

Hannah J. Seemann/Albert Marsik

Expressing the degree of confidence and attitude in Czech and German

Keywords: stance marking; modal particles; Czech; German; parallel corpora

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 their degree of confidence or their attitude with modal particles.¹ We define these particles as non-inflected sentence modifiers that add expressive meaning to an utterance (Zimmermann 2011; Nekula 2017), see (1). In this example, the particle indicates that in the opinion of the person answering in (1a-b), the speaker asking the question already knows (or should know) Paul is sick. Our definition of stance follows the definition given by Biber et al. (1999): Stance markers convey a speaker's feelings, attitudes, and judgments.

- (1) Didn't Paul want to be here? (slightly adapted from Nekula 1996)
- a. Pavel je *přece* nemocný.
 - b. Paul ist *doch* krank.
'Paul is PART sick.'

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 (Masařic 1982; Birckmann 1991; Spěváková 1992; Nekula 1996; Rinas 2006, 2010). 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 exclamative (Rinas 2010) and *asi* in declarative (Masařic 1982) sentences.

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 languages² and aligned on a sentence level. The parallel data are available through the open parallel corpus OPUS (Tiedemann 2012).³ In total, the parallel corpus of Czech and German transcripts contains 154,626 sentence pairs (~4.49M tokens). We use a sub-corpus of 600 sentence pairs (200 sentence pairs per particle, 13,486 tokens).

This corpus allows us to study how speakers of different languages translate a speaker's degree of confidence and attitude expressed in English into Czech and German. We use the aligned transcripts to contrast the different translations, comparing whether a sentence that

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

² <https://www.ted.com/participate/translate>

³ <https://opus.nlpl.eu/index.php>

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 confidence 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*.

	snad	wohl	vielleicht	other particles⁴	other adverbs	phrases⁵	other⁶	none
declarative	163 (81.5)	7 (3.5)	48 (24.0)	7 (3.5)	37 (18.5)	13 (6.5)	17 (8.5)	34 (17.0)
question	37 (18.5)	0	1 (0.5)	11 (5.5)	3 (1.5)	1 (0.5)	3 (1.5)	18 (9.0)
total	200 (100)	7 (3.5)	49 (24.5)	18 (9.0)	40 (20.0)	14 (7.0)	20 (10.0)	52 (26.0)

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

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 confidence or attitude (cf. Grosz 2022).

References

- Biber, Douglas/Johansson, Stig/Leech, Geoffrey/Conrad, Susan/Finegan, Edward (1999): The Longman grammar of spoken and written English. London: Longman.
- Birckmann, Holger (1991): Die tschechischen Äquivalente zu den deutschen Partikeln denn, doch und ja.
- Grosz, Patrick (2022): Scalarity as a meaning atom in wohl-type particles. In: Gergel, Remus/Reich, Ingo/Speyer, Augustin (eds.): Particles in German, English, and Beyond. Amsterdam: John Benjamins Publishing Company (Studies in Language Companion Series), pp. 243–268. <http://doi.org/10.1075/slcs.224>.
- Masařic, Zdeněk (1982): Vorüberlegungen zu den sog. Satzadverbien im Deutschen und Tschechischen. In: Brünner Beiträge zur Germanistik und Nordistik, 3, pp. 21–35.
- Nekula, Marek (1996): System der Partikeln im Deutschen und Tschechischen: unter besonderer Berücksichtigung der Abtönungspartikeln. Tübingen: Niemeyer (Linguistische Arbeiten, 355).
- Nekula, Marek (2017): MODIFIKAČNÍ ČÁSTICE. CzechEncy - Nový encyklopedický slovník češtiny. Edited by Petr Karlík/Marek Nekula/Jana Pleskalová. https://www.czechency.org/slovník/MODIFIKAČNÍ_ČÁSTICE (last access: 24 January 2023).
- Rinas, Karsten (2006): Abtönungspartikeln, das Mittelfeld und die Wackernagelposition. In: brücken: Germanistisches Jahrbuch Tschechien - Slowakei, pp. 331–339.

⁴ Including other modal particles, focus particles, and intensifying particles.

⁵ E.g., *denke ich* ('I think'), *ich hoffe* ('I hope'), *wenn nicht sogar* ('if not (even)'), etc.

⁶ E.g., using a different sentence mode, tag questions, etc.

ICLC-10 2023 - Document Template for Revised Abstracts

- Rinas, Karsten (2010): Partikeln als semantische Mumien: Zur Bedeutung von wohl, vielleicht und snad in Ausrufesätzen. In: Acta Facultatis Philosophicae Universitatis Ostraviensis. *Studia Germanistica*, 6, pp. 113–128.
- Spěváková, Dana (1992): Zu einigen Übersetzungsmöglichkeiten der deutschen Abtönungspartikeln ins Tschechische. In: *Brünner Beiträge zur Germanistik und Nordistik*, 8, pp. 43–57.
- Tiedemann, Joerg (2012): Parallel Data, Tools and Interfaces in OPUS. In: *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012)*.
- Zimmermann, Malte (2011): Discourse Particles. In: Portner, P./Maienborn, C./Heusinger, K. von (eds.): *Semantics*. Berlin: Mouton de Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft HSK, 33), pp. 2011–2038.

Contact information

Hannah J. Seemann

Ruhr-Universität Bochum

hannah.seemann@rub.de

Albert Marsik

Univerzita Karlova

marsikal@ff.cuni.cz