A PRAGMATIC ANALYSIS OF GERMAN IMPERSONALLY USED FIRST PERSON SINGULAR 'ICH'

Sarah Zobel

Abstract

The German first and second person singular pronouns *ich* and *du* allow for a referential use and an impersonal use. In their impersonal use, both pronouns behave like the impersonal pronoun *man* (Engl. *one*) in generic sentences. I argue that the aspect of impersonally used singular personal pronouns that distinguishes them (i) from each other, (ii) from impersonal pronouns, and (iii) from "ordinary" generic sentences is their pragmatic effects. The semantic contribution of the three pronouns and their containing utterances is discussed before a comparative analysis of the pragmatic effects of impersonally used *ich* and *du* and impersonal *man* is given. The analyses are illustrated with naturally occurring data from a self-compiled data collection. Turning to a more practical topic in the second part of the paper, I discuss a methodological issue regarding corpus-based analyses of low-frequency phenomena, such as impersonally used *ich* in the second part of this paper by reporting a small-scale corpus study.

Keywords: German 1st person singular; Non-standard uses of personal pronouns; Impersonal/generic 1st person singular; Pragmatic effects; Corpus-based analysis; Comparative study.

1. Introduction[1]

Languages commonly provide at least four strategies to convey *general statements about people*, i.e., statements that are true for (practically) all members of the set of people. These are (i) sentences containing a universal quantifier (like English *everyone*), (ii) bare plural or indefinite singular generic sentences (cf. Krifka et al. 1995), (iii) sentences containing dedicated impersonal pronouns (like English *one*), and (iv) sentences containing impersonally used second person singular pronouns (like the impersonal use of English *you*).[2] Compare the sentences in (1).[3]

---

[1] General thanks go to Cleo Condoravdi, Eva Csipak, Regine Eckardt, Magda Kaufmann, Florian Schäfer, Thomas Weskott, and audiences at various workshops. In addition, I thank Bettina Kluge, Barbara De Cock and an anonymous reviewer for helpful comments and criticism on previous versions of this paper. Parts of this paper are based on my dissertation (Zobel 2014) which was written at the University of Göttingen as part of the free-floater group "Noun phrases in intensional contexts" with support from the Courant Research Center "Text Structures".

[2] For reasons of space, I leave aside non-standard uses of plural personal pronouns although they definitely merit their own in-depth discussion and comparison to impersonally used singular pronouns.

[3] All naturally occurring examples in this paper are marked with an asterisk (*) followed by the

(1)
      a. Everyone looks forward to Christmas.   (universal quantifier)
      b. People look forward to Christmas.   (generic bare plural)
      c. One looks forward to Christmas.   (dedicated impersonal pronoun)
      d. You look forward to Christmas.   (impersonally used 2nd sg pronoun)

Regarding their semantics, the four strategies in (1) do not behave uniformly: Statements containing a universal quantifier (1a) express strict universal quantification, and hence differ greatly from generic sentences (1b), which express quasi-universal quantification that allows for exceptions (cf. Krifka et al. 1995). Regarding this fundamental semantic division, sentences containing dedicated impersonal pronouns (1c) or impersonally used personal pronouns (1d) have to be grouped with generic sentences (cf. Condoravdi 1989; Chierchia 1995; Malamud 2006; see also Section 2). However, while the last three strategies in (1b-d) are semantically identical, they differ in their pragmatic effects (cf. Kitagawa and Lehrer 1990; Malamud 2012).

The aim of this paper is to add one more strategy to this list – impersonally used first person singular pronouns as found in German, and compare its pragmatic effects to those of the last three strategies presented above. The central questions I address are conceptual and methodological in nature.

- Do sentences containing impersonally used first person singular pronouns have to be grouped semantically with generic sentences? What does this mean regarding their semantic behavior?
- What are the pragmatic effects of impersonally used first person singular pronouns? How do they compare to the pragmatic effects found with impersonally used second person pronouns and impersonal pronouns?
- What challenges arise if one wants to investigate phenomena like this impersonal use on the basis of naturally occurring data? How much help are corpora in this respect?

German provides a richer paradigm of impersonally used singular pronouns than is usually found:[4] The dedicated impersonal pronoun *man* (2a) and impersonally used second person singular *du*[5] (2b) behave just like their better-studied, corresponding

---

type of **online** source in parentheses after the English translation. All examples that lack an asterisk are constructed.

    [4] The semantics and pragmatics of impersonal pronouns and impersonally used second person (singular) pronouns have been studied in various detail for various Indo-European languages: e.g. English (e.g. Kitagawa and Lehrer 1990; Malamud 2012; Moltmann 2012), Italian (Cinque 1988; Chierchia 1995), and Spanish (Alonso-Ovalle 2002). In addition, Siewierska (2004: 212) and Gruber (2011: 340) list e.g. Dutch, Swedish, French, Afrikaans, Romanian, Greek, Mandarin Chinese, Indonesian, Cree, Godi, Gulf Arabic, Hindi, Kashmiri, Koromfe, Koyra Chin, and Kurdish, among other languages that allow for an impersonal use of second person singular pronouns. Given this list of typologically diverse languages, impersonal uses of second person singular pronouns seem to be a cross-linguistically pervasive phenomenon. Kitagawa and Lehrer (1990: 753ff), however, conjecture that languages need to have a small (preferably binary) address system for impersonally used second person to be available. This would exclude, for instance, Japanese.

    [5] In this volume, Kluge discusses the pragmatic effects of non-referentially used second person singular *du* in several languages, among them German, while Gregersen and Jensen discuss Danish *du* and *man*.

English expressions, impersonal *one* and impersonally used *you* as illustrated in (1c) and (1d), respectively.[6]

(2)   a. *Man freut sich einfach auf Weihnachten.*   (≈ 1c)
         one looks-forward himself simply on Christmas
      b. *Du freust dich einfach auf Weihnachten.*   (≈ 1d)
         you look-forward yourself simply on Christmas

In addition, however, German provides an impersonal use of first person singular *ich* 'I' (Zobel 2010, 2014). A naturally occurring example of this use is given in (3).[7,8]

(3)   *Ich kann doch als Brautpaar nicht von meinen Gästen erwarten, dass sie mir quasi die Feier finanzieren!*
      I can PRT as bridal-couple not of my guests expect that they me more-or-less the party finance
      'Bridal couples can't expect their guests to more or less pay for the party.'*(forum discussion)

The impersonal use of *ich* is part of spoken German, and compared to the referential use of *ich*, it has a considerably low occurrence frequency (see Section 4 for discussion). Given the data presented in this paper, it can nevertheless be concluded, though, that impersonally used *ich* is a productive way of expressing general statements in German. It is a genuine alternative to the dedicated impersonal pronoun *man* and the impersonal use of second person singular *du*.

The paper is organized in two parts. The first part discusses the semantic (i.e. truth-conditional) and pragmatic properties of impersonally used first person singular *ich* in comparison to second person singular *du* and the dedicated impersonal pronoun *man*: Section 2 focuses on the semantic aspects and the connection between the impersonal use and the referential use of *ich* and *du*. Section 3 reports results on the pragmatic effects associated with *ich* on the basis of a data collection compiled from Google and corpus searches. The findings are compared to the pragmatic effects observed in the literature for impersonally used *du* and *man*. This part summarizes and extends a proposal made in Zobel (2014). The second part addresses a methodological consideration regarding corpus-based research in semantics and pragmatics with respect to low-frequency phenomena like the impersonal use of *ich* (Section 4). Section 5 concludes.

Preliminary note 1: The term "impersonal" is used differently by different authors and among different traditions. In this paper, I use the term "impersonal use" only to mean the use of personal pronouns in sentences that express statements about "people in

---

[6] In this paper, only declarative sentences are considered. Impersonal uses of personal pronouns in interrogatives and imperatives are left aside since the semantics and pragmatics of these clause types add additional complications compared to declaratives. Hence, their interaction with the impersonal uses is left for further research.

[7] The string "PRT" is used to gloss German discourse particles.

[8] Given the semantics of German impersonally interpreted pronouns (cf. Section 2), the German data is translated as bare plural generic sentences (see (1b)), whenever possible. English impersonal pronouns are not used in the translations for impersonally used *ich* to not suggest the "wrong" pragmatic effects (cf. Section 3 for the discussion of the pragmatic effects). The only exception is example (9) for which *one* is used to translate impersonally used *ich* since it provides the most adequate translation.

general" as exemplified in (1)-(3). This means that in this terminology, "impersonal uses" of pronouns are non-referential. Hence, other non-standard uses of personal pronouns identified in the literature in which the pronouns refer to specific people different from those specified by the pronoun's person and number do not count as "impersonal uses" in this sense, and are left aside (cf. Helmbrecht 2015 for a recent overview).

Preliminary note 2: Most of the examples in this paper are taken from a data collection of approximately 80 naturally occurring uses of impersonal *ich* that was compiled with Google and corpus searches (*DECOW2011* - beta version[9]) by looking for specific target strings (see Appendix A for a detailed description). The data collection consists of utterances in their full contexts of use, which I employed to study and illustrate the pragmatic effects of the pronouns. With respect to register, nearly all of the examples come from texts which are close to spoken German, e.g. transcribed interviews, forum discussions, and comments on newspaper articles.

2. The semantics of impersonally used *ich* and *du*

2.1. *On the role of substitution tests to identify impersonally used personal pronouns*

Before I address the semantics of impersonally used *ich* and *du* in the following subsection, it is instructive to reconsider the role and reliability of substitution tests to identify impersonally used personal pronouns. This discussion also makes the concept of "impersonal use" that I investigate in this paper more precise.

The data on the impersonal use of German first person singular *ich* is novel and surprising insofar as first person singular pronouns are usually claimed to be "pure" or "automatic" indexicals in the sense of Kaplan (1989), and therefore do not allow for an impersonal use (e.g. Moltmann 2010). An exception regarding this stance are Kitagawa and Lehrer (1990). They argue that English first person singular *I* has an impersonal use (comparable to that of *you*) which occurs "mainly in hypothetical contexts" (Kitagawa and Lehrer 1990: 742). Their only example is given in (4).

(4)   We form a frame or script for this kind of situation. ... Thus, in order to be able to take the subway in New York **I** simply need a 'taking a subway' script or frame, if I have one, and supply now relevant specific information about the situation. But at the same time, I may - even if I take the subway daily - be reminded of yesterday's trip when I met this strange man, or last year's when there was a fire in the subway. If I do not have a frame or script, I may well be reminded of the rather vague and remote (i.e., macro-) information from the model I built when some years ago I took the subway in New York. …
(Kitagawa and Lehrer 1990: 741f)

They argue that the use of *I* in (4) cannot be referential since the example was authored by two people, and *I* can be substituted by *one* "without changing the essential message of the text" (Kitagawa and Lehrer 1990: 742). I argue that neither of their arguments necessarily implies that the use of *I* constitutes an impersonal use in the

---

[9] *DECOW2011* is the 2011 beta version of a corpus extracted from German web domains (.de), see references.

restricted sense employed in this paper.

Uses of first person singular pronouns that are non-speaker-referential are not always impersonal uses (unless "impersonal" is defined in this way, of course). For instance, Nunberg (1993) presents the following constructed example of non-speaker-referential uses of *I* (cf. also Grimberg 1994).

(5) Condemned prisoner: ***I*** *am traditionally allowed to order whatever **I** like for my last meal.* (Nunberg 1993: 20)

He argues that *I* in (5) does not refer to the speaker since there cannot be a tradition for a single person regarding their last meal. Hence, *I* must refer to his status as condemned prisoner, and (5) is in fact interpreted as (6).

(6) The condemned prisoner is traditionally allowed to order whatever he likes for his last meal. (Nunberg 1993: 21)

Note that Nunberg does not classify this use of *I* as "impersonal". He argues that the reference of *I* is shifted to whoever is the condemned prisoner in the situations that *traditionally* picks out. Crucially, he argues, the utterance without *traditionally* has to apply to the speaker for this specific non-speaker-referential use to be available. This, I argue, is a sign that *I* is not used impersonally.

In the literature, substitution with dedicated impersonal pronouns, like English *one* or German *man,* is widely used as a test for whether a pronoun is used impersonally or referentially. This substitution test should, however, be handled with care. Hypothetical contexts are not reliable test environments: Even if substitution with a dedicated impersonal pronoun results in an utterance that is true with respect to the given context "without changing the essential message", the pronoun that was substituted may not have been used impersonally by the speaker.[10] To appreciate this point, consider utterances which express specific instances of a general rule, e.g. (7a) and (7b).

Context: Peter/the speaker considers whether to keep on living with his/her parents or to get his/her own flat. A friend of Peters/the speaker may weigh the two options.

(7) a. *If Peter rents a flat, he has to leave a deposit.*
 b. *If I rent a flat, I have to leave a deposit.*
 c. *If one rents a flat, one has to leave a deposit.*

The sentences in (7a) and (7b) are both specific instances of the general rule that is expressed with the help of *one* in (7c). Since the modalities of renting a flat are common knowledge, one would not assume that (7a)/(7b) express a relation between renting a flat and having to leave a deposit which is valid for Peter/the speaker, in particular. But: in both cases, one would probably say that substitution with *one* results in a true utterance, even though one would not want to argue for an impersonal use of the proper name *Peter* in (7a), or – as I argue – an impersonal use for *I*.

---

[10] This criticism depends on one specific interpretation of Kitagawa and Lehrer's (1990: 742) formulation of "essential message" of an utterance as its truth-conditional content. Unfortunately, their formulation is not very precise, and their paper gives no hints regarding which explication of "essential message" they have in mind.

I propose two additional tests to check whether a pronoun is used impersonally in the sense discussed in this paper: The first test is to check whether the sentence that contains the pronoun and that seems to express a general rule can be uttered truthfully in case the referent of the pronoun in its standard referential use does not fall under the generalization that is supposedly expressed. For instance, German *ich* may be used to state generalizations for groups of individuals for which it is clear that the speaker (= the referent in the referential use) is not a member, and/or can never be a member.

(8)   *Wenn **ich** als Mannschaft gewinnen will, dann muss **ich** auch motiviert auf den Platz gehen.*
       if I as team win want then must I also motivated on the field go
       'If a team wants to win, it also has to enter the field motivated.'*(transcribed interview)

Example (8) expresses a generalization about teams using *ich* even though, obviously, the speaker is not a team, and can never be a team by himself. Hence, *ich* passes the first test.

The second test is to check whether simple matrix sentences without conditional structure or other lexical material that may express hypothetical content can express a general rule using the pronoun under investigation.

(9)   *Geht's noch? Was sich manche Menschen erlauben, finde ich unglaublich. **Ich** werfe doch nicht einen fremden Welpen mitten auf der Strasse auf den Rücken.*
       Goes-it still what themselves some people allow find I unbelievable I throw PRT not a strange puppy in-the-middle-of on the street on the back
       'Are you kidding? It's unbelievable what some people take the liberty of doing. One does not throw someone else's puppy on its back in the middle of the street.'*(blog post)

In example (9), neither the final sentence, nor its preceding discourse contain any lexical material that might express hypothetical content and create a hypothetical context. Nevertheless, the final sentence expresses a generalization about people. Hence, *ich* passes the second test.

To summarize, the phenomenon I call "the impersonal use of singular personal pronouns" exhibits the following characteristics:

- Sentences containing an impersonal use always express a generalization about a (possibly explicitly given) group of people.
- The speaker may use the impersonal use of the pronoun even if the generalization that is expressed does not apply to the referent of the pronoun in its referential use.
- The impersonal use is available not only in hypothetical (sentential) contexts, but also in matrix clauses that are not part of a bigger hypothetical context.

Let us return to Kitagawa and Lehrer's example of English *I* in (4): As Kitagawa and Lehrer state, this use seems to be available only in hypothetical contexts (Kitagawa and Lehrer 1990: 742). Hence, it does not pass the second test. Furthermore, translations of German data like example (9) are consistently judged by native speakers as expressing only (in most cases implausible) speaker-referential statements. Therefore,

this use of *I* in (4) does not constitute an impersonal use as discussed in this paper.[11] The same result obtains for Nunberg's examples of non-speaker-referential *I*. Since example (5) without *traditionally* has to apply to the speaker, it does not pass the first test. In comparison, impersonally used *you* passes both tests, as is easy to see from (10).

(10)   Parent/adult: *As a child, you look forward to Christmas.*

2.2. *Impersonally used* ich *and genericity*

The first step of comparing impersonally used *ich* to impersonally used *du* and the dedicated impersonal pronoun *man* is to show that their contribution on the level of truth-conditional semantics is the same. That is, the three pronouns can be used to express the same general statements. If this is indeed the case, any intuitive difference between the three pronouns has to be located outside of their truth-conditional contribution, i.e., on the pragmatic level (see Section 3).

   Impersonally used *ich* occurs in simple or complex matrix clauses, as illustrated in (9) and (11), and in indicative conditionals, as shown in (12).

(11)   ***Ich** kann als Sysadmin ein Zertifikat generieren und wahrscheinlich allen **meinen** Usern im Mozilla installieren, ohne dass die das merken.*
       I can as system-administrator a certificate generate and probably all my users in-the Mozilla install without that they that notice
       'System administrators can generate a certificate, and probably install it for all of their users in Mozilla without them noticing.'*(forum discussion)

(12)   *Wenn **ich** als Mannschaft solche Spiele abliefere, dann zum Boss gehen und **mich** ausheulen ist auch kein Niveau.*
       if I as team such matches deliver then to-the boss go and me cry-eyes-out is also no standard
       'If a team performs this way, running to the boss to complain is not an acceptable behavior,
       either.'*(comment on news article)

   Regardless of their form, sentences containing impersonally used *ich* always express general statements about (a group of) people. As observed for impersonally used *du* and impersonal *man* in the introduction, general statements expressed with these pronouns have to be grouped with generic sentences. This means that these sentences express general statements about people without expressing strict universal quantification. This is shown by comparing sentences containing impersonally used personal pronouns to sentences expressing strict universal quantification, on the one hand, and to generic sentences, on the other hand. Two properties of statements expressing strict universal quantification are (i) that they do not allow for exceptions and (ii) that they may express accidental generalizations.

   Context: This Christmas, every German company happens to donate 10,000 € of

---

[11] The use of first and second person singular pronouns in hypothetical contexts constitutes an interesting phenomenon in its own right, though, which merits its own, detailed investigation.

their profits.

(13)   *Jede Firma spendet zu Weihnachten 10.000 Euro.*
        every company donates at Christmas 10,000 euros
        'At Christmas, every company donates 10,000 euros.'

In the given scenario, the sentence in (13) is true. It describes an accidental generalization for all German companies; the sentence would instantly be false, if one of the companies did not donate 10,000€. The corresponding generic sentence in (14), however, is false in this scenario.

(14)   *Firmen spenden zu Weihnachten 10.000 Euro.*
        companies donate at Christmas 10,000 euros
        'At Christmas, companies donate 10,000 euros.'

Generic sentences always express lawlike generalizations. This includes laws, rules, norms, and other generalizations that are the result of some underlying regularity. Hence, only if the scenario is changed such that there is a tradition for German companies to donate 10,000€ of their profits at Christmas, the generic sentence in (14) is true. Moreover, unlike sentences expressing strict universal quantification, generic sentences allow for (legitimate) exceptions to the generalization they express. That is, in the modified scenario, (14) will still be judged true even if one or two companies do not donate 10,000€ – for instance, because of economic difficulties.

A comparison between (13), (14), and the corresponding sentence with impersonal *ich* in (15) shows that sentences containing impersonally used *ich* pattern with generic sentences, as well.[12]

(15)   **Ich** *spende (doch) als Firma zu Weihnachten 10.000 Euro.*
        I donate PRT as company at Christmas 10,000 euros
        'At Christmas, companies donate 10,000 euros.'

Example (15) expresses a generalization about companies (*als Firma*) in the same way as example (14) above. It states a rule for companies in general, and it allows for the existence of legitimate, exceptional companies that do not donate 10,000€. The tolerance of legitimate exceptions is also illustrated by the possibility to follow up (15) with (16).

(16)   *... aber natürlich nur, wenn ich dadurch nicht Bankrott gehe.*
        … but obviously only if I because-of-that not bankrupt go
        '... but obviously only if they don't go bankrupt because of it.'

In sum, sentences containing impersonally used *ich, du*, or *man* constitute a type of generic sentence.[13] Generic sentences contrast with episodic sentences, which report a

---

[12] In example (15), the presence of the discourse particle *doch* is not obligatory for *ich* to be interpreted impersonally. It only supports the impersonal interpretation (cf. Section 2.3). Similarly, for some speakers the impersonal uses improve if the sentence contains a sentence-initial *als*-phrase or *as*-phrase, as illustrated in (10).

[13] Note that the implication holds only in one direction: If *ich* is interpreted impersonally, then the

specific situation or event (cf. Krifka et al. 1995). Hence, it is predicted that the impersonal use of *ich* is unavailable in episodic sentences. Indeed, linguistic material that suggests that a specific situation or event is reported decreases the plausibility of an impersonal use for these pronouns (see Section 2.3).

The connection between generic sentences and sentences containing impersonally used second person pronouns and dedicated impersonal pronouns has already been made early on in the semantic literature (e.g. Condoravdi 1989; Chierchia 1995). Another well-known observation is that impersonally used second person pronouns and dedicated impersonal pronouns are freely substitutable for each other (e.g. Kitagawa and Lehrer 1990: 741). As mentioned in Section 2.1, substitutability by dedicated impersonal pronouns is not a fool-proof test for impersonal uses. Nevertheless substitutability is a valid observation to which German also conforms: impersonally used *ich* may be substituted by *du* or *man* without a change in the generalization that is expressed.[14]

(17)   a. *Du spendest als Firma zu Weihnachten 10.000 Euro.*
            you donate as company at Christmas 10,000 euros
        b. *Man spendet als Firma zu Weihnachten 10.000 Euro.*
            one donates as company at Christmas 10,000 euros

I argue that substitutability means that examples (14), (15), and (17) are truth-conditionally equivalent. By compositionality of sentence meaning, impersonally used *ich*, *du*, and *man* have the same contribution on the truth-conditional level. An analysis of impersonally interpreted *ich*, *du*, and *man* that aims to capture their truth-conditional equivalence is proposed in Zobel (2014).

On a non-truth-conditional level, substitutability of *ich*, *du*, and *man* can, however, be questioned: Although the sentences in (15), (17a), and (17b) convey the same generalization concerning companies on the truth-conditional level, they differ greatly in their pragmatic effects and preferred contexts of use. These effects, I argue, are the result of independently conveyed pragmatic aspects of these pronouns.

A final note on the semantics: The impersonal use of *ich* has to be distinguished from referentially used *ich* in counterfactual or hypothetical statements since these sentences neither require, nor necessarily express the existence of an underlying law, rule, or norm.

(18)   *Ich würde als Firma zu Weihnachten 10,000 Euro spenden.*
        I would as company at Christmas 10,000 euros donate
        'As a company, I would donate 10,000 euros at Christmas.'

Example (18) expresses that the speaker would donate 10,000€ if he were in charge

---

containing sentence is a generic sentence. However, the containing sentence may be a generic sentence without *ich* being used impersonally. Consider, for instance, sentences reporting habits of the speaker, as in (i).

(i)   *Ich esse Müsli zum Frühstück.*
         I eat muesli to-the breakfast
         'For breakfast, I eat muesli.'

   [14] Note, the *als*-phrase *als Firma* 'as a company' precludes (17a) to express a specific instance of the generalization.

of a company – without expressing or implying, for instance, the existence of a tradition that companies donate 10,000€. In fact, the impersonal use of *ich* is completely unavailable in counterfactual statements (see Section 2.3). The same observation can be made for the impersonal use of *du*.

In sum, impersonally used *ich* exclusively occurs in generic sentences, and is truth-conditionally equivalent to impersonally used *du* and *man*. Therefore, any difference in use between the three pronouns has to be located at the pragmatic level, as is argued for in Section 3.

In the following subsection, I show that for sentences which provide the right sentential context, occurrences of *ich* and *du* are ambiguous between a referential use and an impersonal use. This might seem like a trivial observation, which, however, is at the heart of all methodological issues when impersonally used *ich* and *du* are to be investigated with corpus-based methods.

2.3. *The ambiguity between the impersonal use and the referential use*

As is the case for the impersonal use of *du* (or English *you*), impersonally used *ich* is morphosyntactically identical to its referential use. Both uses show the same agreement behavior, and share their ability to bind the same pronominal elements (Kitagawa and Lehrer 1990: 744; Zobel 2014: 27ff). Sentences containing *ich* or *du* are, therefore, in principle ambiguous between an interpretation where the pronoun is interpreted impersonally (the "impersonal reading"), and one where it is interpreted referentially (the "referential reading"). This ambiguity is illustrated in examples (19) and (20), for which both possible readings are given.

(19)   ***Ich** kann doch als Kunde nicht immer davon ausgehen, dass alles seriös abläuft.*
       I can PRT as customer not always of-it assume that everything legitimately proceeds
       Impersonal: 'Customers can't always assume that everything is done in a legitimate manner.'*(interview transcript)[15]
       Referential: 'As a customer, I can't always assume that everything is done in a legitimate manner.'

(20)   *Damals konntest **du** ein guter Mensch sein und trotzdem Wahlkämpfe gewinnen.*
       back-then could you a good person be and nevertheless elections win
       Impersonal: 'Back then, one could be a good person and nevertheless win elections.'
       Referential: 'Back then, you (addr.) could be a good person and nevertheless win elections.'
       (Malamud 2012: 10, constructed)

The only restriction that differentiates the impersonal use from the referential use is that the impersonal use is more or less restricted to the nominative forms of the pronouns (cf. Malamud 2006, 2012; Zobel 2014).

While *ich* and *du* are completely ambiguous morphosyntactically, not all sentences containing *ich* or *du* are ambiguous between a referential or impersonal reading. Co-occurring linguistic items may make either the impersonal or the referential reading less

---

[15] The context of this naturally occurring example disambiguates it towards impersonally used *ich*.

plausible or even unavailable. The discourse context[16] may also disambiguate which use was intended for a given utterance, see (19).

From my data collection, I compiled a list of lexical elements that support the impersonal reading of a sentence containing *ich*: *als*-phrases, discourse particles (e.g. *doch* or *ja*), and modal verbs.[17] Crucially, this supporting material neither forces the impersonal reading, nor blocks the referential reading. In example (21), all of the mentioned items are present. Nevertheless, the ambiguity between the referential and the impersonal reading persists.

(21) **Ich** *kann doch als Dienstleister meine Kunden nicht so ignorieren*.
I can PRT as service-provider my customers not like-that ignore
Impersonal: 'Service providers can't ignore their customers like that.'*(online review)
Referential: 'As a service provider, I can't ignore my customers like that.'

So, none of the above items are sufficient for the impersonal reading to arise. Like example (21), example (22) is still ambiguous, and discourse context has to be relied on in order to disambiguate.

(22) *Sowas kann* **ich** *doch nicht machen!*
something-like-that can I PRT not do
Impersonal: 'One can't do something like that!'*(forum discussion)
Referential: 'I can't do something like that!'

However, *als*-phrases may make the referential reading implausible if the nominal argument of *als* cannot be said to hold of the speaker (in the case of *ich*) or of the addressee (in the case of *du*). For instance, it is implausible to assume for (23) that the speaker wants to convey that the nominal argument of the *als*-phrase, *Mannschaft* ('team'), holds of herself, i.e., that she is a team. Hence, *als Mannschaft* completely resolves the ambiguity of *ich* in (23) in favor of the impersonal use.

(23) **Ich** *muss als Mannschaft auf bestimmte Spielsituationen umstellen können.*
I must as team on certain match-situations change can
Only impersonal: 'A team has to be able to adapt to certain match situations.'*(forum discussion)

Instead of supporting an impersonal reading, co-occurring linguistic expressions can also make an impersonal reading less likely, or even completely block it. Expressions of this kind are certain spatio-temporal adverbs, referentially used personal pronouns, and subjunctive marking on the verb.

Spatio-temporal adverbs and adverbials that restrict the content of the sentence to specific, intuitively "small" points in time or locations make an impersonal interpretation of *ich* practically unavailable, although they do not strictly block it. For

---

[16] I follow Kamp (2008: 3) and in fact the general consensus in formal semantics in regarding the "discourse context" as being the interpretation assigned to the part of the discourse preceding a sentence that also contains the participants, time, place, and manner of the previous utterances.

[17] Since the collection was compiled by searching for specific target strings, however, no collocation analysis could be performed to get a better feeling for co-occurring elements (cf. Section 4 and Appendix A).

instance, example (24), which contains the temporal restriction *heute um halb zwei* 'today at half past one', is only understood as a statement about a specific event involving the speaker.

(24)   *Heute habe **ich** um halb zwei einen Baum gefällt.*
        today have I at half two a tree cut-down
        Only referential: 'Today at half past two, I cut down a tree.'

Similarly, referentially used personal pronouns restrict the content to make the impersonal reading less plausible. For instance, example (25) can only be understood as a speaker-obligation because of the co-occurring second person singular pronoun *dich* 'you' (acc.).

(25)   ***Ich** muss mich um dich kümmern.*
        I must me about you take-care
        Only referential: 'I have to take care of you (add.).'

In contrast, example (26), in which *dich* was substituted by *meine Kinder* 'my children' (a noun phrase with a bound possessive), can also express a moral obligation for people in general.

(26)   ***Ich** muss mich um meine Kinder kümmern.*
        I must me about my children take-care
        Impersonal: 'One has to take care of one's children.'
        Referential: 'I have to take care of my children.'

Subjunctive marking of the finite verb (*Konjunktiv II* as exemplified in (27b)) has the strongest effect. It strictly blocks the impersonal use of *ich*.

(27)   a. ***Ich** trinke doch als Sportler täglich einen Proteinshake.*
           I drink PRT as athlete daily a protein-shake
           Referential: 'Being an athlete, I drink a protein shake daily.'
           Impersonal: 'Athletes drink a protein shake daily.'
        b. ***Ich** würde doch als Sportler täglich einen Proteinshake trinken.*
           I would PRT as athlete daily a protein-shake drink
           Only referential: 'If I were an athlete, I would drink a protein shake daily.'

While the sentence in (27a) can be read as a statement about the speaker (referential reading) or as a general statement about athletes (impersonal reading), (27b) only expresses a counterfactual statement about the speaker. Compare also example (18).

Co-occurring referential and impersonal uses of a pronoun are allowed in multi-clausal sentences as long as the tokens occur in different clauses.

(28)   *Also **ich** habe ja wenig Ahnung von Finanzen etc. aber **ich** kann doch als Bank/Wechselstube nicht selbst einen Preis/Wert für Währung/Geld fest-legen?!*
        Well I have PRT little knowledge of finances etc but I can PRT as bank/exchange-office not self a price/value for currency/money fix
        'Well, I don't really know anything about financial matters, but a financial institution can't choose a price/value for currency.'*(forum discussion)

In the first clause of example (28), *ich* is used referentially. With this clause, the speaker conveys his level of knowledge regarding financial matters. In the second clause of (28) that is conjoined with contrastive *aber* ('but'), the speaker clearly uses *ich* impersonally to talk about financial institutions in general. Similar switches can also be observed for subordinate clauses, see (29).

(29) **Ich** *finde,* **ich** *kann als Tourist noch so viel über etwas lesen, aber Hochglanzvideos in dieser Qualität machen einfach so viel mehr Lust das Ganze selbst zu sehen.*
      I think I can as tourist still as much about something read but high-definition-videos in this quality make simply so much more desire the whole self to see
      'I think that tourists can read as much as there is about something, but high definition videos like this tempt them so much more to see all there is in person.'*(blog post)

To summarize, the interpretative ambiguity of *ich* and *du* that arises from the fact that their impersonal and referential uses share the same morphosyntactic behavior is influenced by co-occurring lexical material, which may disambiguate the sentence in either direction. Since the effect of co-occurring material is not always strong enough to completely disambiguate which use was intended, the discourse context has to be relied on as well.

3. Pragmatic effects: Comparing impersonally used *ich, du,* and *man*

As was established in Section 2.2, impersonally used *ich*, *du*, and *man* are truth-conditionally equivalent. Nevertheless, there is an intuitive difference between the three pronouns. This difference, I argue, has to be located in pragmatics. To show this, I first summarize the literature on the pragmatic effects of dedicated impersonal pronouns and impersonally used second person singular pronouns, and afterwards connect and compare the results from the literature to the pragmatic effects observable for impersonally used *ich*.
     To anticipate the result of this comparison: I argue that impersonally used *ich, du,* and *man* form a class of expressions in the sense that they share some of their pragmatic effects, but differ from each other in other individuating pragmatic aspects. As a class, the three pronouns contrast with indefinite noun phrases in "ordinary" generic sentences, as is shown in the first subsection.

3.1. *Previous observations in the literature on the pragmatic effects*

Moltmann (2006, 2010, 2012) and Zifonun (2000) observe for English *one* and German *man*, respectively, that both pronouns are used to express generalizations to which the speaker had a personal connection of some sort: Either (i) the speaker has personal experiences which support the validity of the generalization that is expressed, or (ii) he takes his personal experiences as the epistemic basis from which to infer a new generalization. In both cases, the speaker utters a generalization for which his personal experiences ensure that he also falls under the generalization, or would fall under the generalization if he had the relevant properties.

To consolidate this observation, Moltmann (2006, 2010, 2012) compares generic sentences containing English *one* to ordinary generic sentences.

(30)   a. *One can see the picture from the entrance.*
        b. *People can see the picture from the entrance.*
        (Moltmann 2006: 258, constructed)

She argues that unlike ordinary generic sentences, sentences containing *one* express "identification" on the part of the speaker with the group of people that the generalization is about. In other words, impersonal *one* signals that the result of the speaker's simulation of himself being an individual that the generalization applies to supports the validity of the generalization. For example, a speaker can utter (30a) after seeing the picture from the entrance himself, and then generalize from his experience. Or he might simulate the situation of standing at the entrance in his mind, and generalize from the results of this simulation. Generalizing from a subjective experience, Moltmann argues, is not enough to utter (30b), though. Ordinary generic sentences require "evidence" that is given independently of the speaker's own, particular experience.[18]

The same difference is observable for German *man*:

(31)   a. *Man kann als Besucher das Bild vom Eingang aus sehen.*
           one can as visitor the picture from-the entrance off see
        b. *Besucher können das Bild vom Eingang aus sehen.*
           visitors can the picture from-the entrance off see

Zifonun (2000) describes the speaker-orientation of *man* as in (32).

(32)   The use of *man* for which generalizability is intended is particularly significant on the pragmatic level, and is commonly used in colloquial speech. In this use, it is communicated:
        a. that the generalization that applies to all (relevant) individuals is also applicable to the speaker and
        b. that the speaker experiences could be experienced in the same way by all other (relevant) individuals, as well.
        (translated and shortened[19] from Zifonun 2000: 242)

For impersonally used second person (singular) pronouns, the focus of the investigation regarding their pragmatic effects usually lies in the connection between the pragmatic effects of the impersonal use and the referent of the referential use, i.e., the addressee. The main question is how the pragmatic effects may be connected to, or

---

[18] Moltmann (2006, 2010, 2012) makes certain conceptual assumptions regarding the denotation of sentences. As a result, the speaker-orientation observable for *one* is considered part of its denotation, and not as a pragmatic effect of the pronoun. In an ordinary truth-conditional semantics, this is not an option (cf. Zobel 2014).

[19] Zifonun (2000: 242): "Pragmatisch besonders bedeutsam und auch in mündlichem Sprachgebrauch üblich […] ist die erste genannte, also die sprecherinklusive Verwendung, bei der aber Verallgemeinerbarkeit intendiert ist. Dabei wird zu verstehen gegeben:
        a) dass, was allen (einschlägigen) Individuen widerfährt, natürlich auch für den Sprecher gilt
        b) dass, was dem Sprecher widerfährt, genauso auch für alle anderen gelten könnte."

even derived from the addressee-referentiality of the referential use. While various derivations have been proposed, the consensus in the literature is that the effect of impersonally used second person singular pronouns is to create closeness between the speaker and the addressee.[20] I briefly summarize some of the proposals found in the literature.

Following Laberge and Sankoff (1979), Kitagawa and Lehrer (1990: 749) specifically propose that the rhetorical function of impersonal *you* (sg.) is to induce "situational insertion" for the addressee. Additionally, *you* expresses assimilation on the part of the speaker to a wider generality. That is, by using impersonal *you*, the speaker aims to create an informal camaraderie between herself and the addressee. Furthermore, it is argued that the pragmatic addressee-orientation of impersonal *you* is connected to the observation that the speaker cannot explicitly exclude the addressee from a generalization that he stated with impersonal *you*. Compare (33) and (34).

(33)   *But I have a gift for teaching... Plus, teaching fiction writing is a lot like writing. **You** have to examine manuscripts, use **your** mind, come up with possibilities, respond to characters in situations. In a lot of ways, it's like working on **your** own work.*
       (Kitagawa and Lehrer 1990: 741)

Excluding the addressee in example (34) results in oddness/infelicity:[21]

(34)   **?**You** have to examine manuscripts — I don't mean **you** personally — use **your** mind.*
       (Kitagawa and Lehrer 1990: 743)

Malamud (2006: 84, 2012: 6) suggests that the impersonal use of *you* (and the impersonal use of German *du*) is an invitation for the addressee to put himself into the shoes of the individuals that the generalization covers, i.e., to empathize with these individuals.[22] This empathy effect, Malamud (2006: 84) argues, is best seen when impersonally used *you* co-occurs with other impersonally used pronouns, e.g., the impersonal pronoun *one*. In this case, an "empathy tracking effect" is observable: The addressee's empathy is directed towards those individuals that *you* represents in the clause.

(35)   a. *In those days in England, one had to show you some respect.*
       b. *?In those days in England, you had to show one some respect.*
       (Malamud 2006: 84)

Malamud argues that the sentences in (35a) and (35b) show different addressee orientation with respect to the two argument positions of *show* filled by *you* and *one*. In

---

[20] In this spirit, Gregersen and Jensen (this volume) suggest that the impersonal use of Danish *du* enhances the addressee's involvement and empathy.

[21] Importantly, Kitagawa and Lehrer (1990) distinguish the impersonal use of *you* (sg.) from a vague use of *you* (pl.) since in the vague use, the speaker may explicitly deny a specific, addressee-referential interpretation of *you*, see (i).

(i)   You're – I don't mean you personally – you're going to destroy us all in a nuclear war.
           (Kitagawa and Lehrer 1990: 743)

[22] Malamud (2012) incorporates Moltmann's (2010) insights on *one,* and argues that *one* marks the bearer of the speaker's empathy, and *you* the bearer of the addressee's empathy.

(35a), the addressee's empathy seems to be directed at the object position, while in (35b), the addressee's empathy is directed at the subject position.

Rojas (2011) investigates the pragmatic effects of impersonally used second person singular pronouns using the example of second person singular *tú* in Chilean Spanish.[23] She focuses on its speaker-oriented pragmatic effects, and argues that the underlying strategy of the speaker when using *tú* is to present a personal experience to others as if it were a general truth – to persuade the others that they could find themselves in a similar situation. While this description does not cover all the strategies that Zifonun (2000) and Moltmann (2006, 2012) describe for impersonal *man* and *one*, respectively, she also addresses the connection to the speaker's personal experiences explicitly.

Tarenskeen (2010: 53) discusses the pragmatic effects of Dutch impersonal *je*, including its speaker-oriented aspect. She argues that impersonal *je* is particularly suitable when speakers want to generalize on the basis of their own experiences since it issues an appeal to the addressee: The addressee is "invited to take the speaker's perspective".

In the next section, I use the observations and intuitions discussed in this section to connect impersonally used *ich* to *du* and *man*, and to build up a pragmatically motivated system of impersonally used singular pronouns.

3.2. *Differences in pragmatic effects between* **ich, du,** *and* **man**

In this section, I argue that the pragmatic effects of impersonally used *ich, du*, and *man* are the result of different combinations of two aspects: (i) a speaker-oriented aspect shared by the three pronouns, and (ii) an intersubjective aspect, which is shared only by *ich* and *du,* and occurs in different "flavors" for the two pronouns.[24]

As presented above, Moltmann (2006, 2012) and Zifonun (2000) identify speaker-orientation as the distinguishing property of impersonal pronouns compared to indefinite subjects in generic sentences. For impersonally used second person singular pronouns, the speaker-orientation has so far only been explicitly observed in Rojas (2011) for Chilean Spanish. I argue, though, that speaker-orientation can also be observed for impersonally used *ich* and *du*. This becomes apparent when sentences containing impersonally used *ich* and *du* are contrasted with ordinary generic sentences.

(36) a. *Ich kann als Besucher das Bild vom Eingang aus sehen.*
         I can as visitor the picture from-the entrance off see
     b. *Du kannst als Besucher das Bild vom Eingang aus sehen.*
         you can as visitor the picture from-the entrance off see
     c. *Besucher können das Bild vom Eingang aus sehen.*
         visitors can the picture from-the entrance off see

---

[23] Interestingly, Rojas (2011) chooses the term "evidential *tú*". That is, she argues that the connection to the speaker's personal experiences found with impersonal uses of second person singular pronouns should be linked to the category of evidentiality, which, to my knowledge, has not been proposed before.

[24] The term "speaker-orientation" is used to indicate that by using *ich*, *du*, and *man* impersonally, the speaker conveys something about her own mental attitude with respect to the content of the utterance. The term "intersubjective" is used to indicate that the speaker communicates something which does not only involve his or her own mental attitudes, but which conveys the speaker's assumptions regarding the attitudes of others with respect to the utterance.

Example (36c) is a neutral description of an ability of potential visitors. It simply conveys a state of affairs which could also be expressed by the passive sentence in (37).

(37) *Das Bild ist für Besucher vom Eingang aus sichtbar.*
  the picture is for visitors from-the entrance off visible
  'The picture is visible for visitors from the entrance.'

In contrast, (36a) and (36b) are not neutral in this respect. If (37) were taken as a paraphrase, it would not capture the intuition that, in some sense, the speaker communicates a certain perspective which she takes with respect to the generalization that is expressed. The sentence in (36b), for instance, can be naturally used in a context where the speaker is telling a friend about his own visit to the exhibition:

(38) *Die neue Klimt-Ausstellung ist toll!* Der Kuss *wird auch gezeigt. Das Bild ist das Herzstück der Ausstellung. Du kannst es als Besucher schon vom Eingang aus sehen.*
  the new Klimt-exhibition is great *The Kiss* is also shown the picture is the centerpiece of-the exhibition you can it as visitor already from-the entrance off see
  'The new Klimt exhibition is great. *The Kiss* is also shown. The picture is the centerpiece of the exhibition. As a visitor, you can see it already from the entrance.'

It would be very odd to follow up the first three sentences with the ordinary generic sentence in (36c) in case the speaker recounts personal experiences. Hence, impersonally used *du* in (36b) can be said to also perform the second pragmatic function that Zifonun (2000: 242) describes for *man*: "It is communicated: What the speaker experiences could be experienced in the same way by all other (relevant) individuals, as well" (see (32)). Impersonally used *du* can also communicate the first pragmatic function of *man*, i.e., to convey that a generalization that applies to all (relevant) individuals is also applicable to the speaker. This function is observable when the context for (38) is changed as follows: The speaker tells a friend about the exhibition after reading a leaflet that describes the outline of the exhibition, and claims that visitors can see the picture from the entrance.

Similarly, (36a) may be used naturally in a context when the speaker uses his personal experience or knowledge to argue against somebody else's claims regarding the exhibition.

(39) a. A: *Ich habe gehört, dass* Der Kuss *in der neuen Austellung nicht gezeigt wird.*
    I have heard that *The Kiss* in the new exhibition not shown is
    A: 'I heard that *The Kiss* is not shown in the new exhibition.'
  b. B: *Das stimmt doch gar nicht! Ich kann das Bild als Besucher sogar schon vom Eingang aus sehen!*
    that is-true PRT really not I can the picture as visitor even already from-the entrance from see
    B: 'That's nonsense! As a visitor, one can even already see the picture from the entrance!'

Using (36c) as the second sentence in (39b) again results in oddness. So in sum, it can be concluded that impersonally used *ich* and *du* convey the same speaker-orientation as *man*. I propose to summarize the speaker-orientation as in (40).

(40) **Speaker-orientation (*ich, du, man*):**
      It is communicated that the speaker believes that the generalization does apply, or would apply to her, if she had the required properties.

Whether the speaker's belief is based on first hand experiences or second hand knowledge is not part of what is communicated by the speaker-oriented aspect. While the speaker's epistemic basis for his claim is often inferable from the context, I argue that it is not part of what is conveyed.

Example (41) once more illustrates the effect of the speaker-oriented component. Since *ich* communicates that the speaker believes that he actually or presumably falls under the generalization that is expressed, an explicit denial of this fact afterwards results in pragmatic oddness.

(41) *Ich kann als Besucher das Bild vom Eingang aus sehen. #Aber ich glaube nicht, dass das für mich gelten würde, wenn ich dahin gehen würde.*
      I can as visitor the picture from-the entrance off see but I believe not that this for me be-valid would if I there go would
      'As a visitor, one can see the picture from the entrance. #But I don't believe that this would be the case for me, if I were to go there.'

The same observation can be made for *du* and *man* when substituted for *ich* in (41).

For the impersonal uses of *ich* and *du*, additional pragmatic effects are observed, which involve the creation of distance or closeness between the speaker and other people. These effects are independent of the shared speaker-orientation; impersonal *man* does not show any of these additional participant-related, "emotive" effects.

For German impersonally used *du,* the same pragmatic effects arise as observed for English impersonally used *you* (Kitagawa and Lehrer 1990; Malamud 2006, 2012): The impersonal reading of *du* aims to create closeness or an informal camaraderie between the speaker and the addressee. In addition, it conveys (i) that the speaker invites the addressee to identify herself with the individuals that the generalization is about, and (ii) that the speaker expects the addressee to accept the generalization as valid.

I propose that the addressee-oriented aspect of generalizations using *du* can be made more precise as follows: *Du* conveys that the speaker invites the addressee to consider the truth of the generalization by checking it against her personal experiences and knowledge. In addition, *du* also conveys that the speaker holds the validity of the generalization to be uncontroversial. Hence together, the addressee-invitation and the speaker's expectation convey that the speaker expects that the addressee's check results in him accepting the generalization. From this it also follows that the speaker expects that the addressee, like herself, assumes that he falls/would fall under the generalization expressed by her utterance. Example (42) illustrates that a retraction of the applicability of a generalization to the addressee results in oddness when impersonal *du* was used.

(42) *Du musst gut auf deine Kinder aufpassen. #Aber das gilt nicht für dich persönlich.*
      you must well to your children attend but that is-valid not for you personally
      'You (imp.) have to attend to your children well. #But that does not apply to you (addr.) personally.'

The oddness of the continuation in (42) is in accord with Kitagawa and Lehrer's claim for English *you* in (34). The same effect is found for example (43), where the second sentence retracts the speaker's expectation that the addressee will accept the generalization as valid.

(43)   *Als Erzieherin musst du gut mit Kindern umgehen können. #Aber ich bin mir sicher, dass du da anderer Meinung sein wirst.*
   as childcare-specialist must you good with children interact can but I am sure that you on-that different opinion be will
   'As a childcare-specialist, you (imp.) have to be able to interact well with children. #But I am sure that you (addr.) will have a different opinion about that.'

I argue that the intuition that impersonally used *du* and *you* create closeness between the speaker and the addressee is a result of the interaction between the invitation to consider the validity of the generalization, the speaker's expectations regarding the outcome of that consideration, and the speaker-oriented aspect. Together, these components convey that the speaker expects the addressee and himself to share experiences, knowledge, or opinions. This is the "we're not that different" effect of impersonally used second person singular pronouns.

(44)   **Intersubjective aspect of *du*:**
   It is communicated that the speaker holds the validity of the generalization to be uncontroversial, and invites the addressee to check the generalization against her own knowledge and experiences.

In contrast to the closeness induced by the impersonal use of *du*, the intuitive effect of using impersonally interpreted *ich* is to signal distance between the speaker and other people (but not necessarily other discourse participants). It is most often used to criticize, or to complain about the actions or opinions of others, or to state "unpopular" opinions. The first type of utterance is illustrated in (45), the latter in (46).

Context: Speaker A complains about the service at a hairdresser where she was ignored by the staff until she left without getting her hair cut.

(45)   A: *Hat man dort Angst vor Kunden? Ich fand es wirklich peinlich und unverschämt mit Kunden so umzugehen. Ich kann doch als Dienstleister meine Kunden nicht so ignorieren.*
   has one there fear of customers I found it really embarrassing and rude with customers like-this treat I can PRT as service-provider my customers not like-this ignore
   A: 'Are they afraid of customers there? I found it really embarrassing and rude to treat customers like that. A service-provider can't ignore his customers like that.'*(online review)

Context: Discussion about a soccer match. Speaker A responds to the comment of another speaker who, apart from other things, claims that the referee is to blame for the bad performance of the team under discussion.

(46)   A: *Ansonsten ist, meiner bescheidenen Meinung nach, nicht der Schiri schuld an der grottenschlechten Leistung unserer Mannschaft. Der Schiri ist nun mal Gott auf dem Platz und ich muss als Mannschaft das beste draus machen. Egal wie es kommt.*

apart-from-that is my humble opinion after not the referee to-blame on the abominable
performance of-our team the referee is PRT PRT god on-the field and I must as team the
best out-of-it make no-matter how it comes
A: 'Apart from that I don't think that the referee is to blame for the abominable
performance of our team. The referee is God on the field – that's just the way it is – and a
team has to make the best of it. No matter how it is.'*(forum discussion)

In addition, the impersonal use of *ich* is usually connected to increased emotional
involvement on the part of the speaker, i.e., if the speaker fully supports the validity of
the generalization expressed by his statement, he will react with heightened emotion
when the validity is questioned. So, when a speaker uses *ich* impersonally, she may
additionally convey emotions like irritation, anger, or incredulity with her tone of voice
or her facial expressions. These are, however, not an integral part of the participant-
related component of *ich*.

Just as for impersonally used *du*, I argue that the intersubjective aspect of the
impersonal use of *ich* can be split up into two interacting parts. First, impersonally used
ich conveys that the speaker holds the validity of the generalization to be
uncontroversial (shared with impersonally used *du*). And second, it conveys that the
speaker believes that there is someone in the utterance context or the preceding
discourse who contests the validity of the generalization.

(47) **Intersubjective aspect of *ich*:**
    It is communicated that the speaker holds the validity of the generalization to be
    uncontroversial, and that he believes that someone (nevertheless) contests its alidity.

As for impersonally used *du*, the intersubjective aspect that is conveyed cannot be
explicitly retracted. As illustrated in (48), an explicit retraction of the speaker's full
support of the generalization conveyed with impersonal *ich* results in pragmatic
oddness.

(48) *Ich kann doch als Dienstleister meine Kunden nicht ignorieren. #Aber eigentlich können
    Dienstleister meinetwegen machen, was sie wollen.*
    I can PRT as service-provider my customers not ignore but actually can service-provider
    for-all-I-care do what they want
    'As a service-provider, one can't ignore one's customers (I truly think this is wrong).
    #But for all I care, service-providers can behave however they want.'

The increased emotional involvement that sometimes accompanies the impersonal
use of *ich* is a result of the two conflicting aspects of the intersubjective aspect of this
use: The beliefs that the validity of the generalization is uncontroversial, and that
someone contests the validity stand in direct conflict to each other, and may therefore
trigger the negative emotions listed above. Furthermore, the conflict induces the
distancing effect since the speaker openly conveys that his opinions and those of
another person differ in a way that the speaker finds incomprehensible.

Since *man* does not have an intersubjective aspect, the intersubjective aspects of *ich*
and *du* are best seen in sentences which show Malamud's (2006, 2012) "empathy
tracking effect". In sentences showing empathy tracking effects, the relevant empathy is
directed at the group of people denoted by the pronoun that has an intersubjective
aspect. Examples (49) and (50) illustrate the empathy tracking effect for *du* and *man*,

and for *ich* and *man*, respectively.

(49)   a. *Man hat dich damals dafür ins Gefängnis werfen können.*
            one has you back-then for-this in-the jail throw can
            'Back then, one could have thrown you in jail for that.'
       b. *Du hast einen damals dafür ins Gefängnis werfen können.*[25]
            you has one back-the for-this in-the jail throw can
            'Back then, you could have thrown one in jail for that.'

    As in Malamud's empathy tracking example in (35), the invitation issued for the addressee to consider the validity of the generalization by checking his own experiences and knowledge is directed to the position in which the second person singular pronoun *du*/*dich* is realized. For instance in (49a), the addressee is asked to empathize with those possibly thrown in jail.

(50)   *Aaaaber mal ehrlich, wenn ich als Arbeitgeber eine Stellenanzeige aufgebe, dann checke ich doch wenn das Ding online ist, ob auch alles stimmt. Jedenfalls, wenn ich ein Arbeitgeber bin, bei dem man arbeiten will.*
       but now honestly if I as employer a job-ad place then check I PRT when the thing online is whether also everything is-right in-any-case if I a employer am with who one work wants
       'But wait, honestly, if an employer places a job ad, then he has to double-check online whether the ad is fine. At least, if it's an employer one wants to work for.'*(forum discussion)

    In both sentences of example (50), the speaker uses impersonal *ich* for employers in general. In the second sentence of (50), he uses impersonal *man* to refer to employees in general, and to contrast them with the employers. Since the speaker complains about negligent employers, the set of individuals the speaker's complaint focuses on is the set of employers. Hence, the speaker's emotional focus is reflected in the choice of pronominal form regarding employers and employees.

    In sum, impersonally used *ich* and *du*, as well as the impersonal pronoun *man* express a generalization for which the speaker is sure that it applies to him, or would apply to him, if he had the required properties. The additional effects that are associated only with the impersonal uses of *ich* and *du* affect how the generalizations are presented, i.e., whether the speaker expects his opinion regarding the validity of the generalization to be shared, or not: Impersonally used *ich* is employed to signal distance, while impersonally used *du* aims to build closeness between the speaker and the addressee. I argue that the different pragmatic meaning aspects of impersonally used *ich*, *du*, and *man* indicate that the three pronouns integrate their containing utterance differently into the discourse context, and therefore, serve different communicative purposes.

    The exact grammatical status of the proposed pragmatic meaning components is, at the moment, still unclear. What is certain, however, is that the components are not part of the content that is contributed to the truth-conditions of the containing sentence. Furthermore, it seems to be the case that the speaker-oriented and intersubjective aspects are conventionally associated with the different pronouns, and that the

---

[25] German *man* only occurs in the nominative. For all other cases, the indefinite pronoun *einer* has to be used.

pragmatic effects which they induce are detachable. From this it can be inferred that the pragmatic effects that arise from uttering sentences that contain impersonally used *ich, du,* or *man* are not conversational implicatures: Conversational implicatures are neither conventionally associated with a lexical item, nor detachable in a given context (cf. Grice 1975). A detailed analysis of the type of additionally conveyed "not-at-issue" meaning (cf. Potts 2005; Amaral et al. 2007) is given in Zobel (2014). The reader is referred to this work for details.

3.3. *Conditions on the contexts of use for* **ich, du,** *and* **man**

A detailed analysis of the collected data with respect to how the sentences containing impersonal *ich* or *du* are used in their discourse contexts revealed various patterns for their contexts of use. One result of this investigation is that sentences containing impersonally used *ich* and *du* differ in their preferences regarding their contexts of use. Sentences containing *man*, on the other hand, show no contextual preferences, and can be used unrestrictedly.

The impersonal use of *ich* prefers contexts in which the generalization, rule, law, or norm that is expressed was openly violated, or in which the validity of the generalization was questioned. In other words, the actions or opinions of another individual in the context have to be in conflict with the content of the sentence containing impersonal *ich*. I call this a "*negative context*". The context-data-pair in (51) gives an example of impersonally used *ich* in its original, negative context of use.

Context: Forum discussion on "How much money does one give as a present at a wedding?" - The initial question is whether 100 euros is enough. One user argues that it is customary to adjust the amount of money relative to the size and cost of the wedding party held by the bridal couple. Another user takes issue with this claim with the following utterance:

(51)   *Ich find das ist ein total doofes Argument! Ich kann doch als Brautpaar nicht von meinen Gästen erwarten, dass sie mir quasi die Feier finanzieren!*
I find this is a totally stupid argument I can PRT as bridal-couple not of my guests expect that they me more-or-less the party finance
'I think this is a totally stupid argument! Bridal couples can't expect their guests to more or less pay for the party!'*(forum discussion)

From the other user's previous post, the speaker of (51) infers that this user thinks that a bridal couple can, in fact, expect their guests to pay for the party expenses. Thus, the speaker's opinion expressed in the second sentence in (51) is presented as in conflict with the utterance of the other user. This conflict constitutes the negative context.

Non-verbal, extra-linguistic contexts may also qualify as negative contexts, see (52).

Context: A and B see a person knocking over bikes, and running off.

(52)   A: *Was soll das?! Sowas kann ich doch nicht machen!*
what should this something-like-that can I PRT not do

A: 'What does that guy think?! You (imp.) can't do that!'/ 'One can't do that!'

The speaker's exclamation is a direct reaction to the actions of the person knocking over bikes. These actions openly violate the rule of conduct that underlies the speaker's exclamation in (52), i.e., the rule that one should respect the private property of others.

The preference of impersonally used *ich* for negative contexts is further supported by the discourse particles that co-occur with it. Examples (51) and (52) both contain the German particle *doch,* which is analyzed in the literature as signaling a conflict between the content of its containing sentence and an action or utterance in the discourse context (cf. Thurmair 1989; Zimmermann 2011). Note, however, that the status of a context of use as a "negative context" is independent of the presence of the particle *doch*. This is shown by example (53).

Context: Forum discussion on finding good employees. A previous speaker argues that employers have to train their employees accordingly. Speaker A argues against the view that only employers have to care about further training of their employees, and states that employees have to be proactive to learn new skills.

(53)   A: *Ich kann als Arbeitnehmer nicht erwarten, dass mir mein Arbeitgeber alles vorlegt.*
I can as employee not expect that me my employer everything presents
A: 'An employee cannot expect his employer to present him with everything.'*(forum discussion)

The negative context for the utterance in (53) is the conflicting view expressed by the previous speaker, i.e., that employers are fully responsible for their employee's qualifications. Since (53) does not contain the particle *doch*, the presence of the negative context cannot be attributed to it.

While the presence of the discourse particle *doch* is not necessary, it naturally co-occurs with impersonal *ich*: *doch* can be inserted in any of the examples of impersonally used *ich* that are presented in this paper. This fits with the observation that discourse particles support the impersonal interpretation of personal pronouns (see Section 2.3). Further support for the connection between impersonally used personal pronouns and discourse particles come from an informal survey: All native speakers that I consulted claimed that sentences for which an impersonal interpretation of *ich* is intended improve when the discourse particle *doch* is inserted.

Impersonally interpreted *du* has the same preferred contexts of use as impersonally used *you*. Kitagawa and Lehrer (1990: 752) describe the contexts that impersonally used *you* prefers as contexts in which the speaker expects the hearer to readily share his opinion and subscribe to his statement. These are contexts, I suggest, in which the validity of the opinion expressed by the speaker has not been questioned, and nothing points towards a possible objection. I call these contexts "*positive contexts*". The context-data-pair in (54) shows an example of impersonally used *du* in its original, "positive" discourse context.

Context: Transcript of an interview with the coach of the German ice hockey national team. He talks about the frequent and regular occurrence of situations in professional sports in which weaker teams beat stronger teams. He argues that these situations will continue to happen, and cannot be prevented. He says that in a match,

being the stronger team never guarantees a victory, and continues with:

(54)   *Du musst als Mannschaft einfach mehr gewinnen wollen als der Gegner.*
         you must as team simply more win want than the opponent
         'As a team, your wish to win simply has to be greater than your opponent's.'*(interview)

    The sentence in (54) elaborates on the statement that being the stronger team is not enough to automatically win, which is said to be supported by ample evidence from all types of professional sports. That is, the speaker sees this as an established fact about teams in general. The speaker treats his utterance in (54) on the same level, i.e., as an established fact that neither the interviewer nor anyone else is going to question. In addition, there is no open conflict to which the speaker responds. Therefore, the context in (54) can be classified as a positive context.

    As the reader may have noticed, the discussion on the preferred contexts of use shows various similarities to the discussion of the pragmatic effects of the impersonal use of *ich* and *du* in the previous section. This is not an accident since the components that induce the pragmatic effects of impersonally used *ich* and *du* are, I argue, also responsible for the pronouns' preferences on their contexts of use. This is not a new idea, though, but is based on discussions of the preferred contexts of use for discourse particles, like *doch* and *ja.* For discourse particles, their pragmatic function is also argued to restrict their preferred contexts of use (Zimmermann 2011). The parallel between impersonally used *ich* and *du* and discourse particles that is suggested by this comparison is further supported by the observation that both types of expressions comment on and integrate their containing utterance into the current discourse context (see Zimmermann 2011 and Section 3.2).

    The specific preferences of impersonally used *ich* and *du*, I argue, arise from their intersubjective components. Since these components convey beliefs of the speaker, it is to be expected that these beliefs must not be obviously false for the pronouns to be used felicitously. In other words, the contexts of use need to be compatible with these beliefs. For the impersonal use of *ich*, the preference for negative contexts fits with the observation that the speaker believes that someone contests the validity of the generalization that is expressed. This preference is quite strong: In non-negative contexts, the use of impersonally interpreted *ich* seems needlessly emotional and forceful, and is usually perceived as pragmatically odd. This is illustrated by the oddness of example (55b) compared to (55a) in the given context.

    Context: Taken from a website which reports the rules of football/soccer as established by the International Football Association Board to an interested public.

(55) a.   *Ein Spieler, der vor Spielbeginn des Feldes verwiesen wird, kann nur durch einen der gemeldeten Auswechselspieler ersetzt werden.*
            a player who before match-start the field expelled is can only by a of-the registered substitute replaced get
            'A player who is expelled from the field before the start of the match can only be replaced by a registered substitute.'*(original formulation)
      b.   #*Ich kann als Spieler, der vor Spielbeginn des Feldes verwiesen wird, nur durch einen gemeldeten Auswechselspieler ersetzt werden.*
            I can as player who before match-start the field expelled am only by a registered

substitute replaced get
c. *Du kannst als Spieler, der vor Spielbeginn des Feldes verwiesen wird, nur durch einen gemeldeten Auswechselspieler ersetzt werden.*
you can as player who before match-start the field expelled am only by a registered substitute replaced get

In the context for (55), the speaker can assume that the validity of the rule is not questioned given its source, and no open violation of this specific rule is part of the context. Hence, the context is perfectly compatible with the use of an ordinary generic sentence, as in (55a). In contrast, the impersonal use of *ich* in (55b) expresses that the speaker believes that there is someone who contests the validity of the rule. Since this is not the case, the use is perceived as pragmatically odd.

The preference for positive discourse contexts found with impersonally used *du* is also a result of the pronoun's intersubjective aspect. For instance, impersonally used *du* may occur in the context of example (55) to report the rule, see (55c), but the sentence would sound less official than the ordinary generic sentence in (55a) – which is in line with the pragmatic effects of impersonally used *du* discussed above.[26] The preference for positive contexts is, however, not as strong as the preference for negative contexts for the impersonal reading of *ich*. The pragmatic effects of impersonally used *du* are also compatible with a use in negative contexts: The invitational effect can be used to appeal to the addressee to reconsider an opposing stance or opinion on the topic under discussion. This is illustrated by the dialogue in (56), in which impersonally used *ich* and *du* are used to argue for two opposing points of view.

Context: Forum discussion about a news item: A 10-year-old Belgian girl is pregnant. The father is her 13-year-old friend. Speaker B thinks the parents have breached their duty of supervision. Speaker A doubts that this is necessarily the case and continues:

(56) a. A: *Ich meine - du kannst dein Kind ja nicht auf Schritt und Tritt verfolgen.*
I mean you can your child PRT not wherever-he/she-goes follow
A: 'I mean you can't always follow your child around.'*(forum discussion)
b. B: *Klar können die Eltern ihr Kind nicht auf Schritt und Tritt beaufsichtigen, aber ich muss doch als Eltern merken, wenn mein Kind sich schon über solche Sachen Gedanken macht.*
of course can the parents their child not wherever-he/she-goes supervise but I must PRT as parents notice if my child himself/herself already about such things thought makes
B: 'Of course the parents can't supervise their child wherever she goes, but parents have to notice if their child is already wondering about such things.'*(forum discussion)

Given B's judgment regarding the parents' behavior, the context of use for impersonally used *du* in the first sentence is a negative context: B believes that the parents simply did not pay enough attention to their daughter. However, using *du* in this context results in the feeling that A appeals to B to reconsider her judgment about the

---

[26] De Cock (this volume) discusses the genre and register-specific aspects of impersonally used *you*.

girl's parents since A's utterance implies that there is only a limited amount of time that parents may dedicate to their children. For B's answer, A's utterance provides the negative context: Even though B concedes that A's utterance is true, she is not willing to revise her opinion about the girl's parents and to adopt A's less accusing stance.

Note that the pronouns in (56a) and (56b) can also be interchanged. When they are exchanged for each other, the "roles" of the speakers A and B in (56) are also swapped. That is, A is then perceived as distancing herself from B's accusations, and B tries to change A's mind about the validity of her point of view, i.e., that the parents have breached their duty of supervision.

To summarize, the preferences regarding the contexts of use for impersonally used *ich* and *du* are connected to their pragmatic effects, and are a result of their intersubjective aspects. Specifically, they reflect the speaker's assumptions and expectations regarding the addressee's or another person's attitude towards the generalization that is expressed. The "strength" of the different preferences differs for *ich* and *du*. I argue that this is the case since the distancing effect induced by *ich* does not lend itself as well as the closeness inducing effect of *du* for "paralinguistic effects" (cf. Zimmermann 2011 on paralinguistic effects found with discourse particles).

4. A note on the use of corpora for semantic and pragmatic research

4.1. *An overview of the methodological issues*[27]

As stated in the introduction, the original investigation presented in the previous section (cf. Zobel 2014) was done on the basis of a data collection that was compiled from results of Google and corpus searches of fixed target strings. Analyses of the context-data-pairs in the collection provided crucial pointers towards the generalizations regarding the semantics and pragmatics presented in this paper. Nevertheless, it is still fair to argue that the proposed generalizations are the result of my intuitions regarding the data and those of the native speakers that were consulted at various points with respect to this biased data sample. For obvious reasons, no quantitative studies were performed on this data collection: The method of compiling the data does not meet the necessary standard required for quantitative corpus studies, e.g. random sampling.

So why not use a method of compiling data that may be used for statistical analyses? Quantitative corpus analyses are not part of the "standard repertoire" of formal semanticists, which still mainly rely on introspective judgments regarding naturally occurring and constructed examples, i.e., the kind of data presented in the previous sections. Being part of this tradition, I did not consider using corpus-based methods only until after the research presented in the previous sections was done. What I had to learn, however, was that the task of compiling data that may be used e.g., for collocation analyses of supporting material for the impersonal uses (see Section 2.3), or to give a clean, quantitative overview of the pronouns' preferred discourse contexts (see

---

[27] The following discussion reflects the experiences of a formal semanticist who strives to understand how available corpora could be used for formal semantic research questions, and, therefore, features only a small-scale corpus study for illustrative purposes. A large-scale corpus linguistic study is performed in Gregersen and Jensen (this volume), who investigate the pragmatic effects of impersonally used Danish second person singular *du*, and compare it with the effects of the impersonal pronoun *man*.

Section 3.3) is not an easy one.

The wealth of naturally produced data in corpora is a veritable gold mine for any type of linguistic research. However, with respect to conducting quantitative corpus studies, there are some research questions – especially in the areas of formal semantics and pragmatics – for which freely available corpora are not suited, and also do not seem to provide any effective advantage over looking for illustrative examples on Google. Research questions of this kind concern the investigation of a specific interpretation of an ambiguous expression that is not distinguished from the other interpretations by morphological, syntactic, or lexical criteria. Phenomena that fall under this description are e.g. impersonal uses of personal pronouns, the generic reading of bare plurals, and relevance conditionals.

The main reason why corpora do not provide a distinct advantage is that the ambiguity of these linguistic expressions can usually only be resolved clearly in the discourse context. Moreover, the uses that need to be filtered out are sometimes equally frequent, or even more frequent than the use of interest. Since most corpora standardly employ part-of-speech (POS) tagging with little to no semantic information encoded in the tagset[28], no targeted search for tokens of the relevant use can be performed. Therefore, to build an unbiased data collection that could be used for collocation analyses or other quantitative analyses, one has to sort through a possibly quite large random sample of tokens, and identify and tag the relevant tokens by hand.[29]

In those cases where the relative frequency of the different uses of some expression is nearly the same, sorting through data by hand might still be effective. However, some phenomena are low-frequency uses that may be so rare compared to the other uses that finding enough relevant cases cannot be done efficiently (see also Nogué and Bladas, this volume), and building an unbiased data collection will, therefore, be a highly time-consuming enterprise. In addition, some uses may be confined to specific registers and contexts of use, and therefore require corpora that provide enough texts of this kind.

Below I show that the impersonal use of *ich* is one such low-frequency phenomenon. As is the case for first person singular pronouns in other languages, *ich* is one of the most frequent words in German, and is predominantly used referentially. In addition, impersonally used *ich* is a phenomenon of spoken German, and is nearly exclusively employed in colloquial speech. As the analysis of the pragmatic effects and the preferred contexts of use in Section 3 showed, impersonally used *ich* is even more constrained since it seems to occur only in argumentative discourse. So in sum, one needs an efficient way to discard substantial amounts of referentially used tokens in corpora that contain argumentative texts in colloquial German to get the desired examples.

To my knowledge, corpora are not specialized in this way. Most openly available

---

[28] An exception is the differentiation of uses of a single expression that result in different syntactic behavior. The STTS tagset, for example, distinguishes between "substituting demonstrative pronouns" and "attributive demonstrative pronouns" which codes the difference between demonstratives that occur anaphorically with no argument noun and demonstratives that occur together with an argument noun. See http://www.ims.uni-stuttgart.de/projekte/corplex/TagSets/stts-table.html

[29] Tagging data samples extracted from corpora by hand introduces the problem of *inter-rater reliability*, that is, the degree of agreement regarding the kind of use of a lexical item among different raters. As Tarenskeen (2010: 31ff) and Gregersen and Jensen (this volume) show, in the case of impersonal and other non-referential uses of personal pronouns, defining strict criteria for identifying a certain use may not be possible, and indeed there is no consensus among different raters.

existing corpora for German are compiled from balanced amounts of text of controllable origin. This means they contain nearly exclusively examples of written German, e.g., literary texts, newspaper articles, or scientific texts. Compared to corpora of written German, corpora of spoken German that are publicly available are relatively rare, and are much smaller than the written corpora. One exception is the corpus for spoken German "Gesprochene Sprache", which is part of the DWDS corpora, and contains transcripts of recorded speeches, political debates, interviews, and discussions; and more recently, since 2014, the "Datenbank für Gesprochenes Deutsch" (database for spoken German) provided by the IDS Mannheim.

In the following subsections, the difficulties one has to face when using corpora for a certain type of semantic and pragmatic research question are presented and discussed using impersonally used *ich*. Specifically, the results of an attempt to determine the relative frequency of the impersonal use of *ich* are reported and commented upon.

4.2. *Determining the relative frequency of impersonally used* ich

4.2.1. Details of the investigation

To determine the relative frequency of occurrence for the impersonal use of *ich,* random samples of 150 tokens of *ich* from three DWDS corpora were checked with respect to their use/interpretation. I used two corpora of written German, "Kernkorpus 20" and "DIE ZEIT", and the corpus for spoken German "Gesprochene Sprache": "Kernkorpus 20" is compiled from literary, journalistic, scientific, and non-fictional texts. "DIE ZEIT" exclusively contains journalistic texts, which also includes transcribed interviews. "Gesprochene Sprache", as stated above, contains transcriptions of speeches, political debates, interviews, and discussions. For details consult Appendix B.

As a first step, the overall frequency of the first person singular pronoun in nominative case was determined for the three corpora.[30] For the results, see Table 1.

	# of tokens of *ich* per 1,000,000 tokens
Kernkorpus 20	~ 4,714.56
DIE ZEIT	~ 2,405.42
Gesprochene Sprache	~ 11,522.8

Table 1: Number of tokens of "ich" per 1 million tokens for the three corpora (as per Nov. 2012)

Brief remark: *ich* occurs a lot more frequently in the corpus for spoken German than in the two corpora for written German. As will be shown below, this difference had no effect on the results regarding the relative frequency of the impersonal use of *ich*.

The random selection of 150 tokens was made by exporting the first 10,000 occurrences of *ich*[31] for each of the three corpora, and selecting 150 items randomly

---

[30] Note that it suffices to look for sentences containing tokens of the nominative form *ich* since the impersonal use nearly exclusively occurs only with first person singular pronouns in nominative case.

[31] Search strings for "Kernkorpus 20" and "Gesprochene Sprache" are case sensitive. For these two

with the help of the statistics program R. For the tokens in the random selection, the items were annotated with respect to the use of the pronoun in the given context.[32] The following levels for "use" were employed:[33]

- clear referential use (0)
- clear impersonal use (1)
- other (9)

Whether an item constitutes a "clear referential use" or a "clear impersonal use" was determined purely intuitively from the context. The label "other" was used for tokens of *ich* that (i) occur in quoted text, titles, headlines, or slogans, that (ii) are part of the nominalized, fixed expression *das Ich* ('the I, the self, the ego') as used in psychology and philosophy, that (iii) occur in non-declarative sentences, and cannot be clearly classified as referential, or (iv) for examples for which the use cannot be clearly disambiguated in the given discourse context.

The examples were annotated only by myself, but twice – once in 2012 and once in 2015 – on the basis of the same annotation guidelines. Hence, the classification is not based on the judgment of multiple raters. Nevertheless, I give the measures of agreement for the two annotation rounds.[34]

	Kernkorpus 20	DIE ZEIT	Gesprochene Sprache
Measures of inter-annotator agreement	Percentage: 97.3% Cohen's κ: 0.625	Percentage: 96.7% Cohen's κ: 0.532	Percentage: 100.0% Cohen's κ: 1.000

Table 2: Inter-annotator agreement measures for the annotations made in 2012 and 2015

For all items with diverging annotations, I reconsidered both annotations, choosing the one that conformed better to the annotation criteria given above.[35] The relatively high inter-annotator agreement values result purely from the mass of unambiguously referential instances of *ich* (see Figure 1).

Figure 1 presents the absolute and relative frequencies for the classes 0, 1, and 9 as found in the 150 randomly selected items from each corpus. For each of the corpora, the results are discussed in detail in the following subsections.

---

corpora, the first 5,000 occurrences of "ich" and the first 5,000 occurrences of "Ich" were exported and merged, and the sample was selected randomly from this collection. "DIE ZEIT" searches were not case sensitive when the study was conducted.

[32] The context for the exported tokens are ~1 sentence before and after the sentence containing the token for "Kernkorpus 20" and "DIE ZEIT" and ~2 sentences before and after the sentence containing the token for "Gesprochene Sprache".

[33] The specific numerical values (0,1,9) do not have any special significance, but were chosen to code the three classes.

[34] During the rating process, I experienced the uncertainty with respect to how to annotate the examples as reported in Tarenskeen (2010) only for the truly ambiguous cases (see footnote 29).

[35] These cases nearly exclusively involved the question whether an item was to be classified as "other", or not.
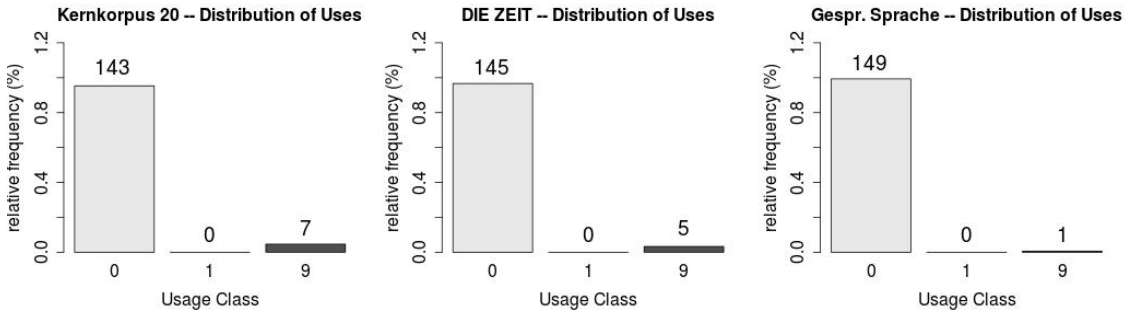
Figure 1: Absolute and relative frequencies of the usage classes in the data

4.2.2. Results for "Kernkorpus 20"

Of the 150 tokens of *ich,* 143 tokens (95.3%) were clearly used referentially, no token was annotated as a clear impersonal use, and 7 tokens (4.7%) were judged as "other". In 5 of the items that were judged as "other", *ich* occurs in quoted text, book titles, or slogans. These tokens were not classified as either a clear referential or a clear impersonal use. Another sentence judged as "other" contained the fixed expression *das Ich* 'the I'. Only one item in the "Kernkorpus 20" sample was clearly ambiguous between a referential and an impersonal use of *ich*:

(57) *Er stimmt insofern, als er ja den Imperfekt benutzt, aber er stimmt insofern nicht, als er es ja immer noch tut. Versetze **ich** den Satz aber in das Präsens, "Dieser Satz steht im Imperfekt", stimmt er gar nicht mehr.*
he is-true insofar as he PRT the imperfect-tense uses but he is-true insofar not as he it PRT always still does put I the sentence however in the present-tense this sentence is in-the imperfect-tense is-right he even no more
'It is true insofar as the imperfect tense is used, but it is false insofar as the sentence is still in imperfect tense. If **I** put the sentence in the present tense, "This sentence is in the imperfect tense", then it is not true any more.'[36]

In the given context (the first sentence of the example), the token of *ich* in (57) can be judged neither as a clear example of a referential use, nor of an impersonal use. While substitution of *ich* with *man* is possible, and does not change the "essential message" of the sentence, *ich* may still be used referentially since it occurs in a hypothetical context (see Section 2.1). Since the given preceding discourse context also does not disambiguate the use of *ich*, (57) is classified as "other".

4.2.3. Results for "DIE ZEIT"

In the sample taken from "DIE ZEIT", 145 tokens (96.7%) of *ich* were clearly used referentially, none of the tokens were clearly used impersonally, and 5 tokens (3.3%)

---

[36] From the second sentence, it can be inferred that the pronoun *er* ('he'; translated as 'it') in the first sentence refers to the sentence *Dieser Satz stand im Imperfekt* ('This sentence was in the imperfect tense').

were judged as "other". Of these 5 tokens, two occurred in headlines which were both direct quotes, two in titles of books, and one token occurred in the fixed expression *das Ich*.

4.2.4. Results for "Gesprochene Sprache"

Of the 150 tokens of *ich* extracted from "Gesprochene Sprache", 149 tokens (99.3%) are clearly used referentially, none is used clearly impersonally, and one token (0.7%) was judged as "other". This example is ambiguous between a referential and an impersonal use. It is given in its full context in (58) without glossing. It is the transcript of a contribution of one participant of the literary program "Literarisches Quartett", a format in which four literary critics discuss literary works (see Appendix B):

(58)   *Es hat etwas, was ich auch selten finde im deutschen Gegenwartsroman, es hat wirklich ein großes kindliches Staunen. Was aber nicht ein so naives Staunen ist, also das Kind steht vor der Welt und weiß nicht, was es ist, sondern es ist ein ganz tiefsinniges Staunen, was man wenig findet so in kleinen Dingen. **Ich** stehe in einem Selbstbedienungsrestaurant, alles ist irgendwie eklig, und ich frage mich plötzlich: Was mache ich hier? Warum bin ich hier? Dieses ursprüngliche Staunen davor, dass die Welt so ist, wie sie ist.*
'It has something that I rarely find in German contemporary novels; it really has a great, childlike amazement. It's however not naive amazement, as if a child is standing in front of the world and doesn't know what it is, but it's a very profound amazement that one rarely finds in small things. **I** stand in a self-service-restaurant, everything is somehow disgusting, and suddenly I ask myself: What am I doing here? Why am I here? This primal amazement that the world is the way it is.'[37]

As in the case of (57), the use of *ich* in (58) cannot be classified clearly as containing either a referential or an impersonal use. In its discourse context, the sentence containing the relevant token of *ich* seems to be completely off topic at first. The only possibility to integrate it into its preceding context is to understand it as the description of a setting in which one would experience the "profound amazement" that the speaker mentions in the first two sentences. That is, the speaker switches seamlessly into a hypothetical context. Hence, example (58) comes close to Kitagawa and Lehrer's (1990) example of "impersonal *I*" discussed in Section 2.1. Another property that (58) shares with Kitagawa and Lehrer's example and (57) is that *ich* can be substituted by *man*.

4.2.5 Discussion

If the two ambiguous examples from "Kernkorpus 20" and "Gesprochene Sprache", which I annotated as "other", are counted as impersonal uses of *ich*, one could conclude that the relative frequency of the impersonal use in these samples is one in 150 (0.67%) for the two corpora, or two in 450 (0.44%) if all three samples are combined. Whether

---

[37] *Es* ('it') in the first sentence refers to the book under discussion.

this result reflects the "actual" relative frequency of this use is questionable. For corpora which contain a larger amount of texts that conform to the (stylistic) environments of impersonally used *ich*, it cannot be excluded that the relative frequency in a similar sample of 150 tokens will be higher. The referential use of *ich* can, however, be expected to be the overall predominant use, independently of the corpus.

The three corpora that were used in this investigation seem to contain only few texts of the types that provide the relevant contexts for impersonally used *ich* to arise. So, one might wonder whether any clear examples of impersonal uses can be found in these corpora in the first place. That this is the case is illustrated in (59)-(61) without glosses.[38]

Context: Newspaper article on legal actions against a German TV station. The article contains part of an interview with P. Tamm, the CEO of a publishing house and a shareholder of the TV station.

(59) *Der Abschluß des Riesengeschäfts trotz der Ablehnung im Aufsichtsrat verstößt gegen alle Regeln des Wirtschaftslebens. "Mir verschlägt's noch immer den Atem", empört sich Tamm, "**ich** kann doch als Kaufmann nicht nach Belieben über das Vermögen der Gesellschafter verfügen."*
'That the giant deal was closed in spite of its rejection by the board of directors infringes upon every rule in the economic world. "I'm still shocked", says Tamm appalled, "a businessman can't just dispose of the shareholders' assets at will."' (Kernkorpus 20)

Context: Transcription of an interview with the head of a German bank. The topic are the low base rates offered by the bank, which earned the institution a lot of criticism. The interviewer comments that offering low base rates is only wise if it nevertheless leads to a profit for the institution. The head of the bank answers that his bank will not have any problems and continues with:

(60) ***Ich** kann als Bank bei den Kreditzinsen immer dann Zugeständnisse machen, wenn ich mit dem Kunden Zusatzgeschäfte vereinbare, zum Beispiel für ihn den internationalen Zahlungsverkehr abwickle.*
'A bank can always make concessions regarding the interest on credits in those cases where additional business has been arranged with the customer, for instance handling his international money transactions.' (DIE ZEIT)

Context: Transcript of a conversation with a miner on his life in the village adjacent to the mine.

(61) ***Ich** kann als Arbeitgeber auch zu den einzelnen hingehen: "Komm Sie heut nachmittag auf mein Büro..." Schriftliche Kündigung kann ja sein, aber die machens nicht, damit dat Volk auf dem Werk nich rebellisch wird!*
'An employer can come up to each single person, "Come to my office this afternoon..." A written notice of termination is also possible, but they don't do it to not make the people working at the mine rebellious!' (Gesprochene Sprache)[39]

---

[38] The elaborate contexts were determined by further investigations of the origins of the sample sentences (e.g. googling for substrings of the examples). The classification of the examples as impersonal uses was also checked against the original texts.

[39] Note that the orthographic idiosyncrasies in the transcription suggest that the speaker is from the

These instances of impersonally used *ich* were found with more sophisticated versions of the strategies that were employed to compile the data collection used in Sections 2 and 3.[40] The full results of the two main searches (for upper and lower case) are given in Table 3 (cf. Appendix A).

	Referential uses	Impersonal uses	Other
"@ich #0 muss #2 als #0 $p=NN" - "I must (max. 2 words) as (noun)"			
Kernkorpus 20	1	0	0
DIE ZEIT	10	1	2
Gesprochene Sprache	2	0	1
"@ich #0 @kann #2 als #0 $p=NN" - "I can (max. 2 words) as (noun)"			
Kernkorpus 20	15	1	2
DIE ZEIT	45	6	9
Gesprochene Sprache	7	1	0

Table 3: Results of two searches of specific target strings in the three DWDS corpora

The relative frequency of the unambiguously impersonal uses for "DIE ZEIT" are one in 13 (7.69%) and six in 60 (10%) for the two searches. Since this result was determined from searches for specific target strings, the result cannot be seen as reflecting the relative frequency of the impersonal use in "DIE ZEIT". The result, however, illustrates the effect of searching for a restricted sentential context which is known to support an impersonal use of *ich*. Given the increase in the percentage of relevant examples found in the three corpora, searching for specific target strings seems to be a promising method to (at least) find illustrative examples. This method, of course, presupposes that one is aware of which restricted sentential context will increase one's success rate. In case of the impersonal use of *ich*, the sentential contexts listed in Appendix A are based on my own and my informants' native speaker intuitions, and were collected gradually.

Searching for specific target strings is not without problems, though. By restricting the sentential context in the searches to a small subset of the possible sentential contexts in which the phenomenon under investigation occurs, potentially interesting data which does not conform to the search string is systematically disregarded. Employing different search strings which cover all possible sentential contexts may be a possible remedy for this situation, but requires a thorough analysis of the phenomenon beforehand.

In corpora which are more tailored to the texts in which impersonally used *ich* are usually found, one might expect a similar increase in the relative frequency of the impersonal use. A corpus of this kind, to my knowledge, does not exist at the moment, but might consist of forum discussions, transcribed interviews, blog posts, and

---

Berlin area.

[40] Specifically, the strategy is to search for specific target strings with part-of-speech tag sequences for variable content.

comments on online newspaper articles, i.e., written texts that are close to spoken language. Whether this expectation is borne out remains to be checked.

4.3. *Checking the pragmatic effects of impersonally used* ich*?*

Given that examples of clear impersonal uses of *ich* can be found in the corpora, could these examples be used to check or refine the analysis of the pragmatic effects put forth in Section 3? To determine whether the examples support the proposal, one needs to look at the broader discourse context. However, as stated above, the DWDS corpora only provide a limited pre- and post-context (see footnote 32) which does not suffice to determine whether a discourse context should be classified as a positive, or a negative context.

The restricted access to the discourse context is an issue that is not unique to the DWDS corpora: Online corpora usually cannot provide full texts – which provide the much needed full discourse contexts – for copyright reasons. This limitation greatly constrains research on phenomena of the type described in Section 4.1, i.e., phenomena where ambiguous linguistic expressions need their discourse contexts to be disambiguated. For these phenomena, data found via Google has the advantage that the entire discourse context is available. However, the search possibilities of Google are limited and often biased, and it is needless to say that no serious quantitative study can be made based on Google results. The "obvious solution", i.e., to build one's own corpus or to get one's hands on corpora that provide full texts, is not always a viable option for each individual researcher since it strongly depends on the available resources, man-power, and infra-structure.

In sum, the way in which I perceive the possibilities provided by corpus-based analyses for formal semantic and pragmatic research is as follows: Corpora are invaluable sources of data. For qualitative and exploratory studies, corpora have – even for low-frequency phenomena like impersonally used personal pronouns – the advantage of providing great user-interfaces and allowing for precise searches. To construct a data sample that may be investigated with quantitative methods to test hypotheses, however, will be highly time-consuming for low-frequency phenomena. Additionally, any discourse pragmatic studies suffer from the lack of availability of the full discourse contexts in freely available corpora.

5. Summary

The aim of the paper was to shed light on the semantic and pragmatic behavior of the impersonal use of German first person singular *ich*. Like German second person singular *du*, *ich* provides a referential use and an impersonal use, which may result in ambiguity of their containing utterances. The impersonal use of *ich* was investigated by comparing it to the impersonal use of *du* and the dedicated impersonal *man* based on previous observations on *du* and *man* from the literature and novel data that was compiled for *ich*.

The comparison in Section 2 showed that impersonally used *ich*, *du*, and impersonal *man* share the same semantic behavior; substitution of one of the pronouns for another

does not change the truth-conditions of the containing sentence. Furthermore, it was shown that sentences containing *ich* and *du* can be fully ambiguous between a "referential reading" and an "impersonal reading". This ambiguity is sometimes resolved by certain expressions or the discourse context.

With respect to the meaning aspects that induce the pronouns' pragmatic effects, it was shown in Section 3 that the three pronouns share a speaker-oriented aspect, but overall differ in which pragmatic effects they induce and which contexts of use they prefer. All pragmatic meaning aspects, however, convey mental attitudes of the speaker with respect to herself, her interlocutors, or a third person. The preferred discourse contexts for impersonally used *ich* and *du* were also connected to these meaning aspects, specifically the intersubjective aspect, which *man* lacks.

In Section 4, a methodological issue regarding the use of corpora in the investigation of phenomena like the impersonal use of personal pronouns was discussed. The problem one faces when investigating low-frequency phenomena of high-frequency expressions was illustrated by discussing a small-scale study to determine the relative frequency of impersonally used *ich* from three DWDS corpora.

A. Appendix A

The data collection that serves as a basis for this paper was compiled by Google and corpus (*DECOW2011*) searches of fixed string-wildcard-combinations that promised relevant results based on native speaker intuitions. The following string-wildcard-combinations and more specific versions thereof were used:[41]

- "ich * als", "ich muss als", "ich kann als" ('I */must/can as')
- "ich * doch als", "ich * ja als" ('I * PRT as'; *doch* and *ja* are German discourse particles)
- "wenn ich als" ('if I as')

As part of the more specific searches, various nouns that do not denote single individuals (e.g. *Mannschaft* 'team', *Familie* 'family', and *Firma* 'company') were used as arguments of *als* 'as'. This strategy resulted in even more successful, targeted searches. The rationale behind this choice was that the *als*-phrases explicitly state which group of entities the generalization is about, and therefore nouns that do not denote single individuals preclude a referential use of *ich* or *du*. The strategy was, however, also successful for nouns that may apply to the speaker or addressee.

B. Appendix B

The following information is taken from http://www.dwds.de/ressourcen/korpora/ (German, as per Nov. 2012). For more detailed information, please visit the homepage of the DWDS.

Gesprochene Sprache: ~ 2.5 million tokens; among them:

---

[41] The asterisk * stands for an undefined number of unspecified lexical material.

- speeches (~ 600.000 tokens)
- transcripts of the German and Austrian parliaments (~ 500.000 tokens)
- parts of interviews printed in the magazine "Spiegel" (~ 400.000 tokens)
- transcripts of the TV program "Literarisches Quartett" (~ 450.000 tokens)
- other transcripts (~ 300.000 tokens)

Kernkorpus 20: 125.990.080 tokens; among them:
- literary texts (ca. 26%)
- journalistic texts (ca. 27%)
- scientific texts (ca. 22%)
- non-fictional texts (ca. 20%)
- transcribed spoken texts (ca. 5%)

DIE ZEIT: 417.422.714 tokens; all issues of the weekly magazine "ZEIT" from 1946 until 2009

Corpora

Corpora from the Web (*COW*). Freie Universität Berlin. Various corpora extracted from different country-specific web domains. Accessed *DECOW2011*, the corpus extracted from German (.de) domains, in its unshuffled beta version between Dec 2011 and Mar 2012. The corpus is not publicly available in this form. Further information regarding the project and publicly accessible versions of the corpora is available at *http://hpsg.fu-berlin.de/cow/*

Datenbank für Gesprochenes Deutsch (DGD). Institut für Deutsche Sprache (IDS), Mannheim. Available at *http://dgd.ids-mannheim.de*

Digitales Wörterbuch der deutschen Sprache (*DWDS*). Berlin-Brandenburgische Akademie der Wissenschaften. Available at *http://www.dwds.de/*

References

Alonso-Ovalle, Luis (2002) Arbitrary pronouns are not that indefinite. In C. Beyssade, R. Bok-Bennema, F. Drijkoningen, and P. Monachesi (eds.), *Romance Languages and Linguistic Theory 2000*. Amsterdam: John Benjamins Publishing Company, pp. 1-15.

Amaral, Patricia Matos, Craige Roberts, and E. Allyn Smith (2007) Review of 'The Logic of Conventional Implicature' by Chris Potts. *Linguistics and Philosophy* 30: 707-749.

Chierchia, Gennaro (1995) The variability of impersonal subjects. In E. Bach, E. Jelinek, A. Kratzer, and B.H. Partee (eds.), *Quantification in Natural Languages.* Dordrecht: Kluwer Academic Publishing, pp. 107-143.

Cinque, Guglielmo (1988) On "si" constructions and the theory of arb. *Linguistic Inquiry* 19: 521-581.

Condoravdi, Cleo (1989) Indefinite and generic pronouns. In E. Jane Fee, and K. Hunt (eds.), *Proceedings of the Eighth West Coast Conference on Formal Linguistics,* pp. 71-84.

De Cock, Barbara (this volume) Non-prototypicality of personal pronouns in typical registers? A cross-linguistic analysis. *Pragmatics* 26.3: 361-378.

Gregersen, Frans, and Torben Juel Jensen (this volume) How did 'man' become 'du', or did it? The pragmatics of generic second person pronouns in modern Danish spoken language. *Pragmatics* 26.3: 417-446.

Grice, Paul (1975) Logic and conversation. In P. Cole, and J. Morgan (eds.), *Syntax and Semantics 3: Pragmatics*. New York: Academic Press, pp. 41-58.

Grimberg, Mary Lou (1994) On Nunberg on indexicality and deixis. *UCL Working Papers in Linguistics* 6: 1-38.

Gruber, Bettina (2011) Indexical pronouns: Generic uses as clues to their structure. *Poznan Studies in Contemporary Linguistics* 47: 331-360.

Helmbrecht, Johannes (2015) A typology of non-prototypical uses of personal pronouns: Synchrony and diachrony. *Journal of Pragmatics* 88: 176-189.

Hugo Rojas, Evelyn (2011) Las formas de segunda persona singular como estrategias evidenciales [The forms of second singular person as evidential strategies]. *Revista de Linguistica Teorica y Aplicada Concepcion (Chile)* 49: 143-167.

Kamp, Hans (2008) Discourse structure and the structure of context. Ms., IMS Universität Stuttgart.

Kaplan, David (1989) Demonstratives. In J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*. Oxford: Oxford University Press, pp. 481-563.

Kitagawa, Chisato, and Adrienne Lehrer (1990) Impersonal uses of personal pronouns. *Journal of Pragmatics* 14: 739-759.

Kluge, Bettina (this volume) Misunderstanding or (deliberately) misinterpreting the second singular – how speakers politely deal with referential ambiguity. *Pragmatics* 26.3: 501-522.

Krifka, Manfred, Francis Jeffry Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia, and Godehard Link (1995) Genericity: An introduction. In G. Carlson, and J. Pelletier (eds.), *The Generic Book*. Chicago: University of Chicago Press, pp. 1-124.

Laberge, Suzanne, and Gillian Sankoff (1979) Anything *you* can do. In T. Givón, (ed.), *Discourse and syntax*. New York: Academic Press, pp. 419-40.

Malamud, Sophia A. (2006) *Semantics and pragmatics of arbitrariness*. Ph.D. dissertation, University of Pennsylvania.

Malamud, Sophia A. (2012) Impersonal indexicals: *One*, *you*, *man*, and *du*. *Journal of Comparative Germanic Linguistics* 15: 1-48.

Moltmann, Friederike (2006) Generic *one*, arbitrary PRO, and the first person. *Natural Language Semantics* 14: 257-281.

Moltmann, Friederike (2010) Relative truth and the first person. *Philosophical Studies* 150: 187-220.

Moltmann, Friederike (2012) Two kinds of first-person-oriented content. *Synthese* 184*: 157-177.

Nogué, Neus, and Òscar Bladas (this volume) "Que bé, tu!" ("That's great, you!"): Non-prototypical uses of the second person singular pronoun *tu* (you) in spoken Catalan. *Pragmatics* 26.3: 473-500.

Nunberg, Geoffrey (1993) Indexicality and deixis. *Linguistics and Philosophy* 16: 1-43.

Potts, Chris (2005) *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.

Siewierska, Anna (2004) *Person*. Cambridge: Cambridge University Press.

Tarenskeen, Sammie (2010) *From you to me (and back): The flexible meaning of the second person pronoun in Dutch*. Master's thesis, Radboud University Nijmegen.

Thurmair, Maria (1989) Modalpartikeln und ihre Kombinationen. *Linguistische Arbeiten 223*. Tübingen: Niemeyer.

Zifonun, Gisela (2000) Man lebt nur einmal. Morphosyntax und Semantik des Pronomens "man". *Deutsche Sprache* 28: 232-253.

Zimmermann, Malte (2011) Discourse particles. In C. Maienborn, K. von Heusinger, and P. Portner (eds.), *Handbook Semantics (Handbücher zur Sprach- und Kommunikationswissenschaft HSK 33.2)*. Berlin: Mouton de Gruyter, pp. 2011-2038.

Zobel, Sarah (2010) Non-standard uses of German 1st person singular pronouns. In K. Nakakoji, Y. Murakami, and E. McCready (eds.), *JSAI-isAI, LNAI 6284,* pp. 292-311.

Zobel, Sarah (2014) *Impersonally Interpreted Personal Pronouns*. Ph.D. dissertation, University of Göttingen. Available at http://hdl.handle.net/11858/00-1735-0000-0023-991B-2