

Expressing certainty and attitude in Czech and German

We present a corpus study investigating how different stance markers are translated from English to Czech and German. In the latter two languages, speakers may express certainty or attitude with modal particles.¹ We define these particles as noninflected sentence modifiers that add expressive meaning² to an utterance (Nekula, 2017; Zimmermann, 2011), see (1).

(1) Didn't Paul want to be here?

a. Pavel je *přece* nemocný.

b. Paul ist *doch* krank.

c. Paul is PART sick.

(slightly adapted from Nekula, 1996)

Between 1980 and 2010, there has been some research on the equivalence of (modal) particles in Czech and German. Researchers identify and discuss equivalent pairs of particles in both languages, sometimes based on corpora of translated literature (Birckmann, 1991; Masařic, 1982; Nekula, 1996; Rinas, 2006, 2010; Spěváková, 1992). While most of the existing literature focuses on the comparison of German *denn*, *doch*, and *ja* to possible Czech equivalents like *přece*, *vždyť*, *však*, and *copak*, other pairs have been studied less frequently. One of these combinations is the German particle *wohl* and its possible equivalents *snad* in exclamatives (Rinas, 2010) and *asi* in declaratives (Masařic, 1982).

Since previous corpus studies have been based on corpora of translated literature, we conduct a corpus study comparing the use of the modal particles *asi* and *snad* in Czech and *wohl* in German using the recent TED2020 corpus. TED2020 contains transcripts of TED and TED-X talks from July 2020 that have been translated by volunteers to many different languages³ and aligned on a sentence level. The parallel data are available through the open parallel corpus OPUS (Tiedemann, 2012).⁴ In total, the parallel corpus of Czech and German transcripts contains 154,626 sentence pairs (~4.49M token). We use a sub-corpus of 750 sentence pairs (250 sentence pairs per particle, 13,986 token).

This corpus allows us to study how speakers of different languages translate speaker certainty and attitude expressed in English into Czech and German. We use the aligned transcripts to contrast the different translations, comparing whether a sentence that is translated to Czech using *asi* or *snad* is translated to German using *wohl* or any other type of stance marker. In the same way, we compare sentences with *wohl* to the Czech translations.

Our results show that there is variety in the marker of certainty or attitude translators choose, ranging from the expected (or other) modal particles to adverbs, phrases, or no realization at all, see table 1 for an overview on parallel translations of *snad*.

This corpus study cannot be compared to a translation task that asks translators to choose the most literal translation, and therefore cannot account for whether native speakers would translate a Czech sentence containing *snad* with *wohl* in German or not. However, it shows the variety of stance markers used when translating a sentence from English, and that modal particles are used, even though English itself does not use modal particles for expressing certainty or attitude (cf. Grosz, 2022).

¹ Other frequently used terms are 'discourse particles' in English, 'Abtönungs-', 'Diskurs'- or 'Modalpartikeln' in German or 'částice modifikační' in Czech.

² In this example (a-b), the particle indicates that the speaker asking the question already knows Paul is sick.

³ <https://www.ted.com/participate/translate>

⁴ <https://opus.nlpl.eu/index.php>

	snad	wohl	vielleicht	other particles ⁵	other adverbs	phrases ⁶	other ⁷	none
declarative	204 (78.46)	7 (2.69)	65 (25.00)	8 (3.07)	42 (16.15)	15 (5.76)	10 (3.84)	57 (21.92)
question	56 (21.53)	0	2 (0.76)	16 (6.15)	4 (1.53)	1 (0.38)	6 (2.30)	27 (10.38)
total	260 (100)	7 (2.69)	67 (25.76)	24 (9.23)	46 (17.69)	16 (6.15)	16 (6.15)	84 (32.30)

Table 1: Translations chosen in German for sentences that were translated to Czech using *snad*. Absolute and relative count, divided by sentence type.

References

- Birckmann, H. (1991). *Die tschechischen Äquivalente zu den deutschen Partikeln denn, doch und ja*. [Magisterarbeit an der FU Berlin].
- Grosz, P. (2022). Scalarity as a meaning atom in wohl-type particles. In R. Gergel, I. Reich, & A. Speyer (Hrsg.), *Particles in German, English, and Beyond* (Vol. 224, pp. 243–268). John Benjamins Publishing Company. <https://doi.org/10.1075/slcs.224>
- Masařic, Z. (1982). Vorüberlegungen zu den sog. Satzadverbien im Deutschen und Tschechischen. *Brünner Beiträge zur Germanistik und Nordistik*, 3, 21–35.
- Nekula, M. (1996). *System der Partikeln im Deutschen und Tschechischen: Unter besonderer Berücksichtigung der Abtönungspartikeln*. Niemeyer.
- Nekula, M. (2017). MODIFIKAČNÍ ČÁSTICE. In P. Karlík, M. Nekula, & J. Pleskalová (Hrsg.), *CzechEncy—Nový encyklopedický slovník češtiny*. https://www.czechency.org/slovník/MODIFIKAČNÍ_ČÁSTICE
- Rinas, K. (2006). *Die Abtönungspartikeln doch und ja. Semantik, Idiomatisierung, Kombinationen, tschechische Äquivalente*. Peter Lang.
- Rinas, K. (2010). Partikeln als semantische Mumien: Zur Bedeutung von wohl, vielleicht und snad in Ausrufesätzen. *Acta Facultatis Philosophicae Universitatis Ostraviensis. Studia Germanistica*, 6, 113–128.
- Spěváková, D. (1992). Zu einigen Übersetzungsmöglichkeiten der deutschen Abtönungspartikeln ins Tschechische. *Brünner Beiträge zur Germanistik und Nordistik*, 8, 43–57.
- Tiedemann, J. (2012). Parallel Data, Tools and Interfaces in OPUS. *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012)*.
- Zimmermann, M. (2011). Discourse Particles. In P. Portner, C. Maienborn, & K. von Heusinger (Hrsg.), *Semantics* (Vol. 2, pp. 2011–2038). Mouton de Gruyter.

⁵ Including other modal particles, focus particles, and intensifying particles.

⁶ E.g., *denke ich* ('I think'), *ich hoffe* ('I hope'), *wenn nicht (sogar)* ('if not (even)'), etc.

⁷ E.g., using a different sentence mode, tag questions, etc.