

Yulia Clausen | Tatjana Scheffler

# A corpus-based analysis of meaning variations in German tag questions

## Evidence from spoken and written conversational corpora

Suggested citation referring to the original publication:  
Corpus Linguistics and Linguistic Theory (2020)  
DOI <https://doi.org/10.1515/cllt-2019-0060>  
ISSN (print) 1613-7027  
ISSN (online) 1613-7035

Postprint archived at the Institutional Repository of the Potsdam University in:  
Postprints der Universität Potsdam  
Humanwissenschaftliche Reihe ; 706  
ISSN 1866-8364  
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-467882>  
DOI <https://doi.org/10.25932/publishup-46788>



Yulia Clausen\* and Tatjana Scheffler

# A corpus-based analysis of meaning variations in German tag questions

## Evidence from spoken and written conversational corpora

<https://doi.org/10.1515/cllt-2019-0060>

**Abstract:** This paper addresses semantic/pragmatic variability of tag questions in German and makes three main contributions. First, we document the prevalence and variety of question tags in German across three different types of conversational corpora. Second, by annotating question tags according to their syntactic and semantic context, discourse function, and pragmatic effect, we demonstrate the existing overlap and differences between the individual tag variants. Finally, we distinguish several groups of question tags by identifying the factors that influence the speakers' choices of tags in the conversational context, such as clause type, function, speaker/hearer knowledge, as well as conversation type and medium. These factors provide the limits of variability by constraining certain question tags in German against occurring in specific contexts or with individual functions.

**Keywords:** German, tag questions, discourse functions, pragmatic variability, corpus annotation

## 1 Introduction

Tag questions (TQs) are a crosslinguistically attested phenomenon typical of spoken conversations. We adopt the definition from König (2017: 236) and define a question tag as a formulaically fixed expression that attaches to a reference utterance (the *anchor*) and can, under certain conditions, elicit a knowledge- or comprehension-related reaction from the hearer.<sup>1</sup>

---

<sup>1</sup> In this paper, we use “tag question” to refer to the entire utterance and “(question) tag” to refer to only the tagged-on fixed word or phrase attached to the anchor.

---

**\*Corresponding author: Yulia Clausen**, Department of Linguistics, University of Potsdam, Potsdam, Germany, E-mail: [yulia.clausen@uni-potsdam.de](mailto:yulia.clausen@uni-potsdam.de)

**Tatjana Scheffler**, Department of Linguistics, University of Potsdam, Potsdam, Germany, E-mail: [tatjana.scheffler@uni-potsdam.de](mailto:tatjana.scheffler@uni-potsdam.de)

<https://orcid.org/0000-0001-7498-6202>

German exhibits a large inventory of possible tags, which are always fixed in form and stem from various literal meanings, such as truthfulness (e. g. *nicht wahr*), alternatives (e. g. *oder*) or confirmation (e. g. *ja, richtig*)<sup>2</sup>:

- (1) – *nich(t)/nech/näch/net/nitt*
- *ne/nä/na/nö*
- *nich(t) wahr/nu wor*
- *wa*
- *stimmt's*
- *gell/gelle/gel/gelt/ge/gö/gä/gäu*
- *woll*
- *wonnich*
- *he*
- *ja/nu*
- *was/wat*
- *nor ni*
- *richtig*
- *oder*
- *oder (etwa) nich(t)/etwa nich(t)/oder was/oder (etwa) doch*

These tags can often be used interchangeably:

- (2) [Emma and her family returned from a vacation. It is Monday morning, her mother comes to the kitchen and finds Emma watching movies instead of getting ready for school.] Mother:

*Du musst heute nicht in die Schule, ne/gell/oder?*

You must today not in the school, TAG?

'You don't have to go to school today, do you?'

On the other hand, a survey reported in (Elspass and Möller 2011) indicates that the choice of tags can depend on the properties encoded in the anchor of a TQ. For example, in contexts where the speaker is unsure, such as in (3), *oder* is most common, although other tags are also possible. In situations where information expressed in the anchor can be easily verified, as in (4), *gell* and *ne* are the most frequent variants across Germany.

---

<sup>2</sup> We compiled this list based on the resources on German TQs referenced in this paper and information from native speakers. However, we do not exclude the existence of other tag variants.

- (3) *Er wohnt doch bei seiner Tante, ...?*  
 He lives PART at his aunt, ...?  
 ‘He lives at his aunt’s, ...?’ (Elspass and Möller 2011)
- (4) *Das ist ein herrliches Wetter heute, ...?*  
 It is a wonderful weather today, ...?  
 ‘The weather is wonderful today, ...?’ (Elspass and Möller 2011)

In fact, individual tag variants in German show a specialization for certain contexts and cannot always be easily switched for each other. For example, when a speaker brings up a fact from their own experience, it is possible to use *ne*, whereas *oder* would be infelicitous (5).<sup>3</sup> This is parallel to example (4), where *oder* was very rarely reported.

- (5) [Sophie is walking her new dog when she runs into Tim. She is expecting that he would congratulate her on the new dog, but he doesn’t mention it. She isn’t sure whether he actually realizes that she got a new dog.] Sophie:  
*Ich habe einen neuen Hund, ne/\*oder?*  
 I have a new dog, TAG?  
 ‘I have a new dog, eh?’ (cf. Wiltschko and Heim 2016: 27)

Examples like (2) suggest that tags in German can all be used to ask for confirmation on the truth of a proposition. However, the examples also demonstrate that certain contexts are incompatible with or favor specific tag variants.

Studies on German have so far only addressed individual tags, leaving aside the variability offered by the vast range of tag variants present in the language. In this paper, we carry out a systematic corpus study of German TQs in different corpora. We address three main questions: First, as a phenomenon typical of conversations, can we find TQs in the spoken as well as the written medium, and does the medium have an effect on the distribution of tag variants shown in (1)? Second, can we identify which context properties influence the choice of question tags?<sup>4</sup> And third, which tag variants are interchangeable in which contexts, and what are their individual syntactic and pragmatic constraints? By

<sup>3</sup> We adapted this example to German from (Wiltschko and Heim 2016). We are aware of a possibility of certain situations, such as a speaker’s sarcastic or snippy attitude, where using *oder* would be conceivable. However, the described situational context aims to represent a straightforward hint by the speaker to the fact of a new dog in order to start a conversation. This meaning, we propose, cannot be well communicated using this particular tag variant.

<sup>4</sup> We understand the context of an utterance in the sense of Birner (2013: 29): linguistic (what has been uttered before) or extralinguistic (interlocutors, their epistemic beliefs etc.).

answering these questions, we systematize the empirical basis for a formal analysis of tags in German and lay groundwork for a crosslinguistic comparison of this phenomenon.

## 2 Previous research on tag questions

TQs are discourse markers that are often used to elicit a kind of confirmation from the hearer (but see further discussion in this section). Most work has focused on English variant and invariant tags (i.a. Cattell 1973; Ladd 1981; Holmes 1984; Tottie and Hoffmann 2006; Kimps 2007; Kim and Ann 2008; Moore and Podesva 2009; Dehé and Braun 2013; Takahashi 2014; Malamud and Stephenson 2015; Pichler 2016). In this section, we look at existing research on TQs in order to identify relevant context distinctions that may help us differentiate the types of the German tags under consideration.

For English, a study by Cattell (1973) notes a difference in the scope of negation in reverse and same polarity TQs and argues that these two types of TQs differ in whether the anchor clause proposition is the *stance of the speaker* (reverse polarity) or not (same polarity). He claims that in reverse polarity TQs, the speaker is to some extent committed to the anchor clause proposition, whereas this is typically not the case in same polarity TQs (Cattell 1973: 5–6).

Drawing on this negation scope difference, Ladd (1981) comes to distinguish two types of tags based on their prosody. He argues that one type expresses “*true uncertainty or doubt*” and asks for confirmation, while the other is a hedged assertion, which, in the absence of speaker doubt, merely asks for acknowledgement from the hearer. Lack of speaker knowledge is also identified in two of the six tag functions by Tottie and Hoffmann (2006: 300): true information questions and biased confirmation questions.

In contrast to reverse polarity TQs, which often ask for affirmation or confirmation of the anchor proposition that the speaker is unsure about, same polarity tags usually indicate that the hearer has authority over the situation or is seen as the potential source of the expressed proposition (Cattell 1973; Farkas and Roelofsen 2012; Malamud and Stephenson 2015). This kind of TQs may for example be used to express the speaker’s surprise by repeating something that was just said (Huddleston and Pullum 2008: 895). Kimps (2007: 281) discusses the functions of same polarity tags in detail proposing two basic functions (evidential modification and turn allocation) as well as secondary “attitudinal” uses.

## 2.1 Functions of invariant tags

In most languages, question tags are invariant in form, including German. While the crosslinguistic picture of the uses of these tags is still not complete, certain recurring functions and contexts for their use can be identified.

English offers a range of invariant tags such as *innit*, *right*, *eh* and others, in addition to the canonical ones. Andersen (2001) studies *innit* and classifies its uses according to three main categories: speaker's attitude (whether the speaker *S* is certain or uncertain about the anchor), *S*'s evaluation of the hearer's relation to the anchor (e. g. whether the hearer is expected to know about or share the anchor proposition), and whether the anchor constitutes an opinion or a fact.

Columbus (2010a,b) considers many different invariant tags in three varieties of English, and comes to distinguish 17 functions. She concludes that most instances of tags in her corpora have one of four functions: narrative, confirmation checking, following a fact or opinion statement, and emphatic. The main differences are that narrative and emphatic tags do not necessarily elicit a response from the hearer (the narrative function has been called “non-turn-yielding” by Andersen (2001)), whereas confirmation checks are clearly meant to indicate uncertainty of the speaker and a need for a verbal response.

Speaker and hearer attitudes toward the anchor proposition emerge as recurring themes in the classification of tag variants, as well as the question whether a response is required. In this vein, Gómez González (2018) distinguishes speaker centered from addressee centered and exchange centered tags. In her study of British English tags, she finds that information and confirmation seeking uses are most frequent overall, but uses where the speaker is certain and does not seek interaction from the hearer are also frequent for invariant tags.

Wilson et al. (2017) study invariant (and variant) tags in Trinidadian English in different types of conversational data. They find that invariant tags are mostly speaker centered and fulfill mainly facilitative (speaker is certain of the truth of the anchor proposition, but encourages a response) and punctuational (speaker is sure and neither is seeking information nor expecting an answer) functions, often being multifunctional. They also report correspondence between tag function and text type, where, for example, phone calls contain more confirmatory tags (speaker is unsure of the truth of their utterance or wants the interlocutor to endorse what they had said) than face-to-face conversations.

Finally, in a recent paper on the Canadian English tag *eh*, Wiltschko et al. (2018) consider the syntactic context (in particular, the anchor clause type) as central for determining the function of a TQ in discourse.

In Japanese, the sentence-final tags *deshoo* and *janai ka* are differentiated by whether they require a response from the hearer: while *deshoo* asks for

confirmation of the truth of a proposition or of a judgement, uses of *janai ka* merely require recognition or acknowledgement of a fact, and therefore hearer reaction is not necessary (Asano 2008). Mithun (2012) finds that the question tag *mahi* in Mohawk has both epistemic uses (indicating reduced commitment of the speaker to the anchor) and interactive uses (i. e. requiring a response), based on the functional inventory from (Tottie and Hoffmann 2006). Narrative uses where the tag is merely used to ensure continued attention of the hearer are frequent in her data. Similarly, Tomaselli and Gatt (2015) specify seven functions for Italian tags, including confirmation requests, establishing common ground, or checking understanding, but also discourse structuring functions such as emphasizing or closing a topic.

For Cypriot Greek, Erotokritou (2014) distinguishes two invariant tags: *oksa* is used when the speaker has doubt or is unsure and *en en* is used when the speaker is certain. The speaker vs. hearer perspective is echoed in the discussion of the Northern Mandarin sentence-final particle *ha*, which has three uses: confirming the speaker's stance, confirming the hearer's stance, and a non-confirmational narrative use (Yang and Wiltschko 2016).

The studies of invariant tags across languages show that they are multifunctional discourse markers, which are influenced by their context and the text type they occur in (Aijmer 2013; Wilson et al. 2017; Wiltschko et al. 2018). However, we can also discern recurring themes which guide our study of the German tags: First, the *epistemic attitude of the speaker* to the anchor proposition (is the speaker uncertain, or are they reporting on their own attitude or stance). Second, the *clause type* and *speech act* of the anchor (e. g. Yang and Wiltschko 2016; Wiltschko et al. 2018). Third, it is important to distinguish instances of TQs where the *target of confirmation* is the truth of the anchor proposition from narrative or discourse structuring cases, where ensuring continued attention is the goal of the tag use. In the following sections, we look at these distinctions with respect to our German data.

## 2.2 Tag questions in German

Corpus-based studies on German TQs have mostly focused on individual tags: *oder* (Drake 2016), *ne* (Hagemann 2009; König 2017), and *gell* (Heim 2019). Many of the described functions are found to overlap to a certain degree between these tags.

A theoretical attempt at comparing *ne*, *oder*, and *gell* by (Zifonun et al. 1997: 384–385) proposes functions relevant for the corpus studies we address in this section. The authors distinguish the uses of *oder* and *ne*, while *ne* and



*gell* are attributed similar functions. They propose that *ne* strives to draw the hearer's awareness to the anchor proposition, and preferably elicit a confirmational response, while *gell* marks the proposition as (future) shared knowledge. *Oder* is described as associated with turn-taking and used in cases where a speaker is less sure and considers the possibility of an alternative to the anchor proposition.

Drake (2016) tests the theoretical assumptions in (Zifonun et al. 1997) about turn-final *oder* in a corpus of telephone conversations. Based on prosodic features, she determines that *oder* can either elicit a confirmation or a disconfirmation as a response or license an alternative to what is expressed in the anchor. Drake (2016: 189) also documents certain differences in the syntactic type of the anchor clause with respect to the different uses of *oder*.

Hagemann (2009) distinguishes different types of tags based on their prosody and position in a collection of recordings of casual conversations and university seminars. König (2017) studies *ne*'s intersubjective functions and its role in discourse structuring based on telephone and Skype conversations. Hagemann (2009) claims that *turn-final* tags elicit a verbal response, whereas *turn-internal* tags require the hearer to acknowledge the truth of the anchor proposition without requesting a "true reaction".

For utterance-final *ne*, König (2017) defines two possible functions based on the speaker and hearer epistemic status: asking for confirmational response (higher status of the hearer) and requesting to document understanding/continuing attention (higher status of the speaker). In her data, *ne*-TQs are more often confirmed than rejected, which suggests that despite asking for confirmation, a speaker is generally sure of the expressed proposition.<sup>5</sup> Following Hagemann (2009), she further states that *ne* in narrative contexts does not assume a reply, as the speaker usually continues with the turn without pausing.

Heim (2019: 130) contrasts *gell* with other tags, such as *oder* and *wa*. On the basis of selected corpus examples and a forced-choice response experiment, he compares turn-medial and turn-peripheral uses of *gell* and finds that while the former are used to check for mutual beliefs in the Common Ground (narrative *gell*), the latter ask for confirmation. For the confirmational *gell* he distinguishes two pragmatic functions similar to those determined by König (2017) for *ne*. Based on the epistemic status of the interlocutors, *gell* can either request confirmation about the truth of the expressed proposition (higher status of the hearer) or demand a confirmation regarding the speech act (higher status of the speaker).

These studies show that the tags behave similarly to each other, and the distinctions between their functions are of a subtle nature. This supports

---

<sup>5</sup> In this respect, König (2017: 241) mentions a contrast to *oder*, but without going into detail.

our hypothesis that all German tags share a common function of requesting confirmation. Different functions of turn-final tags are based on categories such as the epistemic status of the interlocutors and speaker certainty. Several of the studies also indicate a possible central difference between the tags *ne/gell* and *oder*. Even though prosodic distinctions are sometimes mentioned as relevant, they do not correspond directly to usage differences in German (cf. Hagemann 2009: 150–151; König 2017: 240). For this reason, and since prosody is unavailable for written text, they will not be a part of our study.

### 3 Data and method

The phenomenon of TQs is particularly characteristic of conversations. In this study, we investigate the discourse variability of German TQs in three conversational corpora: private telephone conversations in the CallHome German corpus (Karins et al. 1997), in-person dialogs between strangers in the GECO corpus (Schweitzer and Lewandowski 2013), and a corpus of Twitter messages (Scheffler 2014) that captures informal language in written form. We include Twitter because it has been shown to be highly interactive with frequent use of filled pauses, backchannel signals, and TQs (Rehbein et al. 2016).

Given that TQs are an indispensable part of conversational speech, we assume that they can be found in all of its types. However, due to specific demands of particular media, the functions of discourse markers may change “to meet the special demands of this medium” (Aijmer 2013: 51). We therefore expect quantitative as well as functional differences in tag usage across different types of conversations. For example, due to the absence of visual contact, turn-taking becomes more challenging in telephone speech than in face-to-face dialogs (cf. Aijmer 2013: 55). This might result in a more extensive use of tags that check for continuing attention or signal turn-taking in the former. In Twitter, we expect turn-taking to be less central and that tags might rather be used to ask for information or prompt for replies.

#### 3.1 Corpora

##### 3.1.1 CallHome

CallHome (CH) is a corpus of telephone speech collected and distributed by the Linguistic Data Consortium. It contains transcripts of 10 contiguous minutes,

each from 100 telephone conversations ranging up to 30 min in duration. The interlocutors are native speakers of German (who have been living in North America for a short time) and their family members or close friends overseas, typically in Germany.<sup>6</sup> The manually transcribed conversations include spontaneous discussions on various topics and cover standard German as well as Northern, Swabian, Hessian, and Bavarian accents. The corpus comprises ca. 235,000 word tokens.<sup>7</sup>

### 3.1.2 GECO

The GECO German Conversations Database (GC)<sup>8</sup> consists of 46 dialogs of about 25 min length each between previously unacquainted females between 20 and 30 years of age, mostly students. 22 dialogs took place in a unimodal setting, where participants could not see each other, while the remaining 24 dialogs were recorded in a multimodal setting with interlocutors facing each other. The conversation topics were chosen freely by the participants. The data cover standard German with some dialectal variation. The conversations were manually transcribed by the authors of the corpus. The corpus counts ca. 262,000 word tokens.

### 3.1.3 Twitter

The Twitter corpus (TW) provides a nearly exhaustive snapshot of German-language Twitter messages collected during April, 2013. It contains over 24 million tweets and includes a large majority of all German tweets sent during the specified time period. Therefore, it can be regarded as a representative sample of German Twitter content of that time. We used the entire corpus for this study, though only a subset of the tweets are part of conversations (= reply-chains). The tweets for this corpus were retrieved using a list of German high-frequency terms and filtered by applying a language identification module. See Scheffler (2014) for more details on the corpus construction and characteristics. The size of the corpus is ca. 248 million word tokens.

---

<sup>6</sup> Sometimes, the calls were placed outside Germany or, in rare cases, outside Europe.

<sup>7</sup> The number of word tokens for each corpus was counted automatically with the SoMaJo tokenizer: <https://github.com/tsproisl/SoMaJo>

<sup>8</sup> Access to GECO can be requested at <https://www.ims.uni-stuttgart.de/en/research/resources/corpora/ims-geco/>

### 3.2 Tag question extraction

From the three corpora, we automatically extracted all turns containing one or more TQ candidates based on the listing in (1). In the speech data, informal or dialectally colored pronunciations had been replaced with “standard German variants” (GECO) and “standard orthography” (CallHome) in the transcripts. This may also have affected the spelling of tags, so that variants such as *nä/nö* might have been transcribed as *ne* and *net/nitt* as *nicht*. The Twitter data are likely to contain various additional non-standard spellings of the tags (e. g. *geeell*, *nee*) that we do not cover in this study. In addition to the forms listed in (1), we only include *stimmts* in our search, since we consider the spelling without an apostrophe specifically relevant for Twitter for convenience and typing speed.<sup>9</sup> In general, we are positive that our list accounts for most of the existing German tag variants, and since some alternative spellings would result in a very high number of false positives (e. g. *nee* meaning ‘no’), we restrict our search to the listed options.

In this study, we focus on sentence-final tags that occur both turn-finally and turn-internally. A turn is a speaker’s contribution to the dialog, whose boundaries are marked by a speaker change.<sup>10</sup> Our analysis does not cover sentence-internal tags such as *gell* in (6). These tags differ from the ones we are interested in here, as they “do not engage the hearer about the propositional content” (Heim 2019: 132). They appear in narrative sequences where they have a discourse-structuring function (cf. König 2017: 248–250). Further, we do not include stand-alone tags such as *oder nicht* in (7), as we assume that they behave differently from the tags attached to an anchor.

- (6) ...*das ist kein Dörfchen, gell, da tobt der Bär, das ist voll ne*  
 ...this is no village, TAG, there rages the bear, this is really a  
*große Stadt und so :-D*  
 big city and such :-D  
 ‘...this is not a village, eh, there is always something happening here, it’s  
 totally a big city :-D’<sup>11</sup> (TW)

<sup>9</sup> Indeed, in our data, we found that *stimmts* occurred ten times more often than *stimmt*’s in Twitter.

<sup>10</sup> We use the terms speaker/hearer in a generic sense for the sake of simplicity, corresponding to author/addressee in Twitter. A turn in Twitter is equivalent to one tweet.

<sup>11</sup> All translations of corpus examples are our own. The proposed English tag equivalents represent one of many possibilities.

- (7) ...und die Sprache war wohl...Chinesisch. Oder nicht?  
 ...and the language was probably...Chinese. TAG?  
 ‘...and the language was probably...Chinese. Or not?’ (TW)

We extracted the TQ candidates using a regular expression that matches a tag followed by at least one question mark or a full stop, or any combination of those with an exclamation mark.<sup>12</sup> In German, many of the words used as tags can appear at the end of sentences while having other meanings. For instance, this is a very common case for *nicht* (‘not’):

- (8) Oh, du kennst den Inhalt meiner Dis nicht?  
 Oh, you know the content my dissertation not?  
 ‘Oh, you don’t know the content of my dissertation?’ (CH)

To count true TQs, we manually disambiguated the TQ candidates by applying the definition introduced in Section 1. For the spoken data (CallHome and GECO), we checked each extracted TQ candidate for the presence of a tag. For the Twitter data, which has a very large number of extracted turns, we randomly selected a subset of TQ candidates, which we disambiguated until at least 1,000 TQs were identified. We used the obtained information to estimate the number of TQs by extrapolating the counts in the sample to the size of the corpus. An overview of the TQ counts per corpus is presented in Table 1.

**Table 1:** Counts of TQs per corpus. From left to right: number of automatically extracted TQ candidates, number of manually checked TQ candidates, number of true TQs in the sample, number of true TQs per corpus (estimated for Twitter), number of TQs per 10,000 words.

Corpus	TQ candidates extracted	TQ candidates checked	True TQs (%)	(Estimated) TQs/corpus	Normalized TQs / 10,000 words
CallHome	3,104	3,104	1,509 (49)	1,509	64.21
GECO	2,171	2,171	588 (27)	588	22.44
Twitter	12,320	3,854	1,024 (27)	3,326	0.041

The counts for each of the tags found in at least one of the corpora are given in Table 2. The data for several of the variants are quite sparse. In our analysis, we therefore focus on the five most frequent tags across the three corpora: *oder* (31%), *ne* (29%), *ja* (12%), *nicht* (8%), and *gell* (7%).

<sup>12</sup> In the spoken corpora, punctuation is included in the transcriptions.

**Table 2:** Distribution of question tags per corpus. The counts for *stimmt's* also include *stimmts*. The normalized counts are given per 10,000 words.

Tag	CallHome			GECO			Twitter		
	#	%	norm.	#	%	norm.	#	%	norm.
<i>ge</i>	0	0	0	0	0	0	3	0.3	0.0001
<i>gel</i>	0	0	0	2	0.3	0.08	3	0.3	0.0001
<i>gell</i>	128	8.5	5.45	63	10.7	2.40	18	1.8	0.0007
<i>gelle</i>	4	0.3	0.17	0	0	0	3	0.3	0.0001
<i>gö</i>	0	0	0	0	0	0	1	0.1	0.00004
<i>he</i>	0	0	0	0	0	0	2	0.2	0.00008
<i>ja</i>	240	15.9	10.21	88	15.0	3.36	60	5.9	0.0024
<i>nä</i>	0	0	0	0	0	0	3	0.3	0.0001
<i>na</i>	12	0.8	0.51	1	0.2	0.04	0	0	0
<i>ne</i>	709	47.0	30.17	120	20.4	4.58	82	8.0	0.0033
<i>nech</i>	0	0	0	0	0	0	3	0.3	0.0001
<i>nich</i>	0	0	0	0	0	0	1	0.1	0.00004
<i>nicht</i>	244	16.2	10.38	1	0.2	0.04	16	1.6	0.0006
<i>nicht wahr</i>	8	0.5	0.34	0	0	0	12	1.2	0.0004
<i>nö</i>	0	0	0	0	0	0	1	0.1	0.00004
<i>oder</i>	93	6.2	3.96	281	47.8	10.73	602	58.8	0.0242
<i>oder etwa nicht</i>	0	0	0	0	0	0	1	0.1	0.00004
<i>oder nich</i>	0	0	0	0	0	0	1	0.1	0.00004
<i>oder nicht</i>	4	0.3	0.17	2	0.3	0.08	30	2.9	0.0012
<i>oder was</i>	42	2.8	1.79	9	1.5	0.34	67	6.5	0.0027
<i>oder wat</i>	0	0	0	0	0	0	2	0.2	0.00008
<i>oder wie</i>	18	1.2	0.77	21	3.6	0.80	34	3.3	0.0013
<i>richtig</i>	1	0.1	0.04	0	0	0	5	0.5	0.0002
<i>stimmt's</i>	0	0	0	0	0	0	15	1.5	0.0006
<i>wa</i>	0	0	0	0	0	0	27	2.6	0.0010
<i>was</i>	6	0.4	0.26	0	0	0	31	3.0	0.0012
<i>woll</i>	0	0	0	0	0	0	1	0.1	0.00004

## 4 Annotation

In this section, we describe our annotation process.<sup>13</sup> In CallHome and GECO, all occurrences of the five tags were annotated. In Twitter, we annotated the occurrences in the sample of 1,024 TQs described above.

<sup>13</sup> The authors of this paper created an annotation manual available at [https://github.com/TScheffler/TagQuestions/raw/master/Annotation\\_guidelines\\_TQs.pdf](https://github.com/TScheffler/TagQuestions/raw/master/Annotation_guidelines_TQs.pdf). The annotation was performed by a native speaker of German with linguistic background. During the annotation phase, we conducted regular meetings with the annotator to discuss and resolve any difficult cases. If any new cases of frequently occurring patterns emerged, we updated the manual accordingly.

We investigate the usage patterns of these tags based on syntactic as well as pragmatic information. This data is encoded in our annotation categories, which are informed by the findings of previous works, described in Section 2. We assume that the categories that distinguish between tag uses have an influence on the choice of particular tag variants.

For each turn containing a TQ, we annotate the following: anchor clause type, anchor speech act, speaker certainty, target of confirmation, and position in turn. The categories anchor clause type and anchor speech act are determined based on the anchor clause without taking into account the presence of a tag.<sup>14</sup> The other categories are determined by considering the complete sentence including the tag. The annotator also had access to preceding and following sentences in cases where this could help disambiguate.

## 4.1 Anchor clause type

The anchor clause type captures the syntactic form of the anchor clause. We explore the impact of this category on the choice of tag variants based on previous evidence that particular types of TQs can be incompatible with certain clause types. For example, Heim (2019: 134) found that interrogative clauses are compatible only with TQs requesting confirmation, whereas imperatives are used with demanding TQs instead.

We consider four clause types that represent common sentence types: *declarative* (9), *imperative* (10), *interrogative* (11), and *exclamative* (12). Further, we distinguish *ambiguous* (13) clauses, whose syntactic form conforms to both declarative and interrogative clause types with topic drop (cf. Drake 2016: 174).

### (9) DECLARATIVE

*Die gibt es aber in Deutschland auch, ja?*  
 they exist EXPL but in Germany too, TAG?  
 ‘But they do also exist in Germany, right?’ (CH)

<sup>14</sup> For this part of the annotation, our annotator was instructed to disregard the tag and treat the utterance as if it contained no tag.

## (10) IMPERATIVE

*Dann drückt mal die Daumen, dass ich uns die 3000 Öcken*  
 then press once the thumbs, that I us the 3000 bucks  
*rausklage, ne?*

sue for, TAG?

Then keep your fingers crossed that I successfully sue for the 3000 bucks,  
 ok? (TW)

## (11) INTERROGATIVE

*Ja, ist die Hand noch geschwollen oder?*

yes, is the hand still swollen TAG?

‘Yes, the hand is still swollen, is it?’

(CH)

## (12) EXCLAMATIVE

@user *Was ich nicht alles kann, gell? ;-)*

@user what I not all can, TAG? ;-)

@user What I am able to do, huh? ;-)

(TW)

## (13) AMBIGUOUS

*Hat er von dir geerbt dann, ne?*

has he from you inherited then, TAG?

‘(Has) he inherited from you then, right?’

(CH)

Finally, we found a number of incomplete sentence *fragments* as anchors (14). These TQs do not provide enough information to determine the syntactic type of the clause fragment or decide on the other semantic/pragmatic categories we are interested in. They should be addressed separately and are therefore not part of our further analysis in this study.

## (14) FRAGMENT

*Dieses Video von DreamingWeaponStudios, ne? Ich mag das voll.*

this video from DreamingWeaponStudios, TAG? I like this full.

‘This video from DreamingWeaponStudios, eh? I like it a lot.’

(TW)



## 4.2 Anchor speech act

There is no universally agreed upon inventory of possible speech acts (Levinson 2017). For our work, we consider the central classes of speech acts mentioned by Searle (1975): *assertion* (9); different types of directives: *directive* (10) (= all flavors of requests), *query* (11), *wish* (15), *exhortation* (16) (= call for joint action), *commissive* (17) (= self commitment), as well as *exclamation* (18), and *topic introduction* (19). The latter category can refer to cases of not previously mentioned information being introduced into the conversation as well as to the reactivation of knowledge from the Common Ground. This speech act is different from assertions and can ensure the interlocutor's attention for the upcoming information, as for example reported by König (2017: 250).

### (15) WISH

@user *Haha, wenn es doch nur so wäre, nicht? :’D*  
 @user *haha, if it PRT only so was, TAG? :’D*  
 ‘@user *Haha, if only that was true, huh? :’D* (TW)

### (16) EXHORTATION

@user *Dann lass uns mal gucken wegen Samstag ja?*  
 @user *Then let us once see about Saturday TAG?*  
 ‘@user *Then let’s see about Saturday, okay?* (TW)

### (17) COMMISSIVE

@user *ich schreib bei facebook ja?*  
 @user *I write on facebook TAG?*  
 ‘@user *I’ll write on facebook, okay?* (TW)

### (18) EXCLAMATION

*Was für Cooler Papa oder? Siehe Auflösung: <url>*  
*what for cool dad TAG? See resolution: <url>*  
 ‘*What a cool dad, right? See resolution: <url>* (TW)

### (19) TOPIC INTRODUCTION

*Das Gespräch mit dem Urs grad, ne? :’D*  
*The conversation with the Urs now, TAG? :’D*  
 ‘*The conversation with Urs just now, huh? :’D* (TW)

### 4.3 Speaker certainty

Speaker certainty reflects the notion of the epistemic status of the interlocutors. It encodes whether a speaker is sure about the proposition expressed in the anchor of a TQ. This category can be compared to the source of knowledge (i. e. speaker or hearer) notion, whose importance for the distinction of different types of TQs has been demonstrated in previous literature (e. g. Wilson et al. 2017; König 2017; Heim 2019).

There are two possible values: *sure* and *unsure*. Without access to the speakers themselves, we must decide on the speaker certainty from the available text only. We therefore consider a speaker to be necessarily sure of the expressed proposition in two cases. First, if they are talking about their own experiences (20). Second, if the speaker seems to assume that all interlocutors share the same opinion or can easily directly verify the proposition (21). The speaker is judged as unsure in all other cases, in particular when they request information (22).

(20) SURE

*Ist ja eigentlich perfekt, ne?*  
 Is PRT actually perfect, TAG?  
 ‘That’s actually perfect, right?’ (GC)

(21) SURE

*Es ist so abgefahren, oder?*  
 It is so cool, TAG?  
 This is so cool, isn’t it? (TW)

(22) UNSURE

@user *Du meinst jetzt aber den Camembert-ähnlichen, oder?*  
 @user you mean now but the camembert-like, TAG?  
 ‘@user But you mean the camembert-like one, do you?’ (TW)

### 4.4 Target of confirmation

Further, we annotate the target of confirmation which the speaker is requesting. This category reflects the expectation on the hearer’s (*H*) reply with respect to the proposition (*p*) expressed in the TQ anchor. We define the following four possibilities that emerged though the analysis of our data:

*p* is true, where the hearer is asked to confirm whether the anchor proposition is true.

- (23) *Willst dann noch was sehen vom Land, gell?*  
 want then still something see of the country, TAG?  
 ‘(You) want to see something of the country, don’t you?’ (CH)

*H* believes *p*, where the hearer is asked to confirm that they (also) believe the proposition, i. e. show agreement with it.

- (24) *Alles klar, besser als nichts aber, ne?*  
 all right, better than nothing but, TAG?  
 ‘All right, but that’s better than nothing, isn’t it?’ (GC)

*H* perceives *p*, where the hearer is asked to express their awareness of the proposition, i. e. that they heard what was said.

- (25) *Joa. Dann gute Nacht, ne?*  
 yes. then good night, TAG?  
 ‘Yeah. Well good night then, huh?’ (TW)

*H* accepts *p*, where the hearer is asked to confirm that they accept the command expressed in the anchor.

- (26) *@user Ach, schon wieder. Melde dich, wenn’s vorbei ist, ja? pff*  
 @user Ach, yet again. Report yourself, when it over is, TAG? pff  
 ‘@user Ach, again. Let me know when it’s over, ok? pff’ (TW)

A request for the confirmation of the truth of the proposition has been mentioned in several studies on TQs (cf. König 2017; Heim 2019; Wilson et al. 2017). Our categories *H* accepts *p* and *H* believes *p* appear to capture the aspects of the function of *gell* demanding for the confirmation of the appropriateness of a speech act (Heim 2019: 133). They both represent the speaker as the source of knowledge, and *H* accepts *p* assumes a response by performing an action. Similar information- and action-seeking functions of English TQs are addressed in (Gómez González 2018).

## 4.5 Position in turn

Finally, we differentiate between *turn-internal* (27) and *turn-final* (28) positions of sentence-final tags. In tweets, we consider a tag followed by one or more emoticons, hashtags, or URLs turn-final. By taking into account this category, we

aim to identify whether the position of a tag in a turn is connected to functional differences (cf. König 2017).

(27) TURN-INTERNAL

...mm mm *aber auch ein komisches Gefühl, ne? Und was macht*  
 ...mm mm but also a strange feeling, TAG? And what does  
*er beruflich?*

he professionally?

‘...mm mm a strange feeling, huh? And what does he do for a living?’ (CH)

(28) TURN-FINAL

B: *Dann darf ich dich aber vom Bahnhof abholen oder? A: Ja.*

B: Then may I you but from station pick up TAG? A: Yes.

B: ‘Then I can pick you up from the station, right?’ A: ‘Yes.’ (CH)

## 5 Analysis

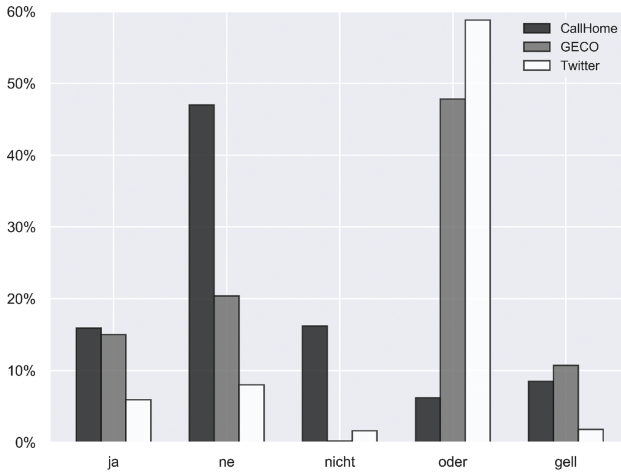
Our analysis of the annotated data reveals several differences among the corpora as well as the individual tags. We test these results for significance using logistic regression.<sup>15</sup> The results for several tags reveal their specific functions in certain contexts and conversation types.

### 5.1 Corpus differences

Figure 1 (based on Table 2) depicts the relative frequency of the five most common tags in the three corpora. The tag distribution in CallHome shows that people make use of all five tags in telephone conversations, of which *ne* appears to be the most common by far (47%). In GECO, *oder* is similarly frequent (48%). And whereas *gell* and *ja* occur with comparable frequency in these two corpora, we find a much lower usage of *ne/nicht* in GECO. Twitter displays a very different picture in terms of tag distribution: *oder* is used in more than half of the cases (58%), whereas the other tags appear with a considerably low frequency (< 8% each).

---

<sup>15</sup> We use the multinomial logit model from the Python package *statsmodels*, v. 0.9.0: [https://www.statsmodels.org/dev/generated/statsmodels.discrete.discrete\\_model.MNLogit.html](https://www.statsmodels.org/dev/generated/statsmodels.discrete.discrete_model.MNLogit.html). We describe the dependent and independent variables in the header of each table. Because of sparse data for many combinations, we could not fit a complete model using all variables.



**Figure 1:** Distribution of the tags *ja*, *ne*, *nicht*, *oder*, and *gell* across the corpora; see Table 2.

The corpora also differ in the overall frequency of TQs with CallHome containing the highest and Twitter the lowest count per 10,000 word tokens: 64 TQs in CallHome, 22 TQs in GECO, and only 0.04 TQs in Twitter (cf. Table 1). The low per-token count of TQs in Twitter is also due to the fact that not all tweets in the corpus are part of conversations. A significant portion of tweets is made up of statements, stories, and other monological content (Scheffler 2014).

From these results, we draw two conclusions: On the one hand, TQs are an important part of all types of spontaneous conversations, whether spoken or written. On the other hand, the conversation type has a large influence on the overall frequency of tags. Private conversations between family or friends (CallHome) contain the highest numbers of TQs, dialogs between strangers (GECO) contain fewer TQs, and the lowest numbers are found in written conversations in computer-mediated communication (Twitter).

## 5.2 Contexts of German question tags

In this section, we analyze the tags' contexts of use based on the annotated categories. An overview of the counts of the annotated TQs is shown in Table 3.

### 5.2.1 Anchor clause type and anchor speech act

In all three corpora, tags are predominantly attached to declarative clauses (Table 4) corresponding to assertions (Table 5). These results go in line with the

**Table 3:** The number of annotated TQs with the tags *ja*, *ne*, *nicht*, *oder*, and *gell* per corpus. Fragments in CallHome and GECO are excluded from these counts. For Twitter, the data already contain no fragments: since more data are available in Twitter, we were able to annotate the necessary number of TQs without fragments.

Tag	CallHome	GECO	Twitter
<i>ja</i>	145	75	60
<i>ne</i>	589	97	82
<i>nicht</i>	190	1	16
<i>oder</i>	83	219	602
<i>gell</i>	108	36	18
total	1,115	428	778

**Table 4:** Distribution of anchor clause type across the corpora.

		ja	ne	nicht	oder	gell
declarative	CallHome	130	564	187	62	105
	GECO	72	97	1	149	34
	Twitter	43	81	16	598	17
imperative	CallHome	6	0	0	0	2
	GECO	2	0	0	0	2
	Twitter	11	0	0	0	0
exclamative	CallHome	0	0	0	0	0
	GECO	0	0	0	0	0
	Twitter	0	1	0	0	1
interrogative	CallHome	8	7	0	13	0
	GECO	1	0	0	70	0
	Twitter	5	0	0	4	0
ambiguous	CallHome	1	18	3	8	1
	GECO	0	0	0	0	0
	Twitter	1	0	0	0	0

findings in previous literature (see Drake (2016: 174) for *oder*; Heim (2019: 134) for *gell*; Gómez González (2018: 123) for English).

It is observable across the different corpora that apart from occurring with declaratives, *oder* and, to a lesser extent, *ja* are nearly the only tags associated with interrogatives/queries. Furthermore, *ja* differs from other tags by appearing with imperatives/directives. Syntactically ambiguous clauses (due to topic drop) occur primarily in the spoken CallHome corpus. In contrast to the (also spoken) GECO corpus, the interlocutors in CallHome are familiar with each other, and therefore more likely to use rather informal language.

**Table 5:** Distribution of anchor speech act across the corpora. *Topic introduction* does not appear in this overview since all occurrences thereof coincide with the anchor clause type fragment, which we do not consider in the present analysis.

		ja	ne	nicht	oder	gell
assertion	CallHome	131	581	190	68	106
	GECO	72	97	1	151	33
	Twitter	34	78	16	596	17
query	CallHome	8	5	0	15	0
	GECO	1	0	0	68	1
	Twitter	5	0	0	4	0
exclamation	CallHome	0	2	0	0	0
	GECO	0	0	0	0	0
	Twitter	0	1	0	0	1
directive	CallHome	6	0	0	0	2
	GECO	2	0	0	0	2
	Twitter	15	0	0	0	0
wish	CallHome	0	0	0	0	0
	GECO	0	0	0	0	0
	Twitter	4	0	0	0	0
exhortation	CallHome	0	1	0	0	0
	GECO	0	0	0	0	0
	Twitter	2	0	0	2	0
commissive	CallHome	0	0	0	0	0
	GECO	0	0	0	0	0
	Twitter	0	3	0	0	0

Due to the high predominance of declaratives/assertions and a sparse number of examples for the other types, the performance of regression tests is not possible. More annotated data are needed to verify whether the observed occurrences of *ja* and *oder* with queries and directives are statistically significant.

This first overview shows that although some tags may specialize to a certain degree (e. g. imperative clauses are almost exclusively seen with *ja*), all tags share a basic context of occurrence. Most frequently, they are seen with declarative assertions. This indicates that all tags show considerable overlap in their use.

### 5.2.2 Speaker certainty

The distribution of the tags with respect to speaker certainty about the underlying proposition of the anchor is shown in Table 6. Two tendencies can be seen across the corpora. First, we observe a clear difference between *oder* and the other tags:

**Table 6:** Distribution of speaker certainty in the corpora.

		ja	ne	nicht	oder	gell
sure	CallHome	97	484	154	6	86
	GECO	68	52	8	0	20
	Twitter	17	45	10	111	6
unsure	CallHome	48	105	36	77	22
	GECO	7	45	1	211	16
	Twitter	43	37	6	491	12

*oder* is used mostly in cases of uncertainty, whereas all the other tags occur mainly in contexts where the speaker is sure. Second, we notice that (especially in CallHome) *ne* is predominantly used in the contexts of certainty. This preference of *ne* for certain and *oder* for uncertain contexts is significant across the corpora (Table 7).

**Table 7:** Significance tests: *sure* for *oder/ne* vs. other tags across the corpora (dep. var.: tag, ind. var.: certainty).

	oder				ne			
	coef	SE	z	p	coef	SE	z	p
CallHome	-3.911	0.431	-9.077	0.000	0.900	0.141	6.368	0.000
GECO	-3.980	0.389	-10.229	0.000	1.040	0.237	4.389	0.000
Twitter	-1.259	0.185	-6.819	0.000	1.540	0.241	6.392	0.000
all corpora	-2.841	0.114	-25.020	0.000	1.643	0.099	16.582	0.000

Importantly, Table 8 shows that in spontaneous speech, speakers have a strong tendency to use TQs even in contexts where the speaker is sure of the anchor proposition. This contrasts with Twitter, where the tags *ja* and *gell* appear more often in contexts where the speaker is unsure (Table 6).

**Table 8:** Significance tests: *sure* in CallHome/GECO vs. Twitter (dep. var.: certainty, ind. var.: corpus).

	coef	SE	z	p
CallHome vs. Twitter	2.1915	0.108	20.286	0.000
GECO vs. Twitter	0.4991	0.132	3.793	0.000
Twitter vs. CallHome/GECO	-1.6770	0.099	-16.962	0.000



### 5.2.3 Target of confirmation

The target of confirmation encodes the kind of reaction the speaker is requesting by uttering a TQ. As in the case of speaker certainty, we see a clear difference in the contexts favoring *oder* (Table 9). Speakers show a preference for *oder* when attempting to confirm the truth of the anchor proposition (*p is true*). The probability for *oder* to appear in such cases is significantly higher compared to the other tags (Table 10).

In the spoken corpora, the other four tags are used mostly to elicit a sign of awareness from the hearer about what was said (*H perceives p*). It is noticeable that *oder* is virtually never chosen to affirm hearer awareness. Our detailed annotation allows us to differentiate whether *oder*'s specialization is mainly for speaker uncertainty or for a specific target of confirmation. Since *oder* is sometimes used in the contexts where the speaker is sure (Table 6), but even in those

**Table 9:** Distribution of target of confirmation across the corpora.

		ja	ne	nicht	oder	gell
H accepts p	CallHome	9	2	0	0	2
	GECO	2	0	0	0	0
	Twitter	21	1	0	6	0
H believes p	CallHome	8	92	22	24	13
	GECO	9	29	0	33	17
	Twitter	13	36	11	253	6
H perceives p	CallHome	93	382	129	2	67
	GECO	60	26	0	0	4
	Twitter	16	22	1	2	0
p is true	CallHome	35	113	39	57	26
	GECO	4	42	1	186	15
	Twitter	10	23	4	341	12

**Table 10:** Significance tests: *oder* vs. other tags expressing *p is true* across the corpora (dep. var.: tag, ind. var.: target of confirmation).

	coef	SE	z	p
CallHome	2.1317	0.249	8.567	0.000
GECO	2.5925	0.242	10.709	0.000
Twitter	1.2197	0.187	6.515	0.000
all corpora	1.8175	0.094	19.332	0.000

cases the intention of the speaker is never to merely ask for the hearer's awareness of the proposition (Table 9), we conclude that the crucial constraint for the usage of *oder* in German TQs is the discourse function intended by the speaker.

We find that TQs expressing confirmation requests for acceptance of a command (*H accepts p*) use *ja* consistently across corpora. Occasionally, this category is also expressed by *ne/gell* (CallHome) and *ne/oder* (Twitter). However, these usages are only statistically significant for *ja* in our data (Table 11).

**Table 11:** Significance tests: *H accepts p* as opposed to other possibilities for target of confirmation for *ja* across the corpora (dep var.: tag, ind. var.: target of confirmation).

	coef	SE	z	p
CallHome	2.771	0.608	4.559	0.000
GECO	24.66	0.000	0.000	1.000
Twitter	4.002	0.466	8.580	0.000
all corpora	3.170	0.356	8.907	0.000

Furthermore, it is noticeable that the category *H perceives p* is particularly relevant for spontaneous conversations, especially CallHome. The results of the regression tests for this category appearing in spontaneous conversation corpora are significant (Table 12).

**Table 12:** Significance tests: *H perceives p* vs. other possibilities for target of confirmation in spontaneous speech corpora (dep. var.: target of confirmation, ind. var.: corpus).

	coef	SE	z	p
CallHome vs. GECO	1.744	0.133	13.063	0.000
CallHome vs. Twitter	3.310	0.172	19.270	0.000
GECO vs. Twitter	1.566	0.200	7.847	0.000

#### 5.2.4 Position in turn

The distribution of position in turn across the corpora shows that all tags have a preference for the turn-final position (Table 13).

Whereas the other tags display similar behaviour across the corpora, we find a surprisingly high number of *oder* in turn-internal position in Twitter. A closer look at these cases reveals that they are mainly TQs that contain a judgment of some situation (29) or an assumption about the hearer (30) followed by some kind of justification or elaboration.

**Table 13:** Distribution of position in turn across the corpora.

		ja	ne	nicht	oder	gell
final	CallHome	133	517	159	81	104
	GECO	68	81	1	196	30
	Twitter	34	54	14	467	13
internal	CallHome	12	72	31	2	4
	GECO	7	16	0	23	6
	Twitter	26	28	2	135	5

## (29) JUDGMENT + JUSTIFICATION

*@username so ein Mist, oder? Und ich sitz jetzt (trotz Schirm!)*  
*@username such a crap, TAG? And I sit now (despite umbrella!)*  
*klatschnass im Zug. Kalt :(*  
*soaking wet in train. Cold :(*  
 ‘*@username that sucks, right? And now I am sitting (despite the umbrella!)*  
*soaking wet in the train. Cold :(* (TW)

## (30) ASSUMPTION + ELABORATION

*@username du magst Menschen nicht so, oder? Zumindest nicht in*  
*@username you like people not so, TAG? At least not in*  
*echt, wenn es keinen Blocken-Knopf gibt...?*  
*real, when EXPL no block button exists...?*  
 ‘*@username you don’t really like people, do you? At least not in reality*  
*where there is no block button...?’ (TW)*

In these cases, the justification/elaboration following the TQ has a subordinate role in discourse and does not carry on the conversation. The speaker may very well still expect an answer to the TQ. In contrast, typical turn-internal TQs in spontaneous speech do not require an answer, as in example (27) repeated in (31).

- (31) *...mm mm aber auch ein komisches Gefühl, ne? Und was macht*  
*...mm mm but also a strange feeling, TAG? And what does*  
*er beruflich?*  
*he professionally?*  
 ‘*...mm mm a strange feeling, huh? And what does he do for a living?’(CH)*

### 5.3 Context correlations

Finally, we investigate the correlations between the annotated features (and the exhibited tags).

Figure 2 shows the pairwise correlations between all categories in our data.<sup>16</sup> Positive correlations are red and negative correlations are blue (the latter exist in particular between opposite values such as *sure* vs. *unsure*). Most of the correlations reflect the patterns discussed in the previous section, e. g. a positive correlation between *ne* and the categories *H perceives p* (0.32) and *sure* (0.36) or a negative correlation between *sure* and *oder* (−0.58). Apart from that, we see the relations between various annotated features, such as a strong correlation between *H perceives p* and *sure* (0.71), and correlations between tags/features and the corpora.

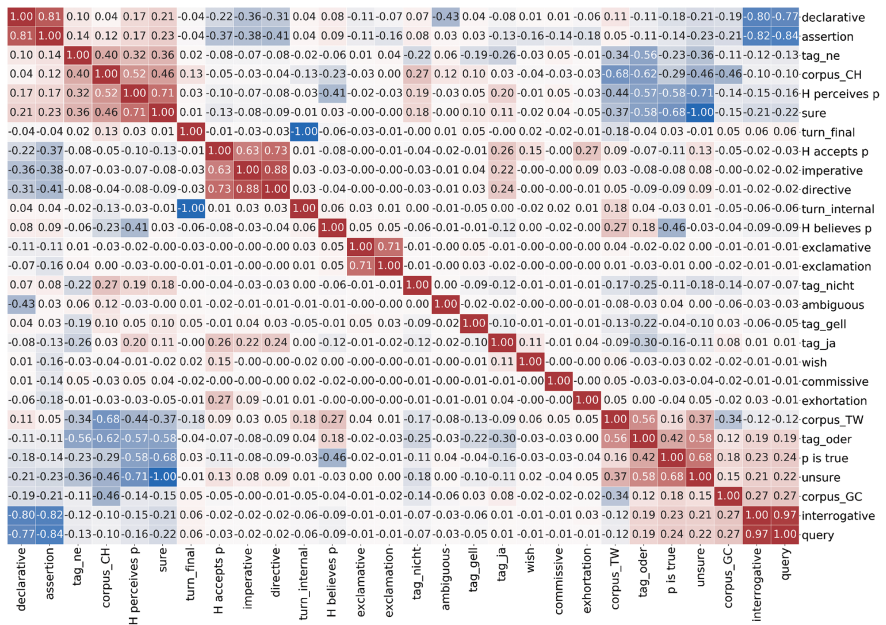


Figure 2: Pearson coefficient correlations between the annotated categories.

<sup>16</sup> The matrix was generated with *seaborn*, v. 0.9.0: <https://seaborn.pydata.org/generated/seaborn.clustermap.html>.

## 6 Discussion and conclusion

In this paper, we analyzed German question tags in three types of conversational corpora: in-person conversations, telephone speech, and computer-mediated communication on Twitter. We found that all corpora document a frequent presence of TQs, showing that this construction is an important characteristic of all kinds of conversations. However, the type of medium is one of the factors that determine the frequency of TQ usage. The most TQs are found in telephone speech, fewer in the in-person conversations and even less in Twitter. Further, the type of corpus has an effect on the tag variants used in it: *ne*, *nicht* and *gell* prevail in telephone conversations, while in Twitter and in-person conversations *oder* is the most common variant. The differences in tag distribution between the two corpora of spontaneous speech can be explained by the different relationship status of the interlocutors: family or close friends (CallHome telephone corpus) vs. strangers (GECO in-person dialogs). The latter may lead to more restrained conversations (cf. Koch and Oesterreicher 1985) with an overall lower usage of TQs.

Our analysis shows that the majority of tags behave similarly across the corpora: they occur with declarative/assertive anchors and ask for confirmation of some kind. However, our results also highlight substantial differences between the analyzed corpora and tags. In our data, tags have a specific function in telephone speech: they occur in contexts where the speaker is sure of the expressed proposition and asks the hearer to confirm their awareness of it. This function is especially relevant for the tags *ne*, *nicht*, and *gell*, where it accounts for the majority of occurrences. Essentially, the same functionality of tags was delineated in (König 2017) for *ne*. This use of tags is all but missing from the Twitter corpus, in part explaining the large difference in the overall frequency of TQs across the corpora. However, the low per-token count of tags in Twitter is also due to the fact that not all tweets in the corpus are part of conversations.

We have identified clearly differentiated contexts for the tags *oder* and *ne*, confirming the contrast between these two tags alluded to before (Zifonun et al. 1997; König 2017). Further, our data showed that the behavior of *ne* is mimicked by *nicht* and *gell*. We found that regarding *ne*, several claims made in the described literature hold only partially. Although it is true that *ne* is preferred in cases where the speaker is sure of the expressed proposition, this is not exclusively so. Contra König (2017), *ne* also quite frequently appears in cases of uncertainty, where it could easily be replaced by *oder*. Furthermore, Hagemann (2009) states that turn-internal tags do not aim for a “true reaction”, and although we have discovered such cases (*H perceives p*) in our data, we also found turn-internal occurrences of *ne* that are used to prompt a confirmation of whether the proposition is true or whether the hearer believes it.

TQs with *oder* are predominantly used to signal doubts about the truth of the speaker's assumptions. Apart from that, speakers use this tag to find out whether the hearer has the same opinion on some matter of judgment. The fact that *oder* mostly occupies the turn-final position goes in line with the uncertainty of the speaker and indicates that a clarifying response is anticipated before the conversation can be resumed. These functions of *oder* seem to be clearly established across the different corpora. However, our results show that *oder* prefers but does not require speaker uncertainty. Instead, the defining characteristic of *oder*-TQs is their discourse function, viz. the request for the hearer's confirmation of the truth of or their agreement with the anchor proposition. We found that in spontaneous speech, all tags except for *oder* are used with significant frequency in contexts where the speaker is sure and/or merely asks the hearer for their acknowledgement of the proposition.

Our current analysis of TQ contexts is based on corpus annotations. In further work, we are augmenting the analyzed TQs with their replies and annotating the response types, in order to tap into the hearers' interpretation of the TQs.

While we do not focus on a semantic account of the lexical meaning of tags in this paper, our findings are compatible with an approach that characterizes the meaning of an expression such as a question tag as interacting with its context, as defined in the notion of meaning potentials (Norén and Linell 2007; Aijmer 2013). A similar approach is taken by Wiltschko et al. (2018) who analyze Canadian *eh* as having a principal function that interacts with its syntactic/pragmatic context to yield several contextual functions. We leave this semantic account of the German tags for future work. Such an account must however go beyond the existing theory in that it explains how and why certain tags, such as *oder* and *ne*, can sometimes, but not always, substitute for each other.

With this corpus analysis, we provide a systematic empirical study of German tag variants across three types of conversational data. We have been able to partially confirm existing analyses of individual German tags in spoken corpora, but also found usages that have not been previously described. Furthermore, we add to the overall study of TQs by analyzing a corpus of written conversations from Twitter, where we show that tags are also used but exhibit different functions than in spoken data. In this way, our results expand the crosslinguistic picture of the discourse behavior of question tags.

**Acknowledgment:** We thank our undergraduate research assistant Lukas Paschen for annotating the data. We gratefully acknowledge the detailed and valuable comments of the reviewers. We thank Manfred Stede for his helpful observations and suggestions. All remaining errors are our own.

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project number 317633480, SFB 1287, project A03.

## References

- Aijmer, Katrin. 2013. *Understanding pragmatic markers: A variational pragmatic approach*. Edinburgh: Edinburgh University Press.
- Andersen, Gisle. 2001. *Pragmatic markers and sociolinguistic variation: A relevance-theoretic approach to the language of adolescents*. Amsterdam: John Benjamins.
- Asano, Yuko. 2008. Semantic analysis of tag questions in Japanese: *deshoo* and *janai ka*. In T. J. Curnow (ed.), *Selected papers from the 2007 conference of the Australian Linguistic Society*, 1–14. University of Adelaide: The Australian Linguistic Society.
- Birner, Betty J. 2013. *Introduction to pragmatics*. Malden, MA: Wiley–Blackwell.
- Cattell, Ray. 1973. Negative transportation and tag questions. *Language* 49(3). 612–639.
- Columbus, Georgie. 2010a. ‘Ah lovely stuff, eh?’ – invariant tag meanings and usage across three varieties of English. *Language and Computers: Studies in Practical Linguistics* 71. 85–102.
- Columbus, Georgie. 2010b. A comparative analysis of invariant tags in three varieties of English. *English World-Wide* 31(3). 288–310.
- Dehé, Nicole & Bettina Braun. 2013. The prosody of question tags in English. *English Language and Linguistics* 17(01). 129–156.
- Drake, Veronika. 2016. German questions and turn-final *oder*. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 17. 168–195.
- Elspace, Stephan & Robert Möller. 2011. Atlas zur deutschen Alltagssprache: *gell?/oder?* <http://www.atlas-alltagssprache.de/runde-2/f19a-b/> (accessed 17 May 2019).
- Erotokritou, Vasiliki. 2014. *The semantics and pragmatics of tag questions in Cypriot Greek*. Universität Konstanz MA thesis.
- Farkas, Donka F. & Floris Roelofsen. 2012. Polar initiatives and polarity particle responses in an inquisitive discourse model. Manuscript. <https://sites.google.com/site/inquisitivesemantics/documents/polar-initiatives-july15.pdf>.
- Gómez González, María de los Ángeles. 2018. ‘God that came out quick didn’t it eh’: Variable and invariable tag questions in spoken British English. In M. Á. Gómez González & J. Lachlan Mackenzie (eds.), *The construction of discourse as verbal interaction*, 189–241. Amsterdam: John Benjamins.
- Hagemann, Jörg. 2009. Tag questions als Evidenzmarker. Formulierungsdynamik, sequentielle Struktur und Funktionen redeguginterner tags. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 10. 145–176.
- Heim, Johannes. 2019. Turn-peripheral management of Common Ground: A study of Swabian *gell*. *Journal of Pragmatics* 141. 130–146.
- Holmes, Janet. 1984. Modifying illocutionary force. *Journal of Pragmatics* 8(3). 345–365.
- Huddleston, Rodney & Geoffrey K. Pullum. 2008. *The Cambridge Grammar of the English language*. Cambridge: Cambridge University Press.
- Karins, Krisjanis, Robert MacIntyre, Monika Brandmair, Susanne Lauscher & Cynthia McLemore. 1997. CALLHOME German Transcripts LDC97T15. Web Download. Philadelphia: Linguistic Data Consortium.
- Kim, Jong-Bok & Ji-Young Ann. 2008. English tag questions: Corpus findings and theoretical implications. *English Language and Linguistics* 25. 103–126.
- Kimps, Ditte. 2007. Declarative constant polarity tag questions: A data-driven analysis of their form, meaning and attitudinal uses. *Journal of Pragmatics* 39(2). 270–291.

- Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe–Sprache der Distanz. *Romanistisches Jahrbuch* 36(85). 15–43.
- König, Katharina. 2017. Question tags als Diskursmarker? – Ansätze zu einer systematischen Beschreibung von *ne* im gesprochenen Deutsch. In H. Blühdorn, A. Deppermann, H. Helmer & T. Spranz-Fogasy (eds.), *Diskursmarker im Deutschen. Reflexionen und Analysen*, 233–258. Göttingen: Verlag für Gesprächsforschung.
- Ladd, D. Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In *Papers from the 17th Regional Meeting of the Chicago Linguistic Society*, 164–171. Chicago.
- Levinson, Stephen C. 2017. Speech acts. In Y. Huang (ed.), *The Oxford Handbook of Pragmatics*, 199–216. Oxford: Oxford University Press.
- Malamud, Sophia & Tamina Stephenson. 2015. Three ways to avoid commitments: Declarative force modifiers in the conversational scoreboard. *Journal of Semantics* 32(2). 275–311.
- Mithun, Marianne. 2012. Tags: Cross-linguistic diversity and commonality. *Journal of Pragmatics* 44(15). 2165–2182.
- Moore, Emma & Robert Podesva. 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38(04). 447–485.
- Norén, Kerstin & Per Linell. 2007. Meaning potentials and the interaction between lexis and contexts: An empirical substantiation. *Pragmatics* 17(3). 387–416.
- Pichler, Heike. 2016. Uncovering discourse-pragmatic innovations: *innit* in Multicultural London English. In H. Pichler (ed.), *Discourse-pragmatic variation and change in English: New methods and insights*, 58–81. Cambridge: Cambridge University Press.
- Rehbein, Ines, Emiel Visser & Nadine Lestmann. 2016. Discussing best practices for the annotation of Twitter microtext. In F. Mambrini, M. Passarotti & C. Sporleder (eds.), *Proceedings of The 3rd Workshop on Annotation of Corpora for Research in the Humanities (ACRH-3)*, 73–84. Sofia: Bulgarian Academy of Sciences.
- Scheffler, Tatjana. 2014. A German Twitter snapshot. In *Proceedings of the 19th International Conference on Language Resources and Evaluation*, 2284–2289. Reykjavik, Iceland.
- Schweitzer, Antje & Natalie Lewandowski. 2013. Convergence of articulation rate in spontaneous speech. In *Proceedings of the 14th Annual Conference of the International Speech Communication Association*, 525–529. Lyon, France.
- Searle, John R. 1975. A taxonomy of illocutionary acts. In K. Gunderson (ed.), *Language, mind, and knowledge. Minnesota studies in the philosophy of science*, vol. 7, 344–369. Minneapolis: University of Minnesota Press.
- Takahashi, Mariko. 2014. A comparative study of tag questions in four Asian Englishes from a corpus-based approach. *Asian Englishes* 16(2). 101–124.
- Tomaselli, Maria Vittoria & Albert Gatt. 2015. Italian tag questions and their conversational functions. *Journal of Pragmatics* 84. 54–82.
- Tottie, Gunnel & Sebastian Hoffmann. 2006. Tag questions in British and American English. *Journal of English Linguistics* 34(4). 283–311.
- Wilson, Guyanne, Michael Westphal, Johanna Hartmann & Dagmar Deuber. 2017. The use of question tags in different text types of Trinidadian English. *World Englishes* 36. 726–743.
- Wilschko, Martina, Derek Denis & Alexandra D’Arcy. 2018. Deconstructing variation in pragmatic function: A transdisciplinary case study. *Language in Society* 47(4). 569–599.
- Wilschko, Martina & Johannes Heim. 2016. The syntax of confirmational: A neo-performative analysis. In G. Kaltenböck, E. Keizer & A. Lohmann (eds.), *Outside the Clause: Form and*



*function of extra-clausal constituents*, 305–340. Amsterdam/Philadelphia: John Benjamins.

Yang, Xiaodong & Martina Wiltschko. 2016. The confirmational marker *ha* in Northern Mandarin. *Journal of Pragmatics* 104. 67–82.

Zifonun, Gisela, Ludger Hoffmann & Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. Berlin, New York: de Gruyter.

## Bionotes

### Yulia Clausen

Yulia Clausen received her master's degree in Computational Linguistics from the University of Heidelberg, Germany. She is currently a Ph.D. student at the University of Potsdam, Germany, where she works as a research assistant in the SFB 1287 "Limits of variability in language", project A03. In her Ph.D., she investigates the role of German tag questions in the Common Ground management in spoken and Twitter data.

### Tatjana Scheffler

Tatjana Scheffler received her Ph.D. in 2008 from the University of Pennsylvania, USA. She currently teaches computational linguistics at the University of Potsdam, Germany. Her research interests center around modelling and processing discourse and dialog, in particular in social media, including coherence relations, non-propositional meaning, and questions.