that one form can express different kinds of modality. Therefore, the semantic content of the modals in question was examined in conjunction with contextual co-occurrence patterns, that is, syntactic criteria that affect their epistemic and non-epistemic interpretation. The uncertainty reading belongs to the domain of epistemic modality because epistemic modality “qualifies the speaker’s commitment to the truth of the modalised proposition” (Huddleston & Pullum, 2002, p. 52), thus marking “to what extent one can rely on the information which is being conveyed by the proposition” (Vold, 2006, p. 65). Modal verbs interpreted epistemically typically occur with inanimate subjects and stative verbs. In contrast, non-epistemic meanings, sometimes termed “root” modality (Coates, 1983; Leech, 2004; Papafragou, 2000), refer to actions and events controlled by humans or other agents, implying that the subject is usually animate and the verb dynamic (Biber et al., 2021, p. 484).

Based on this epistemic-root distinction, only (1a) has been included in this study because here the speaker B uses *could* epistemically together with another uncertainty marker (*I'm not sure*) to clearly communicate a lack of certainty due to having no experience with renting a flat. Also, the modal verb *could* co-occurs with the inanimate subject *it* and the stative verb *have*. Contrarily, in (1b), *could* in the utterance *And could you name any benefits?* is employed in its root meaning to convey a request and co-occurs with the human subject *you* and the verb *name* denoting an activity. With respect to the focus of this paper, only epistemic readings of the modal verbs listed in Table 1 are within the scope of the analysis, namely epistemic possibility of *can*, *could*, *may*, and *might* (e.g., *I think it can be useful but the disadvantage is bigger*), hypothetical possibility/tentative prediction of *would* (e.g., *It would probably help me to, (er) be more independent*) and tentative inference/assumption of *should* (e.g., *Well, maybe, England. It it should be better*) (see Hoye, 2013, p. 82).

Nevertheless, the most challenging structure to interpret was the construction *I think*, which, as mentioned in 2.2, can communicate the speaker's deliberative opinion or the speaker's uncertainty, which is shown in (2a) and (2b) respectively.

- (2) a. <A> *So what do you think about internet social networks in our life?*
 *Okay, (er) so **I think** that internet social networks are really important (er) in in my life I really don't use those social networks like twitter, tumblr and such as these, but but the only one I have is facebook and that's just because most of my friends are there and that's like the cheapest or free way how to contact them and talk to them when we're separated.*
- b. *And the university dining hall? I have no experience.*
 <A> *Me neither. But (er) I, **I think** in university dining hall there's quite good food as I heard it from other students and I'm going to go there a few weeks so I, **I think** there will be good food for students and (er) wide choice of food and [...]*

In (2a), the student expresses his opinion on the role of social networks in his life, not communicating any doubt about the proposition. This example of deliberative *I think* also illustrates its typical context of occurrence—initial position and the presence of the complementizer *that* (see Coates, 1987; Holmes, 1990). On the other hand, in (2b), the student lacks sufficient evidence and thus tentatively communicates what she believes to be true. Following Caprario (2023) and Vold (2006), the construction *I think* is interpreted as tentative if it is possible to substitute it with an explicit uncertainty marker; for example, both utterances containing *I think* in (2b) could be paraphrased *Perhaps/maybe there's (will be) quite good food*. Accordingly, only the occurrences of *I think* at the uncertainty end (illustrated in 2b) are included in the current study (292 instances), while those at the certainty end (illustrated in 2a) are excluded from further discussion (267 instances).

Finally, it is essential to note that this categorization of *I think* may be problematic in some cases because the distinction between the two interpretations of *I think* is not always clear-cut and seems to be fuzzy. As Vold (2006, p. 70) points out, readings of polysemous markers can be related to each other and thus distinguishing between them may not be always relevant. Leech and Coates (1980) or Coates (1995) argue that such cases are not necessarily mutually exclusive and could be discussed in terms of merger.

4 Results and discussion

4.1 Overall findings

The analysis of 114 discussions in CCSSE yielded a total of 640 instances conveying the speaker's uncertainty. Regarding the major categories of uncertainty markers, the results in Table 2 show that lexical verbs are predominant, with the preponderance of *I think*. The other two frequent categories are modal adverbs and modal auxiliaries. This finding is congruous with Gablasova et al.'s (2024) study of L2 spoken data, which reveals that verbal expressions of epistemic stance (lexical verbs) are prevalent, followed by adverbial expressions. In L1 spoken production, Kärkkäinen (2003, p. 37) reports that the most common means to mark epistemic stance are lexical verbs (60.7%), mainly *I think* and *I don't know*, together with a few modal adverbs (18.8%) and epistemic modals (17.8%), which indicates that Czech speakers of English rely on the same categories of epistemic stance markers as L1 speakers and tend to employ them with roughly the same frequency.

Category of uncertainty marker	No. of occurrences (%)	Linguistic form	No. of occurrences	%
Lexical verb	364 (56.9%)	<i>(I) think</i>	292	45.6
		<i>(I) don't know</i>	35	5.5
		<i>(I) guess</i>	23	3.6
		<i>(I) suppose</i>	7	1.1
		<i>seem</i>	4	0.6
		<i>(I) believe</i>	2	0.3
		<i>(I) doubt</i>	1	0.2
Modal verb	129 (20.1%)	<i>would</i>	54	8.4
		<i>could</i>	35	5.5
		<i>can</i>	20	3.1
		<i>should</i>	8	1.3
		<i>may</i>	6	0.9
		<i>might</i>	6	0.9
Adverb	122 (19.1%)	<i>maybe</i>	97	15.2
		<i>probably</i>	22	3.4
		<i>perhaps</i>	2	0.3
Adjective	21 (3.3%)	<i>presumably</i>	1	0.2
		<i>not sure</i>	15	2.4
Noun	4 (0.6%)	<i>possible</i>	6	0.9
		<i>possibility</i>	4	0.6
Total		640		

Table 2: Markers of the speaker's uncertainty in CCSSE

As for the most frequently used forms in CCSSE, Table 2 shows that Czech learners of English favour three items when conveying uncertainty, that is, the lexical verb *think*, the modal adverb *maybe* and the modal verb *would*. The occurrences of these three items make up approximately 69% of all the data, while the remaining 17 forms represent only 31% of the uncertainty markers collected. These results are in line with the prior studies of L2 spoken production by Gablasova et al. (2017; 2024), who observe that a small set of epistemic markers accounts for almost two-thirds of all stance markers identified (2017, p. 630), with *I think* (44.3%) and *maybe* (16.5%) being the most common lexical verb and adverb respectively (2024, pp. 197–198). Likewise, a limited range of high-frequency epistemic stance expressions is found in spoken American English by Kärkkäinen (2003), who thus concludes that the expression of epistemic stance follows “a very regular and routinized pattern in everyday speech” (p. 97). Although *I think* in her research outnumbers the other forms and categories of epistemic stance markers, it represents only 11.5% of the total. However, in CCSSE, *I think* is much more dominant (45.6%), which might suggest that Czech L2 speakers tend to overuse it.

In summary, Precht (2003, 2006) highlights that even though there is a wide range of forms available for us to express our commitments, “it would seem that we are culturally ‘programmed’ to use a very limited, very specific subset of these options” (2003, p. 133). This applies to the analysis of CCSSE as well because Czech L2 speakers prefer a restricted set of linguistic forms to convey their uncertainty (see Table 2) and tend to use them in recurrent patterns (see 4.2).

4.2 Distributional patterns of uncertainty markers in CCSSE

This section deals with the placement of uncertainty markers in utterances, but only those items that show syntactic mobility are taken into account here. This means that all the occurrences of modal verbs (129 instances) and the lexical verb *seem* (4 instances) have been excluded from further discussion because their position in a clause is constrained syntactically, see (3) and (4).

- (3) *(Er) [...] It's hard to say [...] (er) [...] I think that the education **could** be free for everybody, (er)...*
- (4) *they just post everything like what they're doing right now and they have thousands of friends even though they don't know them and they're all around the world so that's kind of crazy, it **seems** like that to me.*

Both the modal auxiliary *could* in (3) and the lexical verb *seem* in (4) are placed after the subject because it is their only appropriate position in these clauses. In

contrast, the other uncertainty markers collected can occur in different positions. For example, the utterances in (5) reflect the syntactic mobility of *I guess*, which commonly appears in initial or final position in the corpus discussions.

- (5) a. *I'm not quite sure but, (er) it will be probably somewhere in the USA I guess.*
 b. *Okay, so I guess you are from economic school.*

Figure 1 demonstrates that the dominant pattern in CCSSE is the placement of uncertainty markers in initial position, that is, before a proposition itself, as in (5b), where *I guess* precedes the propositional information *you are from economic school*, or in (6), where *maybe* communicates the speaker's uncertainty about the subsequent proposition *it's cheaper*.

- (6) *Well, it's [...] (er) easier to get to school and maybe it's cheaper [...] the transportation.*

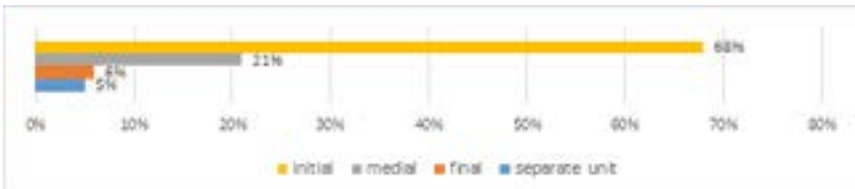


Figure 1: Placement of syntactically mobile uncertainty markers in CCSSE

Additionally, the results in Table 3 show that initial position is preferred with all the categories of syntactically mobile uncertainty markers in CCSSE. Biber et al. (2021, p. 963) explain that the overwhelming preference for this position reflects the primary function of stance markers, that is, serving as a frame that communicates the personal perspective to propositional information. Since speakers generally tend to first identify their stance, initial position is considered to be unmarked (Kärkkäinen, 2003, p. 183).

Regarding epistemic lexical verbs, they typically co-occur with the first-person subject in singular, which overtly marks the speaker's stance, as in (5) or (7) (see Kärkkäinen, 2003; Thompson, 2002). Since they express the speaker's stance towards an issue in the complement clause, Thompson (2002) labels them "complement-taking predicates or CTP-phrases". However, she argues that they can be reanalysed as epistemic parentheticals equivalent to epistemic adverbial phrases due to the fact that they "can and do float away from CTP position" (p. 134) (see Louro & Harris, 2013). This is apparent in (5b), where *I guess* could be moved from initial position (*so, you are from economic*

school I guess) or could be replaced with an epistemic adverb (*so probably you are from economic school*).

	Initial position	Medial position	Final position	Separate unit	Total
Lexical verb	262	51	25	22	360
Adverb	64	50	6	2	122
Adjective	14	6	1	-	21
Noun	3	-	1	-	4
Total	343	107	33	24	507

Table 3: Positions of individual categories of syntactically mobile uncertainty markers in CCSSE

Considering the CTP-phrase *I think*, the most frequent uncertainty marker in CCSSE, 225 out of 292 (i.e., 77%) instances appear in initial position and a common pattern is the absence of the complementizer *that* after *I think*:

- (7) *Well, I would like to try the United States if I would have to choose.*
 <A> *Me too [...] but I think it's (er) expensive.*
 Well, yes but I think this school can give us some money to go there or I am not sure.

This finding is in line with prior studies dealing with *I think* (Aijmer, 1997; Ariel, 2009; Baumgarten & House, 2010; Louro & Harris, 2013; Szczyrbak, 2020), which conclude that the zero complementizer dominates in spoken discourse. Ariel (2009, pp. 20–21) notes that the complementizer is often absent because the CTP-phrase serves as a mere epistemic framing for the complement that contains contextually more relevant information. Aijmer (1997) adds that the recurrent phrase *I think* has been pragmaticalized and “developed into a discourse marker or modal particle which is syntactically a speech-act adverbial” (p. 1) (see Thompson, 2002). Regarding the occasional presence of the complementizer *that* in spoken interaction, Kaltenböck (2009) clarifies that it can be motivated by rhythmic purposes or by bridging a hesitation phase, which implies that the complementizer may act like a filler allowing the speaker to get some time.

The other common position of uncertainty markers in CCSSE is medial position, with 21% of syntactically mobile uncertainty markers (see Figure 1). Table 3 indicates that it is epistemic adverbs that often occur medially (41% of adverbs), that is, within a proposition, as in:

- (8) *That's probably true.*

It has been found out that 33 out of 97 (i.e., 34%) occurrences of *maybe*, the predominant adverb in CCSSE, are placed medially. However, a closer examination of these occurrences reveals that medial position of *maybe* is sometimes inappropriate (9), in contrast to *maybe* in initial position (6).

- (9) a. *So it would be **maybe** little problem but I think it worth to try it...
 b. *you have to **maybe** be with touch, in touch with, with...

Such inappropriate usage of *maybe* may stem from the influence of the students' mother tongue because in Czech the corresponding adverb 'možná' commonly appears in medial position. The general preference for initial position of *maybe* is reported by Castello (2023, pp. 259-260), who observes that *maybe* is by far the most frequently utilized epistemic adverbial in turn-initial position, with L2 speakers (Italian and Swedish learners) employing it more extensively than L1 speakers. Hoyer (2013) explains that although epistemic adverbs are often found in medial position, their most frequent placement is initial, which is associated with their thematic role: "their scope is sentential making them ideal 'scene-setters' for what is to follow" (p. 185).

Two remaining positions, at the end of an utterance and an utterance itself, are rare in CCSSE (6% and 5% respectively) and mostly associated with epistemic lexical verbs (see Table 3):

- (10) *If you are beginner and trainee, I think you have to pass some exams and after this, this work isn't really time time wasting* **I think**.
- (11) (er) *Personally, I would not have the courage to study anywhere but in my hometown. [...] Yeah, because because of this, like, I would be afraid probably.*
 <A> *Yeah, I think so.*

Final position of an epistemic lexical verb, that is, after a proposition, as in (5a) or (10), could be perceived as a kind of afterthought due to the unplanned nature of spoken face-to-face interaction. Biber et al. (2021) state that speakers may realize that their previous proposition "needs qualification, and so add a hedging comment clause" (p. 972). Regarding markers that occur as separate utterances, mainly *I (don't) think so* is used this way to communicate that "one believes that something is true" (Merriam-Webster, n.d., Definition 1), as in (11), where the epistemic formula *I think so* communicates that student A believes that B's proposition is true.

Finally, the analysis of distributional patterns of uncertainty markers shows that they commonly co-occur in CCSSE:

- (12) *(er) yes, I have thi= thought about it, too. But, (er) **I think it would** be [...] very interesting but also dangerous and **maybe I, I I would** go there [...] but only with someone else. **I think** on my own it **would, I would I would** be scared **I think**.*
- (13) *Well, I'm **not so sure** about the pricing of flat, flats nowadays, **I guess** if you split with a lot of people it **would** be, **presumably**, cheaper.*

Examples (12) and (13) reflect co-occurrence of uncertainty markers in the immediate context within one's turn. In such congruous combinations, these markers express the speaker's high degree of uncertainty and hesitation. This finding is in alignment with Kirkham's (2011) claim that co-occurrence of epistemic stance markers may contribute towards constructing "a more tentative stance" (p. 211) or with Szczyrbak's observation (2020) that it may attenuate the speaker's precision of a proposition. Additionally, Examples (12) and (13) illustrate collocations of modal verbs with adverbs of uncertainty/likelihood which, according to Hoye (2013), signal the epistemic reading of the modal. However, only 15 cases of modal-adverb clusters have been identified in CCSSE, with *would + maybe* (5 instances), and *would + probably* (4 instances) being the most frequent ones.

4.3 Functions of tentative *I think* in CCSSE

When examining contexts of occurrence of the preponderant uncertainty marker *I think*, some of its communicative functions attested in previous studies (see 2.2) were identified in the CCSSE discussions as well. The analysis reveals that Czech learners do not employ *I think* only as an uncertainty marker, but also as an interactional device helping to organize their discourse and as a device of politeness. Additionally, these functions of *I think* seem to be often performed simultaneously, meaning that *I think* can signal initiating a turn and communicate the speaker's belief at the same time (15 and 16) or mitigate disagreement alongside marking the speaker's tentative perspective (17 and 18) (see Baumgartner & House, 2010). Nevertheless, some instances convey just the speaker's uncertainty, particularly when *I think* is found in medial position, where it is associated with approximation (see Hoye, 2013):

- (14) *At the beginnings it was hard but after **I think** two months you are used to [...] to speak with mistakes.*

As a boundary marker, tentative *I think* routinely signals turn-taking in CCSSE, typically appearing as a second part of an adjacency pair:

- (15) *But if you rent a flat you can live with people you like. If you don't want them they will not annoy you by by, I don't know, hearing music loud and so...*
 <A> **I think** *I will maybe in three years rent a flat with my friends but really now I don't have any in this city so it's best for me if I live at home.*
- (16) <A> *Yes that's what students usually do. As as our teacher said 'Students lie and students cheat'.*
 I think *the blame is on both sides ... (er) but more on the students because, the agencies, it's their business.*

In (15) and (16), employing *I think* in the starting-point function, the students react to the preceding turn by the other interlocutor and at the same time convey their degree of certainty at that point. Whereas *I think* in (15) expresses the student's uncertainty, in (16) its tentative and deliberative interpretation seem to be merged because it could be paraphrased both as *perhaps the blame is on both sides* and *my opinion is that the blame is on both sides*.

As a politeness device, tentative *I think* is commonly employed in CCSSE when negotiating the point of a discussion, particularly to mitigate potential disagreement:

- (17) <A> *I don't like fast foods and it's not tasty at all. And organic food I guess I've never tried, in in some kind of vegan ... I've never been there.*
 *Okay. **But I think**, on the other hand, that in the, in the university dining hall, there are a lot of students and maybe, (er) you may, (er) you may...*
 <A> *Yeah, it's a bit crowded and you can wait long time, so [...] but it's cheap and some= sometimes there are no people and you get your food as quick as you get there.*
- (18) <A> *But, (er) there are some disav=, disadvantages such as for some students it could be harder than for the other students for example the students for, students of grammar schools.*
 *Well, I take your point, **but** on the other side **I think** (er) some students, (er) from, for which it's harder; (er) they should learn more and then they can pass it easier.*
 <A> *Yes, **but I think** some students aren't able to learn more and improve their self.*

In (17) and (18), the interlocutors do not share the same perspective on the issue discussed, which might be perceived as potentially face-threatening. Therefore, using the connector *but* followed by *I think*, they express partial agreement with the preceding proposition and then justify their belief.

Since this study is concerned only with tentative *I think*, other functions of this epistemic phrase reported in the prior studies (Baumgarten & House, 2009; Kärkkäinen, 2003) have not been observed in the CCSSE discussions. Some of

them, for example, closing one's turn or expressing agreement, could likely be associated with deliberative *I think*.

5 Conclusion

The analysis of student-student discussions in the Corpus of Czech Students' Spoken English indicates that expressing uncertainty in Czech learners' face-to-face interaction tends to be regular and routinized both at the level of linguistic form and in terms of distributional patterns.

The speaker's uncertainty is conveyed through a restricted set of linguistic forms in CCSSE. Advanced Czech learners of English favour the lexical verb *think*, which dwarfs all the other uncertainty markers identified. The other two frequently employed items are the adverb *maybe* and the modal auxiliary *would*. These three preferred uncertainty markers outnumber all the others, accounting for 69% of all the data. Similar tendencies have been observed both in L1 and L2 spoken discourse (see Gablasova et al., 2024; Kärkkäinen, 2003). Distributional patterns and positional preferences of the markers in question appear to be recurrent in the corpus discussions. Nearly 70% of syntactically mobile uncertainty markers are placed in initial position, which implies that the students tend to establish their stance before a proposition itself. Epistemic lexical verbs are typically used with the first-person singular subject and without the complementizer *that*. Additionally, uncertainty markers often co-occur in the immediate context within one's turn and thus upgrade the speaker's degree of uncertainty and hesitation. Some of the findings could be discussed in terms of the speakers' L1 background (Czech), which may account for a preference for specific uncertainty markers as well as positional preferences, for example, favouring the adverbs *maybe* and *probably* and placing them in medial position.

Regarding the usage of *I think*, the predominant uncertainty marker in CCSSE, the investigation confirmed its multifunctionality attested in previous studies. Besides marking the speaker's uncertainty, *I think* can simultaneously act as a discourse-organizing device or a device of politeness. In Baumgartner and House's (2010) words, *I think* expressing subjective evaluation "is usually embedded in an utterance or sequence of utterances which fulfil other functions in the discourse" (p. 1191) (see Aijmer, 1997). In CCSSE, *I think* tends to convey the speaker's uncertainty at a particular point in an exchange alongside signalling turn-taking or it often justifies the speaker's belief when helping to negotiate alignment. Additionally, it has been found that, in terms of expressing degrees of certainty, the phrase *I think* seems to be indeterminate in its nature because some instances could be assigned both the tentative and deliberative reading. Thus, distinguishing between these two readings of *I think* may not be always relevant and some instances could be interpreted as cases of merger.

Finally, limitations of the present study and directions for future research should be mentioned. Although this study uncovers tendencies and preferences in the usage of selected categories of uncertainty markers in spoken production by Czech learners of English, due to its narrow focus, it does not address some related issues. For example, it would be useful to address possible influences of the speakers' L1 (Czech) and to compare systematically the usage of uncertainty markers by Czech speakers against native speakers' performance. Furthermore, as prosody may affect the interpretation of epistemic stance markers, it would be helpful to analyse the transcribed data alongside recordings. Regarding the marker *I think*, since the distinction between its tentative and deliberative use can be greatly blurred, it would be beneficial to explore the usage of this epistemic phrase as such in order to capture all its interactional functions. The generalizability of findings could be increased by extending the corpus and enlarging the scope of research; other speaking tasks, a wider range of forms conveying uncertainty or learners with different proficiency levels could be included in future research. Also, individual preferences in linguistic choices among speakers could be taken into consideration. It might be interesting to examine individual variables, for example, to explore differences in the use of stance markers in terms of gender.

References

- Aijmer, K. (1997). I think—an English modal particle. In T. Swan & O. J. Westvik (Eds.), *Modality in the Germanic languages* (pp. 1–47). Mouton de Gruyter. <https://doi.org/10.1515/9783110889932.1>
- Ariel, M. (2009). Discourse, grammar, discourse. *Discourse Studies*, 11(1), 5–36. <https://doi.org/10.1177/146144560809849>
- Baumgarten, N., & House, J. (2010). 'I think' and 'I don't know' in English as lingua franca and native English discourse. *Journal of Pragmatics*, 42(5), 1184–1200. <https://doi.org/10.1016/j.pragma.2009.09.018>
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5, 97–116. <https://doi.org/10.1016/j.jeap.2006.05.001>
- Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of spoken and written English*. John Benjamins.
- Brezina, V. (2009). “We only say we are certain when we are not”: A corpus-based study of epistemic stance. In J. Levická & R. Garabík (Eds.), *The proceedings of the fifth international conference SLOVKO 2009* (pp. 41–63). Tribun.
- Caprario, M. (2023). Multifunctionality of epistemic stance markers: Variation across disciplines and speaker roles in classroom discourse. *Contrastive Pragmatics*, 4, 213–242. <https://doi.org/10.1163/26660393-bja10066>
- Castello, E. (2023). Stance adverbials in spoken English interactions: Insights from corpora of L1 and L2 elicited conversations. *Contrastive Pragmatics*, 4, 243–273. <https://doi.org/10.1163/26660393-bja10065>
- Coates, J. (1983). *The semantics of the modal auxiliaries*. Croom Helm.

- Coates, J. (1987). Epistemic modality and spoken discourse. *Transactions of the Philological Society*, 85(1), 110–131. <https://doi.org/10.1111/j.1467-968X.1987.tb00714.x>
- Coates, J. (1995). The expression of root and epistemic possibility in English. In J. L. Bybee & S. Fleischman (Eds.), *Modality in grammar and discourse* (pp. 55–66). John Benjamins.
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment—Companion volume*. Council of Europe Publishing. <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Crompton, P. (1997). Hedging in academic writing: Some theoretical problems. *English for Specific Purposes*, 16(4), 271–287. [https://doi.org/10.1016/S0889-4906\(97\)00007-0](https://doi.org/10.1016/S0889-4906(97)00007-0)
- Du Bois, J. W. (2007). The stance triangle. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction* (pp. 139–182). John Benjamins. <https://doi.org/10.1075/pbns.164.07du>
- Gablasova, D., Brezina, V., McEnery, T., & Boyd, E. (2017). Epistemic stance in spoken L2 English: The effect of task and speaker style. *Applied Linguistics*, 38(5), 613–637. <https://doi.org/10.1093/applin/amv055>
- Gablasova, D., Harding, L., Brezina, V., & Dunlea, J. (2024). Expressions of epistemic stance in computer-mediated L2 speaking assessment: A corpus-based approach. *International Journal of Learner Corpus Research*, 10(1), 183–215. <https://doi.org/10.1075/ijlcr.00044.gab>
- Halliday, M. A. K. (2004). *An introduction to functional grammar*. Edward Arnold.
- Holmes, J. (1990). Hedges and boosters in women’s and men’s speech. *Language and Communication*, 10(3), 185–205. [https://doi.org/10.1016/0271-5309\(90\)90002-S](https://doi.org/10.1016/0271-5309(90)90002-S)
- Hoye, L. (2013). *Adverbs and modality in English*. Routledge.
- Huddleston, R., & Pullum, G. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.
- Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford University Press.
- Hyland, K. (1995). The author in the text: Hedging scientific writing. *Hong Kong Papers in Linguistics and Language Teaching*, 18, 33–42.
- Hyland, K. (1999). Disciplinary discourses: Writer stance in research articles. In C. Candlin & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 99–121). Routledge.
- Hyland, K., & Zou, H. (2021). “I believe the findings are fascinating”: Stance in three-minute theses. *Journal of English for Academic Purposes*, 50, 100973. <https://doi.org/10.1016/j.jeap.2021.100973>
- Kaltenböck, G. (2009). Initial I think: Main or comment clause? *Discourse and Interaction*, 2(1), 49–70.
- Kärkkäinen, E. (2003). *Epistemic stance in English conversation: A description of its interactional functions, with a focus on I think*. John Benjamins. <https://doi.org/10.1075/pbns.115>
- Kirkham, S. (2011). Personal style and epistemic stance in classroom discussion. *Language and Literature*, 20(3), 201–217. <https://doi.org/10.1177/0963947011413>
- Leech, G. (2004). *Meaning and the English verb*. Longman.
- Leech, G., & Coates, J. (1980). Semantic indeterminacy and the modals. In R. Quirk, S. Greenbaum, G. Leech & J. Svartvik (Eds.), *Studies in English linguistics* (pp. 79–99). Longman.

- Louro, C. R., & Harris, T. (2013). Evolution with an attitude: The grammaticalisation of epistemic/evidential verbs in Australian English. *English Language and Linguistics*, 17(3), 415–443. <https://doi.org/10.1017/S1360674313000105>
- Merriam-Webster. (n.d.). I think so. In *Merriam-Webster.com dictionary*. Retrieved October 16, 2025, from [https://www.merriam-webster.com/dictionary/I think so](https://www.merriam-webster.com/dictionary/I%20think%20so).
- Myers, G. (2010). Stance-taking and public discussion in blogs. *Critical Discourse Studies*, 7(4), 263–275. <https://doi.org/10.1080/17405904.2010.511832>
- Papafragou, A. (2000). *Modality: Issues in the semantics-pragmatics interface*. Elsevier.
- Precht, K. (2003). Great versus lovely: Stance differences in American and British English. In P. Leistyna & Ch. F. Meyer (Eds.), *Corpus analysis: Language structure and language use* (pp. 133–151). Rodopi. https://doi.org/10.1163/9789004334410_009
- Precht, K. (2006). Stance moods in spoken English: Evidentiality and affect in British and American conversation. *Text & Talk*, 23(2), 239–257. <https://doi.org/10.1515/text.2003.010>
- Szczyrbak, M. (2020). Verb patterns in trial discourse: The case of I think. *Discourse and Interaction*, 13(2), 119–135. <https://doi.org/10.5817/DI2020-2-119>
- Thompson, S. A. (2002). “Object complements” and conversation: Towards a realistic account. *Studies in Language*, 26(1), 125–163. <https://doi.org/10.1075/sl.26.1.05tho>
- Vold, E. T. (2006). Epistemic modality markers in research articles: A cross-linguistic and cross-disciplinary study. *International Journal of Applied Linguistics*, 16(1), 61–87. <https://doi.org/10.1111/j.1473-4192.2006.00106.x>
- White, P. (2008). Modality as dialogue: A Bakhtinian reanalysis of epistemic stance. *WORD*, 59(1–2), 143–167. <https://doi.org/10.1080/00437956.2008.11432584>
- Wu, B., & Paltridge, B. (2021). Stance expressions in academic writing: A corpus-based comparison of Chinese students’ MA dissertations and PhD theses. *Lingua*, 253, 103071. <https://doi.org/10.1016/j.lingua.2021.103071>
- Xu, F., & Cao, R. (2023). Epistemic stance in Chinese L2 spoken English: The effect of grade and genre-specific questions. *Languages*, 8(1), 15. <https://doi.org/10.3390/languages8010015>

Sources

Ježková, Š. (Ed.). (2015). *Corpus of Czech students’ spoken English*. University of Pardubice.

Petra Huschová is Assistant Professor of English Linguistics at the Department of English and American Studies, University of Pardubice, Czech Republic. She specializes in morpho-syntax, stylistics and pragmatics. Her main research fields of interest include modality, hedging devices and speech acts. She currently focuses on analysing spoken learner discourse.

ORCID ID: <https://orcid.org/0000-0001-7157-5301>

Address: Petra Huschová, Department of English and American Studies, Faculty of Arts and Philosophy, University of Pardubice, Studentská 84, Pardubice 530 09, Czech Republic. [e-mail: petra.huschova@upce.cz]