

NO DOUBT, BUT... ON CONNECTIVE FUNCTIONS OF EPISTEMIC MARKERS

LUCIE POLÁKOVÁ¹ – JANA ŠINDLEROVÁ² – BARBORA ŠTĚPÁNKOVÁ³

¹Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics,
Charles University, Prague, Czech Republic (ORCID: 0000-0002-4879-5530)

²Department of Linguistics, Faculty of Arts, Charles University, Prague, Czech
Republic (ORCID: 0000-0002-9610-4618)

³Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics,
Charles University, Prague, Czech Republic (ORCID: 0000-0001-9498-7165)

POLÁKOVÁ, Lucie – ŠINDLEROVÁ, Jana – ŠTĚPÁNKOVÁ, Barbora: No doubt,
but... on Connective Functions of Epistemic Markers. *Journal of Linguistics*, 2025, Vol. 76,
No 1, pp. 75 – 84.

Abstract: Epistemic markers are usually treated on the basis of their primary function to express the level of certainty of a speaker about a given proposition. They are often described as items operating on higher levels than syntax. In this paper, we focus on cases in which epistemic markers actually contribute to the syntactic organization of text, either by developing a text-organizing function or a discourse connective function. Specifically, we address three patterns which appeared in the corpus data: a confrontation of modalities, a function of a topic orientation marker, and a contrastive pattern with concessive features.

Keywords: epistemic markers, connectives, concessive constructions, Czech

1 INTRODUCTION

Epistemic markers (EMs)¹ are function words which express the level of certainty of a speaker towards a proposition. Their predominant function is pragmatic and as such, they belong to the group of pragmatic markers.² The certainty levels expressed range from full certainty to probability, possibility, down to uncertainty or doubt. Typical examples for Czech are expressions like *jistě* ‘certainly’, *asi* ‘perhaps’, *možná* ‘maybe’ or *stěží* ‘hardly’, i.e. words the meaning of which is defined by their subjective, attitudinal and modifying character.

¹ This study is part of a project investigating expressions that are considered to be typical epistemic markers (EMs). According to Komárek et al. (1986), we also include evidential (*ocividně* ‘obviously’) and confirmatory (*samořejmě* ‘of course’) expressions in this group. For the purposes of this study, we use the term ‘EM’ (instead of ‘examined expression’, for example) even when the epistemic modality of a marker is completely or significantly reduced.

² In Czech linguistics, epistemic markers belong to particles, a word class defined through their attitudinal and syn-syntactic character (cf. e.g. Komárek et al. 1986).

EMs often etymologically derive from adverbials (*určitě, jasně, zřejmě*), so they retain word class ambiguity. The same expression can be used either as an autosemantic adverbial (1a), or as a synsemantic pragmatic marker (1b).

(1a) *Její vraždu viděl zcela jasně.*
‘He saw her murder quite **clearly**.’

(1b) *Jasně že ti zavolám.*
‘I will call you **for sure**.’

The shift from adverbial to pragmatic meaning seems to be quite a systematic language change across the respective group of words (cf. Šindlerová et al. 2023; Traugot 1989), the original adverbial meaning may even disappear completely with time.

Secondarily, epistemic pragmatic markers can also have other, derived pragmatic functions, e.g. they can gain an independent syntactic status and become affirmative markers, see (2).

(2) *Zavoláš mi? – Jistě!*
‘Would you call me? – **Sure!**’

In this study, we approach epistemic markers **from a syntactic perspective**, i.e. from the point of view of **their possible connective functions**. During our analysis of these expressions, we came across a certain group of syntactic patterns, or contexts,³ where the original epistemic marker occurs in a position typical of a connective device or a certain part of it, while also exhibiting similar semantics.

The connective function of EMs has not been consistently described using Czech data yet. It is not mentioned either in standard grammars, or dictionaries, including the most recent ones. Nevertheless, individual studies of Czech particles suggest that connective function may be in fact quite common for various pragmatic markers (see e.g. Kolářová 1998 on the connective functions of *teda* (roughly) ‘so’, or Mladová 2008 on the functions of the prototypical focusing particle *také* ‘also’).

We will try to answer the following research questions:

1. What are the types of constructions and contexts for the EMs in connective function?
2. Does the different epistemic strength of individual EMs have an impact on the interpretation of the propositions?
3. What other aspects do the EMs bring into the interpretation of an utterance compared to neutral connective devices?

³ We prefer to speak about contexts, where the setting of the studied marker(s) is inter-sentential, goes beyond the sentence boundary.

We will focus on the description of such patterns with epistemic markers that appear to be the most distinctive demonstration of text linking functions in our data.

2 DATA AND METHOD

The research is inspired by the works on the SEEM-CZ project, within which a lexicon of Czech epistemic and evidential markers SEEMLex is being composed (cf. Štěpánková et al. 2024) via the annotation of the data of the InterCorp corpus, specifically the core part of its Czech and English sections (Rosen et al. 2022). A list of EMs was compiled manually after analysis of the Czech grammars, dictionaries, and corpora. The forty most frequent expressions in the InterCorp corpus were then selected from this list and annotated with 100 examples each. The annotation scheme (see Štěpánková et al. 2024) includes information such as the use of the EM and the presence of contrast in the sentence. During the annotation, the connective function of certain EMs was identified based on these clues. The patterns found in the InterCorp data were then confronted with their occurrences in the large Czech National Corpus (Křen et al. 2024), in the annotated data of the newest version of the Prague Discourse Treebank 4.0 (Synková et al. 2024) and in CzeDLex, the lexicon of Czech discourse connectives (Mírovský et al. 2021).

3 ANALYSIS

In this section, we describe the patterns of the connective use of EMs we found in the corpus data. We proceed from cases where the presence of an EM supports, emphasizes, or builds upon a contrastive syntactic relation, to cases that signal a topic diversion, up to those with a clear connective function.

3.1 Confrontation of modalities

The first pattern contrasts two EMs expressing differing – sometimes even opposing – degrees of certainty, by juxtaposing them in a comparative construction: maybe A, but definitely B, see (3). Typically, in such constructions, the confrontational meaning arises from the opposition between the semantic features of the lexical items themselves. This semantic tension remains strong even in the absence of an explicit contrastive connective. Furthermore, one part of the construction may be negated (yielding opposite polarity, as in (3a), though this is not always the case (3b).

(3a) *V lepším případě možná i šest – ale rozhodně ne víc.* (InterCorp)
‘**Maybe** six at best – but **certainly** no more.’

(3b) *Kniha Viktimologie pro forenzní praxi nepatří podle Ludmily Čírtkové do ruky těm, o kterých v publikaci najdeme následující vtip – možná cynický, ale rozhodně výstižný pro naši dobu [...].* (SYNv13)

‘According to Ludmila Čírtková, the book *Victimology for Forensic Practice* does not belong in the hands of those about whom the following joke can be found in the publication – **perhaps** cynical, **but certainly** appropriate for our times [...].’

Since the conjoining “ability” of the EMs in this case lies only in the cooperation within a semantic (lexical) contrast, we do not consider this a case of proper connective use. Nevertheless, there is a clear text-orienting function present in the construction.

3.2 Topic orientation

A different pattern is typical of the word *každopádně* ‘anyway’ in Czech when it appears in the left periphery of a sentence. There are (at least) two subtypes of this pattern, though the boundaries between them are often unclear. The first subtype typically involves a list of possibilities stated by the speaker, followed by a summarizing conclusion introduced by *každopádně* (4). The conclusion presents an idea that the speaker considers valid regardless of whether the previous context is relevant. In this respect, *každopádně* works as a discourse marker⁴, more specifically, as a marker of topic orientation (cf. Fraser 2009).

(4) *Prosila ji, aby přijela, protože se cíti slabá, nebo je něco v neporádku s domem, každopádně protože potřebuje pomoc.* (SYNv13)
‘She begged her to come because she felt weak or something was wrong with the house, **either way** because she needed help.’

This function of *každopádně* is even more perceivable in an inter-sentential setting (5).

(5) *Možná lod' bloudila vesmírem velmi dlouho, nebo její posádka změnila v zoufálé snaze uniknout nebezpečí kurz. Každopádně se nenašlo nic, podle čeho by šlo zjistit, odkud pocházela.* (SYNv13)
‘Perhaps the ship had been wandering in space for a very long time, or its crew had changed course in a desperate attempt to escape danger. **Either way**, there's nothing to indicate where it came from.’

In other cases, *každopádně* signals a more distinct shift in topic, or a transition from unimportant (irrelevant) information to an important (relevant) one (6). In such instances, the expression's relation to the preceding context is not semantic; the clauses adjacent to the marker are only loosely related in terms of content, or may even be entirely unrelated.

⁴ In accordance with Fraser (2009), we understand “discourse markers” to be a class of pragmatic markers that signal an aspect of the organization of ongoing discourse.

(6) *Než jsem vyrazil, postříkal jsem vnitřek vozu deodorantem s vůní borovic, ale moc rozdílu v tom nebylo. Každopádně jsem myslел pouze na jediné – aby se Georgina neprobudila, dřív než dorazím do Raytonu.* (SYNv13)
 ‘I sprayed the inside of the car with pine scented deodorant before I left, but it didn’t make much difference. **Anyway**, my only thought was to make sure Georgina didn’t wake up before I got to Rayton.’

In both cases, the epistemicity of *každopádně* is weakened. In the former case, the speaker summarizes the previous utterances, which cannot be verified, with a general statement they regard as certain. In the latter case, the speaker dismisses the prior statements as irrelevant in contrast to the following one – they shift their attention from one situation (or an aspect of a situation) to another.

3.3 Concessive pattern

One of the most compelling examples of a connective function of EMs occurs when an epistemic expression stands as a direct **part of the connecting expression** in a correlative contrastive pattern (7).

(7) *Císař možná křesťany podporoval, ale u dvora bylo plno lidí, kteří se na tu novou sektu dívali v lepším případě s pobavením, v horším případě s podezřením či dokonce nepokrytě nepřátelsky.* (SYNv13)
 ‘The Emperor **may** have supported the Christians, **but** the court was full of people who looked at the new sect with amusement at best, suspicion at worst, or even outright hostility.’

Here, the EM stands in the left part of the clause (sentence), taking the position and function of the prototypical Czech connective *sice* in the correlative multiword connective *sice – ale*, cf. (8).

(8) *Císař sice křesťany podporoval, ale u dvora bylo plno lidí, kteří se na tu novou sektu dívali v lepším případě s pobavením, v horším případě s podezřením či dokonce nepokrytě nepřátelsky.*

Sice prototypically signals the speaker’s admission of the validity of the content of the proposition in which it appears, a partial assent. It is non-autonomous in that it presupposes some contradiction in the second proposition. Sentences with *sice – ale* are typically formally analyzed as adversative in Czech grammars (Grepl and Karlík 1998, p. 341), nevertheless their status as a paratactic formulation of a concessive relation is sometimes mentioned (Karlík 1995, p. 112). To our knowledge, English grammars do not recognize a multiword connective of this type. Rather, translations reveal that other linguistic devices are usually employed, including *may* (9), *while*,

although, etc. (cf. Vavřín and Rosen 2015). Compare also example (10), where, in the English translation, the acknowledgment of the content of the first proposition is conveyed through the intensifying use of *did*.

(9) *Já vždycky říkám, že kachna je sice dost tuhá, ale má svou zvláštní chuť.*
‘I always say that duck **may** be tough, **but** it has its own special taste.’ (InterCorp)

(10) *Tato strategie úspor sice vyřešila problémy s bilancí, měla však za následek nízký růst a obrovskou nezaměstnanost.*
‘The savings strategy **did** sort out the imbalances, **but**, in turn, resulted in low growth and increasingly high unemployment.’ (InterCorp)

The concessive interpretation of the constructions with modals is discussed e.g. by Palmer (2001, p. 31), or Leclerq (2024), who offers an account of a concessive *may/might* construction as a newly developed strong pattern in English syntax. Leclerq speaks primarily about the secondary grammaticalization of a modal verb (a potential epistemic marker) as a connective device, thus we believe that other epistemic markers also do have a potential to constitute contrastive and concessive relations.

Other expressions in English, which are, more rarely, used in the same position/setting, are also related to veridicality: apart from *may/might*, literature focused on English primarily mentions the use of: *true, fact, well, indeed, granted...* in these contexts (König 1988, pp. 154–155). In the more conversational or argumentative settings, expressions like *no doubt* or *of course* were identified. Also the German *zwar*⁵ etymologically comes from *es ist wahr* ‘it is true’.

The modal expressions occurring in concessive connective contexts express varying degrees of certainty – from high certainty (*bezpochyby, určitě, rozhodně*, etc.) to low certainty (*možná*). We believe that these markers can indicate different communicative functions and speaker strategies. For example, in (11), the use of the high-certainty expression *bezpochyby* ‘no doubt’ does not primarily indicate the speaker’s strong epistemic conviction regarding the proposition. Rather, it reflects a highly polite attitude toward the interlocutor’s claim, preceding the speaker’s own contradictory statement. Similarly, in (12), *možná* is used to express a directive assumption⁶ about the interlocutor’s state of mind. In both cases, the initial proposition is weakened in favour of the following, contradictory claim.⁷

⁵ The German construction *zwar – aber* is a direct parallel to Czech *sice – ale*.

⁶ By directive assumption we mean a situation in which the speaker firmly assumes a particular (agreeing or disagreeing) stance of their communication partner, but the acknowledgement or denial from the partner is not expected at all.

⁷ Cf. also Rossari (2018) or Ivanová (2019) for the analysis of non-epistemic use of *may be* in constructions of the type *I may be a woman, but I can change a tire*.

(11) *Bezpochyby máš pravdu, že je někdo na palubě, ale to nemusí nutně znamenat, že to má něco společného s námi...* (SYNv13)
 ‘You’re **no doubt** right that someone’s on board, but that doesn’t necessarily mean it’s anything to do with us...’

(12) *Možná že jsem vás nepřesvědčila, ale já věřím, že hlavní slovo v té záležitosti měl pan Dixon.* (InterCorp)
 ‘I **may** not have convinced you **perhaps**, but I am perfectly convinced myself that Mr. Dixon is a principal in the business.’

Analogous structures can as well be found in an inter-sentential setting, see (13). This is particularly interesting, since the original Czech construction *sice – ale*, as well as the German *zwar – aber*, is syntactically strongly constrained, it occurs within one complex sentence unit, otherwise it is considered ungrammatical. It seems that similar constructions with epistemic markers operate independently of sentence boundaries in this respect.

(13) [Context: *Když je někdo opravdu dobrý, tak si poradí.*] *Samořejmě i dnes máme fenomenální osobnosti. Na jednu takovou připadne však devět jiných, které nevhodné školní vzdělávání srazí do průměru.* (PDiT 4.0)
 ‘[Context: If someone is really good, they will find a way.] **Of course**, we still have phenomenal individuals today. For every one of them, **however**, there are nine others whom inadequate schooling pushes down to mediocrity.’

Moreover, a sentence-initial EM can become syntactically independent, separated by a comma, and its function may shift to that of an affirmative particle (14).

(14) *Jistě, skepticismus je někdy na místě. Ale že by dětem někdo podsouval vzpomínky na minulé životy pořád dokola?* (SYNv13)
 ‘**Sure**, scepticism is sometimes appropriate. **But** that someone would keep implanting memories of past lives into children over and over?’

The abovementioned patterns, be it the intra-sentential, inter-sentential, or the independent affirmative markers, share typical characteristics and conversational structure of a concessive relation. Among the typical features shared with concessive structure, the following are mentioned: adversativity and causal relation between the two parts of the sentence, resulting in an unmet expectation, triadic logical structure and polyphonic character (underlying dialogic interaction, the sentence presents opinions of at least two people, there are two “voices” present) (Schwenter 2000; Drobník 2024). Couper-Kuhlen and Thompson (2000, p. 382) work with the following conversational structure of the conceding act:

1st move A: States something or makes some point

2nd move B: Acknowledges the validity of this statement or point (the conceding move)

3rd move B: Goes on to claim the validity of a potentially contrasting statement or point

In our data, the first move is either present in the preceding context (actual dialogue between two subjects) (15), implicitly assumed on the basis of the speaker's experience with the partner or a third person (16), or it is represented as a generally accepted fact (17). The concessive construction then involves the explicitly contrasted 2nd and 3rd move, respectively.

(15) „Musím ji jít hledat.“ „**Jasně.** Ale zatím se posad' tuhle na lavici.“ (SYNv13)
“I have to go look for her.” “**Sure. But** in the meantime, take a seat on the bench here.””

(16) *Strážci jazyka se mnou možná nebudou souhlasit, ale to je v pořádku, nemusejí.* (SYNv13)
‘The guardians of the language **may** disagree with me, **but** that's okay, they don't have to.’

(17) *Z těch dvou možností by nepochybňě byla bezpečnější puška, ale čím víc Frank mysel na horkou krev, stříkající z proříznutého hrdla George T. Nelsona a skrápějící mu ruce, tím víc se mu zamlouvala i druhá možnost.* (SYNv13)
‘Of the two options, the rifle would **undoubtedly** have been the safer, **but** the more Frank thought about the hot blood spurting from George T. Nelson's slit throat and scraping his hands, the more he liked the second option.’

The EM then serves to emphasize the acknowledgement of the validity of the 1st move. In other words, “in the epistemic domain concessive connection expresses the idea that the speaker, in spite of being convinced of the content of the concessive clause, still reaches the opposite conclusion contained in the main clause” (Crevels 2000, p. 318). The same idea is expressed by Karlík (1995), who suggests that concessive sentences should be viewed pragmatically as a way to prevent misunderstanding in communication, to prevent conflict (cf. also Barth 2000, or Čermáková et al. 2019).

The epistemic strength of the EM then indicates the actual amount of credit we give to the 1st move statement, or the level of politeness we wish to grant to its author, cf. (18a) and (18b).

(18a) *Bezpochyby je pravda, že může nastolovat nějaká téma, může je debatovat s vládou, ale jeho pravomoci jsou někde jinde.* (SYNv13)
‘It is **undoubtedly** true that he can raise issues, he can debate them with the government, **but** his powers are elsewhere.’

(18b) *Možná máš pravdu, ale to ho nijak neomlouvá.* (SYNv13)
'You **may** be right, **but** that's no excuse.'

4 CONCLUSION

In this paper, we have examined several patterns of the connective use of Czech epistemic markers (EMs). The first pattern involves two EMs expressing differing degrees of certainty, which are juxtaposed to highlight a contrast in modality. We describe the role of EMs in this context as *text-structuring*. The second pattern concerns *topic orientation* – here, the EM functions as a topic shifter, marking the boundary between irrelevant and relevant utterances, or summarizing preceding, less relevant content with a more general and more relevant idea.

The third, and most prominent, pattern is the *concessive* use, in which the EM takes on the role of the Czech connective *sice* within the multiword connective expression *sice – ale*. Unlike the neutral *sice*, however, the EM also conveys politeness – expressing a greater or lesser degree of respectful acknowledgment of the interlocutor's opinion before presenting one's own, contrasting view.

While observing the connective functions of the EMs in corpora, we can see that the (secondary) pragmaticalization is an ongoing process. Nevertheless, it is important to try to capture the full range of their functions, in order to offer credible and relevant accounts of the epistemic markers in Czech.

ACKNOWLEDGEMENTS

The research has been supported by the Czech Science Foundation under the project GA23-05240S.

References

Barth, D. (2000). "that's true, although not really, but still": Expressing concession in spoken English. In Cause – Condition – Concession – Contrast: Cognitive and Discourse Perspectives. Berlin: De Gruyter Mouton, pp. 411–438.

Čermáková, A. et al. (2019). Diskurzní markery. In: J. Hoffmannová et al. (eds.): Syntax mluvené češtiny, pp. 244–351.

Couper-Kuhlen, E., and Thompson, S. (2000). Concessive patterns in conversation. In Cause – Condition – Concession – Contrast: Cognitive and Discourse Perspectives. Berlin: Mouton de Gruyter, pp. 381–410.

Crevels, M. (2000). Concessives on different semantic levels: A typological perspective. In Cause – Condition – Concession – Contrast: Cognitive and Discourse Perspectives. Berlin: Mouton de Gruyter, pp. 313–340.

Drobník, O. (2024). Přípustka v mluvnicích češtiny. Bohemistyka, 24(3), pp. 323–348.

Grepl, M., and Karlík, P. (1998). Skladba češtiny. Olomouc: Votobia, 603 p.

Fraser, B. (2009). An Account of Discourse Markers. *International Review of Pragmatics*, 1(2), pp. 293–320.

Ivanová, M. (2019). Epistemicita ako koncept deiktického odkazovania na propozíciu. In: J. Kesselová (ed.): Personálna a sociálna deixa v slovenčine. Prešov: Filozofická fakulta Prešovskej univerzity, pp. 153–210.

Karlík, P. (1995). Studie o českém souvětí, Brno: MU.

Kolářová, I. (1998). Významy slova tedy (teda) v souvislých textech. *Naše řeč*, 81(2–3), pp. 118–123.

Komárek, M., Kořenský, J., and Petr, J. (1986). Mluvnice češtiny 2. Praha: Academia.

König, E. (1988). Concessive connectives and concessive sentences: cross-linguistic regularities and pragmatic principles. In: J. Hawkins (ed.): *Explaining Language Universals*. Oxford: Blackwell, pp. 145–166.

Křen, M. et al. (2024). Corpus SYN, version 13 from 27/12/2024. ÚČNK FFUK, Praha. Accessible at: <https://www.korpus.cz>.

Leclercq, B. (2024). The post-modal grammaticalisation of concessive may and might. *Constructions and Frames*, 16, pp. 130–161.

Mírovský, J., Synková, P., Poláková, L., Kloudová, V., and Rysová, M. (2021). CzeDLex 1.0. Data/software, UK, Prague. Accessible at: <http://hdl.handle.net/11234/1-4595>.

Mladová, L. (2008). K problematice vztahu rematizátorů a textových konektorů. (On the Relation between Rhematizers and Discourse Connectives). In *Čeština doma a ve světě*, 3 and 4, pp. 126–133.

Palmer, F. R. (2001). Mood and Modality. Cambridge University Press.

Rosen, A., Vavřín, M., and Zasina, A. J. (2022). InterCorp, Release 15 of 11 November 2022. ÚČNK UK. Accessible at: <http://www.korpus.cz>.

Rossari, C. (2018). The representation of modal meaning of French sentence adverbs in a qualitative and quantitative approach. In *Linguistik online* 92(5), Special Issue, pp. 235–255.

Schwenter, S. (2000). Viewpoints and polysemy: Linking adversative and causal meanings of discourse markers. In *Cause – Condition – Concession – Contrast: Cognitive and Discourse Perspectives*. Berlin: Mouton de Gruyter, pp. 257–281.

Synková, P., Mírovský, J., Paclíková, M., Poláková, L., Rysová, M., Scheller, V., Zdeňková, J., Zikánová, Š., and Hajičová, E. (2024). Prague Discourse Treebank 4.0. Data/software, LINDAT/CLARIAH-CZ. Accessible at: <https://ufal.mff.cuni.cz/pdit4.0>.

Šindlerová, J., Štěpánková, B., and Andrén, I. L. (2023). Epistemická částice zřejmě pohledem paralelního korpusu. *Korpus – gramatika – axiologie*, 27, pp. 37–52.

Štěpánková, B., Poláková, L., Šindlerová, J., and Novák, M. (2024). What Can Dictionaries Tell Us About Pragmatic Markers – Building the Lexicon of Epistemic and Evidential Markers in Czech. In *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress*, Zagreb: Institut za hrvatski jezik, pp. 728–741.

Traugott, E. C. (1989). On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65(1), pp. 31–55.

Vavřín, M., and Rosen, A. (2015). Treq. FFUK. Praha. Accessible at: <http://treq.korpus.cz>.