

Annotation manual

Version: December 2019

Annotating morphosyntactic and semantic properties of the
linguistic context surrounding occurrences of dedicated
impersonal pronouns

Sarah Zobel

University of Oslo

sarah.zobel@iln.uio.no



This manual was developed as part of the project “Looking for the Impersonal Core – Impersonal Pronouns across Germanic languages” (IPG_CORE), which has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 842363.

Contents

1	Introduction	3
2	Theoretical background and motivation	4
2.1	Motivation behind the morphosyntactic and semantic categories	5
2.2	The state of affairs described by a clause: M1–M6 and S1	8
3	Annotation guidelines	10
3.1	Step 1: morphosyntactic properties of the containing sentence/clause	10
3.1.1	M1 – Sentence type	10
3.1.2	M2 – Co-occurrence of a modal verb	11
3.1.3	M3 – Occurrence inside a conditional	12
3.1.4	M4 – Temporal verbal form	13
3.1.5	M5 – Presence of a temporal adverbial	14
3.1.6	M6 – Presence of a spatial adverbial	14
3.2	Step 2: semantic properties of the containing clause	15
3.2.1	S1 – Situation type	15
3.2.2	S2 – Interpretation of an impersonally usable personal pronoun	18
3.2.3	S3 – Inclusion of the speaker/author	20
3.2.4	S4 – Particular cases: intuitions about who is meant	22
4	Possible conclusions for the range of interpretation and uses	23
	References	27
A	Annotation specifics and examples for Norwegian	28
B	Annotation specifics and examples for German	32
C	Recent classifications of the interpretations/uses of human impersonal pronouns	36

I gratefully acknowledge the help and input of my annotators while developing and testing this manual: Tonje Andersen, Judith Lauterbach, and Ingvild Røsok.

1 Introduction

The project IPG_CORE investigates the **possible interpretations of dedicated impersonal pronouns** in German, Norwegian, and Swedish. So, this annotation manual was developed with these three languages in mind. The manual is written in such a way that the general approach can be adapted to investigate other languages in a similar manner.

Dedicated impersonal pronouns (e.g., English *one*, German/Norwegian/Swedish *man*, or French *on*) are pronominal expressions that are primarily or exclusively used to express **impersonal meaning**, which is characterized as backgrounding the subject by not identifying or referring to a specific individual (see e.g., Cabredo Hofherr 2015, Gast and van der Auwera 2013). So, unlike personal pronouns (and other referential expressions), dedicated impersonal pronouns do not refer to individuals that have been introduced in the course of the conversation.¹

In the literature, **three different uses** for dedicated impersonal pronouns are traditionally distinguished, which are linked to **three different interpretations** (see a.o., Cinque 1988, Chierchia 1995, Cabredo Hofherr 2015), of which only the first two constitute impersonal meaning of the type described above.

- A **generic use**, in which the dedicated impersonal pronoun is described as contributing (roughly) the meaning “people in general”.

(1) **One** has to respect one’s parents. (*one* ≈ people in general)

- An **existential use**, in which the dedicated impersonal pronoun is described as contributing (roughly) the meaning “someone”.

(2) **Man** hat mir mein Fahrrad gestohlen. (German)
 ‘MAN stole my bike.’ (MAN ≈ someone)

- A **referential use**, in which the dedicated impersonal pronoun is described as contributing a 1st-person-pronoun-like meaning (i.e., either “I” or “we”).

(3) **On** doit encore faire les courses. (French)
 ‘ON still has to buy groceries.’ (ON ≈ we)

Cross-linguistically, not all of these uses are available for all dedicated impersonal pronouns:

- English *one* and Icelandic *maður* are said to only have the generic use (a.o., Cabredo Hofherr 2015, Fenger 2018).
- German *man* is said to only have the generic and existential uses (a.o., Cabredo Hofherr 2015, Zifonun 2000).
- French *on* and Swedish *man* are said to have all three uses (a.o., Fenger 2018, Egerland 2018).

¹They, therefore, also semantically contrast with personal pronouns that have *impersonal uses*. Personal pronouns have impersonal uses in addition to their primary referential uses (see Malamud 2006, 2012, 2013, Zobel 2014). The semantic closeness between impersonal uses of personal pronouns and dedicated impersonal pronouns is used in the annotation guide and is further addressed below.

The **goal of the annotation described in this manual** is to build an annotated collection of attested corpus examples for a particular dedicated impersonal pronoun of a given language. This data, on the one hand, is intended to provide an overview of how this pronoun is used (in the register/genre of the corpus) and, on the other hand, to help to determine which uses and interpretations are plausibly available for that dedicated impersonal pronoun.

Before turning to the annotation guidelines, the choice of morphosyntactic and semantic properties that will be annotated is motivated in Section 2 and the relevant background is provided. The annotation guidelines in Section 3 are then divided into two steps, which is also the recommended order for annotating new data and when training new annotators:

- Step 1: morphosyntactic annotation
- Step 2: semantic annotation

Section 4 discusses which conclusions can be drawn from the annotation results regarding the available interpretations for the dedicated impersonal pronoun under consideration. The details used for the annotation of Norwegian *man* and German *man* in connection with IPG_CORE are given in the Appendix.

The guidelines are illustrated using English examples containing the dedicated impersonal pronoun *one*, as well as German and Norwegian examples containing the German/Norwegian dedicated impersonal pronoun *man* if no fitting English example could be found (i.e., for episodic contexts). The English examples are taken from the web or constructed based on such examples or were taken from the *Oslo Multilingual Corpus*, which was accessed at the University of Oslo: <https://www.hf.uio.no/iln/om/organisasjon/tekstlab/>. The German examples are also either constructed or were taken from the corpora at the *Digitales Wörterbuch der Deutschen Sprache* (DWDS): <https://www.dwds.de/>. The Norwegian examples were taken from the main corpus study conducted for IPG_CORE using the *Leksikografisk bokmålskorpus* and a support study using the *Oslo Multilingual Corpus*, both of which were also accessed at the University of Oslo.

2 Theoretical background and motivation

The goal of the annotation described in this manual is to build an annotated collection of attested corpus examples for a particular dedicated impersonal pronoun of a given language, which may help to determine which uses and interpretations are plausibly available for that pronoun. The categories that I propose for annotation in Section 3 are based on morphosyntactic properties of the sentence/clause containing the dedicated impersonal pronoun, as well as semantic properties connected to the situation described by that clause and to the interpretation of the pronoun. The categories are introduced and discussed in Sections 2.1 and 2.2 below.

Crucially, neither the interpretation nor the use of the dedicated impersonal pronouns occurring in the corpus items are proposed as annotation criteria. This choice is motivated by the fact that the full range of potential impersonal interpretations for dedicated impersonal pronouns is far from clear and no principled annotation guidelines for them can be given. Recent cross-linguistic investigations of the possible interpretations and uses of *human impersonal pronouns* (a cover term for dedicated impersonal pronouns and impersonally interpreted personal pronouns) argue that it is not enough to distinguish only one generic use from one existential

use, as is standardly assumed (see the introduction). However, even though there is a general consensus that more than two interpretations and uses should be distinguished, the authors do not agree how many and which interpretations and uses should be distinguished instead (compare, e.g., Cabredo Hofherr 2006; Cabredo Hofherr 2015; Gast and van der Auwera 2013; Siewierska and Papastathi 2011). Furthermore, for some uses, there is no consensus whether they should be considered a type of generic or a type of existential use. For instance, Cabredo Hofherr (2015:11–12) conceptually links the interpretation of a human impersonal pronoun in the **corporate use**, see (4), to the generic use, while Gast and van der Auwera (2013:152–153) categorize it as an existential use.

- (4) a. **Ils** ont encore augmenté la TVA. (French)
 ‘They increased the taxes again.’
 b. **Man** hat wieder die Steuern angehoben. (German)
 ‘MAN increased the taxes again.’

The pronouns *ils*, *they*, or *man* in (4) intuitively contribute the relevant people in the government that are authorized to change the effective tax rates (see Cabredo Hofherr 2015:10). This intuitive interpretation is captured neither by the paraphrase traditionally suggested for the generic use (i.e., “people in general”) nor by that suggested for the existential use (i.e., “someone”). This shows that the uses that are distinguished traditionally cannot provide an adequate basis for annotating the interpretation/use of dedicated impersonal pronouns, and that identifying their possible readings via substitution with these paraphrases will not lead to a satisfying classification. As a result, it is currently not possible to give principled guidelines to annotate either the interpretation or use of a given dedicated impersonal pronoun.

In the following subsections, I motivate and discuss the morphosyntactic and semantic categories that are to be annotated. How these categories can help to investigate the available interpretations and uses of dedicated impersonal pronouns will be discussed in Section 4.

2.1 Motivation behind the morphosyntactic and semantic categories

The main motivation behind the choice of categories to be annotated is their connection to the interpretation of human impersonal pronouns. While the full range of possible interpretations/uses and how they are connected are still open questions, the discussions in the literature resulted in a list of morphosyntactic (M) and semantic (S) properties of the sentences/clauses containing the pronouns that affect how the pronouns are interpreted. The corpus data will be annotated with these properties.

The bulk of the annotation categories concern the **type of state of affairs described by the clause** containing the dedicated impersonal pronoun:

- M1 sentence type
- M2 co-occurrence of a modal
- M3 occurrence of the dedicated impersonal pronoun in a conditional
- M4 temporal verbal form of the finite verb in the clause
- M5 presence of a temporal adverbial
- M6 presence of a spatial adverbial
- S1 situation type described by the verbal material in the clause

This group of categories aims at determining whether the clause containing the dedicated impersonal pronoun describes a **particular state of affairs** that actually happened or is currently happening, or whether the clause describes a **necessity, possibility, rule or regularity**. How the different categories listed above help in determining the type of state of affairs is discussed in Section 2.2 below.

The reason for focusing on this aspect of the clause is that it correlates with, or maybe in some cases even determines, the interpretation of a dedicated impersonal pronoun. The relevant observation is that the generic use and the existential use are connected to which type of state of affairs the sentence describes (e.g., Cinque 1988, Chierchia 1995): if the main predicate in the clause describes a rule or regularity about the world, a dedicated impersonal pronoun will occur in its generic use; if the main predicate in the clause describes a particular event, however, a dedicated impersonal pronoun will occur in its existential use. This is illustrated for German *man* in (5) and (6).² The types of predicates occurring in (5) and (6) are called “**generic/habitual**” and “**episodic**”, respectively.

- (5) **Generic/habitual** (German)
- a. **Maria** isst am Wochenende Eier zum Frühstück.
 ‘On the weekend, Maria eats eggs for breakfast.’
 - b. **Man** isst am Wochenende Eier zum Frühstück.
 ‘On the weekend, MAN eats eggs for breakfast.’ (MAN ≈ people in general)

The predicate *eat eggs for breakfast on the weekend* describes a regularity of certain eating events happening regularly each weekend. In (5-b), the dedicated impersonal pronoun *man* in subject position can only be understood as appearing in its generic use.

- (6) **Episodic** (German)
- a. **Maria** rief gestern Abend die Polizei.
 ‘Yesterday evening, Maria called the police.’
 - b. **Man** rief gestern Abend die Polizei.
 ‘Yesterday evening, MAN called the police.’ (MAN ≈ someone)

The predicate *call the police yesterday evening* in (6) describes a particular event of calling the police. In (6-b), *man* can only be understood as appearing in its existential use, contributing a meaning roughly paraphrasable as “someone”.

As stated above, this correspondence does not fully determine the interpretation or use for a dedicated impersonal pronoun, given that dedicated impersonal pronouns have been shown to have more than just two uses. Nevertheless, the correlation between the state of affairs described by the clause (incl. the type of predicate occurring in the clause) and the interpretation and use of the dedicated impersonal pronoun still holds and provides a valuable piece of evidence to map out the possible uses of a given dedicated impersonal pronoun.³

The remaining semantic categories concern the interpretation of the dedicated impersonal

²Corresponding examples could be given for other dedicated impersonal uses allowing both types of uses, e.g., French *on* or Italian *si* (see e.g., Cabredo Hofherr 2015, Cinque 1988).

³The correlation between context and use seems to break down for different impersonally usable personal pronouns. Greek 3rd person plural and Spanish 2nd person singular impersonally used personal pronouns seem to have existential uses in generic sentence contexts and generic uses in episodic sentence contexts (see Condoravdi 1989; Alonso-Ovalle 2002; Cabredo Hofherr 2015).

pronoun directly:

- S2 interpretation of an impersonally usable personal pronoun in the same clause
- S3 inclusion of the author/speaker in the group of individuals described by the dedicated impersonal pronoun
- S4 specification of the group of individuals described by the dedicated impersonal pronoun in the context

These three categories target different aspects of the interpretation of a dedicated impersonal pronoun and are based on different correlations observed in the literature.

The **category S2** concerns the available interpretation of an impersonally usable singular personal pronoun if it is substituted for the dedicated impersonal pronoun in the clause. Impersonally used personal pronouns only have a generic interpretation in addition to their regular, referential interpretation (see e.g., Malamud 2006, 2012, 2013, Zobel 2014). The generic interpretation occurs only with generic/habitual predicates that express a necessity, possibility, rule or regularity. The same contexts correlate with a generic use of dedicated impersonal pronouns. This is illustrated in (7) for English *one* and *you*.

- (7) a. **You** have to respect **your** parents. (both: people in general)
 b. **One** has to respect **one's** parents.

The generic use of an impersonally usable singular personal pronoun may, however, be more restricted than that of a dedicated impersonal pronoun (see Zobel 2014 for German 2nd person singular *du*).

The **category S3** concerns the idea of speaker inclusion, as discussed by Kratzer (1997) and Moltmann (2006, 2010, 2012). Kratzer (1997) observes for German *man* that the group of individuals that are described sometimes contains the speaker and sometimes does not, see (8). She proposes that speaker inclusion vs. speaker exclusion distinguishes two different uses of German *man*.

- (8) a. Wir spielten Karten, weil **man** ja hier nichts Besseres zu tun hatte.
 'We played cards because MAN had nothing better to do here.' (includes speaker)
 b. Im Ministerium wusste **man** ganz genau über mich Bescheid.
 'In the ministry, MAN knew everything about me.' (excludes speaker)

Moltmann (2006, 2010, 2012) argues that speaker inclusion does not necessarily mean that the group of people described by the dedicated impersonal pronoun actually contains the speaker. The speaker may also include themselves in that group of people by signalling that they *identify* with the individuals in that group.⁴ English *one*, Moltmann argues, always involves speaker inclusion, which can be observed by contrasting *one* with the bare plural noun *people*, see (9).

- (9) a. Sometimes **one** receives strange advertisements in the mail.
 b. Sometimes **people** receive strange advertisements in the mail.
 (Moltmann 2006:268)

The speaker in (9-a) identifies with those individuals that receive strange advertisements. This is the case even if the speaker never actually received strange advertisements themselves. So,

⁴A similar observation is reported for German *man* in Zifonun 2000.

speaker inclusion vs. speaker exclusion may be connected to certain interpretations and uses of dedicated impersonal pronouns and may distinguish them from other interpretations and uses.

Lastly, the **category S4** concerns the identity of the group of individuals that the dedicated impersonal pronoun talks about. The annotation guideline restricts this category to corpus items that feature a sentence that plausibly does not describe a rule or regularity since in these cases, the group of individuals might be identifiable more readily for a corporate use than for an existential use (compare (4-b) to (6-b)).

2.2 The state of affairs described by a clause: M1–M6 and S1

The following list describes how the categories M1–M6 and S1 help in determining **the type of state of affairs** described by the clause.

S1 Situation type: Following the literature on genericity and the automatic detection of generic sentences (e.g., [Carlson 2005](#); [Friedrich et al. 2015a,b](#); [Friedrich 2017](#)), the annotation distinguishes three situation types: *habitual*, *episodic*, and *static*. This three-way split distinguishes clauses describing events from those describing states (= *static*), and among the clauses describing events, it distinguishes those describing a particular event (= *episodic*) from those describing an event regularity (= *habitual*).

The distinction between episodic and habitual makes sure that the annotation provides the necessary basis to take advantage of and investigate the correlation between the type of state of affairs and the interpretation/use of a dedicated impersonal pronoun described in Section 2.1.

The sentences describing *static* situations, however, may contain both particular states, as well as necessities, possibilities, and state regularities. Hence, they do not form a homogeneous class and need to be distinguished further by other means. This is accomplished through the annotation of M2 and M3.

M2 Co-occurrence of a modal: The annotation captures any modal verb occurring in the same clause as the dedicated impersonal pronoun. Hence, this annotation will distinguish all clauses that describe modal-verb-based static situations from those that describe static situations without the presence of a modal. This distinction is meaningful since clauses that contain modal verbs do not describe particular actual situations, but possibilities and necessities (see e.g., [Kratzer 2012](#)).

M3 Occurrence in a conditional: The annotation captures whether the clause containing the dedicated impersonal pronoun occurs inside of a conditional structure, or not. This means that the annotation identifies conditional-based static situations, which are grouped with modal-verb-based static situations due to their shared intensional semantics (see e.g., [Kratzer 2012](#)).⁵

M1 Sentence type: The annotation distinguishes the three major sentence types: *declarative*, *interrogative*, and *imperative*.

Identifying a declarative sentence does not help in indentifying the type of state of affairs further since declarative sentences can describe any state of affairs depending on the

⁵One exception to this are clauses that are embedded under factive predicates in a clause that is part of a conditional structure. Factive predicates are, for instance, *know* and *be surprised*; see [Karttunen 2016](#).

linguistic material in the sentence (e.g., presence vs. absence of a modal verb or a conditional structure). In contrast, interrogative sentences and imperatives arguably describe the same type of state of affairs as clauses containing modal verbs (see e.g., [Groenendijk and Stokhof 1984](#); [Kaufmann 2012](#)).⁶

M4 Temporal verbal form: The annotation identifies the temporal verbal form of the finite verb of the clause containing the dedicated impersonal pronoun. There are various correlations between the temporal verbal form of the finite verb and the situation type described by a clause, although there is no strict rule linking a particular verbal form with a particular situation type (see [Krifka et al. 1995](#)).

- In general, habitual situations are mostly described with a present temporal form. Past temporal forms can in principle also describe habitual situations, in particular, they describe regularities that no longer hold.
- In English, sentences with a progressive verbal form can never describe a habitual situation.
- For those languages that distinguish perfective and imperfective past temporal forms (e.g., Romance languages), only the imperfective past temporal forms can be used to describe habitual situations.

In addition, the temporal verbal form together with the situation type determined in S1, allow us to differentiate subtypes of episodic situations. Across languages, human impersonal pronouns seem to differ with respect to the availability of an existential use in different types of episodic sentence contexts. For instance, German *man* is not usable with a meaning close to “someone” in a clause that describes a particular event happening at the time of utterance; in contrast, French *on* can occur in this context, see (10) ([Cabredo Hofherr 2015:18](#)).

- (10) a. #**Man** klopft an der Tür. (German)
 b. **On** frappe à la porte. (French)
 ‘IMP is knocking on the door.’ (IMP ≈ someone)

M5 Presence of a temporal adverbial: The annotation identifies any type of temporal adverbial, which temporally situates the state of affairs described in the sentence. The presence of a temporal adverbial can be used as a heuristic towards judging whether a sentence describes an episodic situation: episodic situations happen at a particular time, which can be specified via a temporal adverbial.

M6 Presence of a spatial adverbial: The annotation identifies any type of spatial adverbial, both locatives and directives. Locative adverbials that situate the eventuality described in the sentence at a specific, restricted location hint at the situation being episodic rather than habitual. In contrast, locative adverbials identifying general areas have been argued to distinguish different types of generic uses: for some impersonally usable personal pronouns (e.g., 3rd person plural pronouns), a generic use is only available if a locative expression co-occurs in the clause. No similar necessity obtains for the dedicated impersonal pronouns English *one*, German *man*, or French *on* ([Cabredo Hofherr 2015:15](#)).

⁶One exception to this are clauses that are embedded under factive predicates inside interrogatives and imperatives may be veridical.

3 Annotation guidelines

This section generalizes the instructions to the annotators for the corpus studies conducted in the course of Workpackage 1 of IPG_CORE and specifies how the guidelines need to be adapted to other languages.

Choice of annotators: Since the semantic annotation in the second step is partly based on the intuition of the annotators, the annotators need to be **native speakers of the language** that is investigated. The guidelines assume only **basic knowledge of grammatical terminology**. Nevertheless, all annotation tasks should be discussed with the annotators beforehand to prevent misunderstandings.

General rules:

- If more than one dedicated impersonal pronoun occurs inside a multiclausal sentence, **only the first occurrence and its clause** are relevant for the annotation.
- For all clause-related annotation properties, the clause containing the dedicated impersonal pronoun should be **treated like a root clause even if it is embedded**.

Preparation of the corpus items: Depending on their level of linguistic training, it might be helpful for the annotators to **identify the relevant occurrence** of the dedicated impersonal pronoun beforehand and to **clarify the clause boundaries** for them. This could either be done (semi-)automatically or while the corpus sample is checked for false hits and fragments.

3.1 Step 1: morphosyntactic properties of the containing sentence/clause

The following six properties are annotated for the sentence/clause containing the first (or only) occurrence of the dedicated impersonal pronoun:

- M1 sentence type
- M2 co-occurrence of a modal
- M3 occurrence of the dedicated impersonal pronoun in a conditional
- M4 temporal verbal form of the finite verb
- M5 presence of a temporal adverbial
- M6 presence of a spatial adverbial

Given the theoretical background discussed in Section 2, the first four properties provide the most information about the meaning and use of a co-occurring dedicated personal pronoun. So, if time is scarce, the **annotation can be restricted to the first four properties** (see Section 4).

3.1.1 M1 – Sentence type

For the annotation of sentence type, the three main sentence types – *declaratives*, *interrogatives*, and *imperatives* – are distinguished.⁷

(11) SENTENCE TYPE: decl (declarative), int (interrogative), imp (imperative)

⁷For languages with more syntactically encoded sentence types, the possible annotation levels need to be extended with the available minor sentence types.

The sentence types are described very minimally below. Any language-specific information about how these different sentence types are encoded in the language of interest should be determined in advance.

DECL Declaratives are prototypically used to make statements. They are usually ended orthographically with a period (.) – although, especially with more colloquial corpus data, sometimes declarative sentences occur with a question mark (?), an exclamation mark (!), or both (!?) in case the author wants to express that they are unsure or surprised about the content.

(12) ONE does not simply walk into Mordor. (annotation: decl)

INT Interrogatives are prototypically used to ask questions. They are usually ended orthographically with a question mark (?).

(13) What does ONE do with celery? (annotation: int)

IMP Imperatives are prototypically used to communicate a demand or order that is directed at the addressee. They are usually ended orthographically with an exclamation mark (!) and do not contain an overt subject.

(14) Read this book!

Because imperatives usually don't have overt subjects and are directed at the addressee, occurrences of dedicated impersonal pronouns inside imperatives are expected to be rare; especially for pronouns that can only occur in subject position, like German/Norwegian/Swedish *man* (see a.o., [Cabredo Hofherr 2015](#), [Fenger 2018](#)). They may, however, occur embedded:

(15) Gib zu, dass man da nichts machen kann! (German)
 'Admit that MAN can't do anything about that!' (annotation: imp)

3.1.2 M2 – Co-occurrence of a modal verb

Does a modal verb occur in the clause containing the dedicated impersonal pronoun?

The exact inventory of modal verbs differs from language to language. So, whichever expressions will be identified here, have to be determined in advance.

(16) CO-OCCURRING MODAL: note down the exact form of the modal or use 0 ("no modal")

Since the relevant domain for annotation is the clause containing the dedicated impersonal pronoun, it is advisable to first identify the relevant clause before doing the annotation. In the examples below, the clauses are underlined (with clauses embedded in those clauses appearing dashed).

English has the modal verbs *can, may, will, shall,* and their past tense forms *could, might, would,* and *should,* as well as the “semi-modal” *ought*.

- (17) I have no qualms, contrary to what ONE **might** expect. (annotation: *might*)
- (18) We’d often talked about it, how ONE **ought** to conduct oneself when a relationship was over. (annotation: *ought*)

The annotation is “0” if there is no modal, or if there is a modal, it does not occur in the same clause as the relevant occurrence of the dedicated impersonal pronoun.

- (19) ONE does not simply walk into Mordor. (annotation: 0)
- (20) If ONE attaches importance to heightened realism, one **can** apply a little blood colour at this point. (annotation: 0)

3.1.3 M3 – Occurrence inside a conditional

Does the dedicated impersonal pronoun occur inside a conditional statement – either inside the conditional clause or the main clause (or both)?

- (21) IN CONDITIONAL: 0 (= no), 1 (= yes)

Conditional statements express that something is / would be the case if something else is/were the case (the condition):

- (22) If it’s raining, you need an umbrella.

So, a conditional structure is made up from two clauses (that can be complex themselves): the clause that expresses the condition (usually called “conditional clause”) and the clause that expresses what happens in case the condition occurs.

Languages differ in how they express conditional statements, that is, which expressions can introduce a conditional clause. So, the inventory for the language that is to be annotated needs to be determined in advance.

If a sentence expresses a conditional statement, the first step in the annotation is to determine which clauses make up that statement. If it is clear which material in the sentence makes up the conditional statement, it is easier to determine whether the impersonal pronoun occurs inside one of the clauses making up that conditional statement, or not. In the examples below, the clauses are underlined.

English has the conditional introductory expressions *if* and *when(ever)*.

- (23) It’s easy to get a bit clumsy, isn’t it, especially **if** ONE isn’t being observed. (annotation: 1)
- (24) **When** all is said and done, ONE doesn’t need to hide all that much from a gecko. (annotation: 1)

For some types of conditional statements, it is also possible to have a conditional clause without

an introductory expression. In these cases, the conditional clause starts with the finite verb.

- (25) When the agent drives a car knowing a crash might ensue (such that, had ONE known of the crash in advance, one should not have started driving), the pedestrian also faces a choice. (annotation: 1)

The annotation is “0” if the corpus item does not contain a conditional clause, or the dedicated impersonal pronoun does not occur inside either of the clauses that are conditionally linked.

- (26) I have no qualms, contrary to what ONE might expect. (annotation: 0)
- (27) But she was very discreet, never mentioned him if I didn't directly ask; ONE could have believed I, not she, was in love with him. (annotation: 0)

3.1.4 M4 – Temporal verbal form

What is the temporal verbal form of the finite verb in the clause containing the dedicated impersonal pronoun?

The annotation of TEMPORAL VERBAL FORM concerns **the morphological form of the finite verb** in the clause containing the dedicated impersonal pronoun. In a multiclausal sentence, the temporal forms of the verbs in other clauses are not relevant. So, it is advisable to identify the relevant clause beforehand. In the examples below, the clauses are underlined (with clauses embedded in those clauses appearing dashed).

Since languages differ in their inventory of possible temporal verbal forms, the levels of annotation will depend on the language that is annotated. The following list provides the traditional temporal verbal forms found in **English**.⁸

- (28) TEMPORAL VERBAL FORM:
pres (present), pres.prog (present progressive), past, past.prog (past progressive), perf (perfect), plup (pluperfect) fut (future) fut.perf (future perfect)
- (29) It's easy to get a bit clumsy, isn't it, especially if ONE isn't being observed. (annotation: pres.prog)
- (30) When all is said and done, ONE doesn't need to hide all that much from a gecko. (annotation: pres)
- (31) When the agent drives a car knowing a crash might ensue (such that, had ONE known of the crash in advance, one should not have started driving), the pedestrian also faces a choice. (annotation: plup)

Note that if a **modal verb** occurs in the sentence, the temporal verbal form connected to the modal verb is annotated.

- (32) I have no qualms, contrary to what ONE might expect. (annotation: past)

⁸For a language with special subjunctive/Konjunktiv verb forms, these can be annotated together with the temporal verbal forms. For instance, for German, Konjunktiv I and Konjunktiv II forms can be added to the list of possible verbal forms.

Since the temporal verbal form of a modal does not necessarily correspond to its temporal interpretation (e.g., the past form *might* expresses that something is currently an epistemic possibility), these cases may need to be discussed with the annotators.

3.1.5 M5 – Presence of a temporal adverbial

Is there an adverbial expressions in the containing sentence that determines a time or temporal interval for the situation that is described in the clause containing the dedicated impersonal pronoun?

(33) TEMPORAL ADVERBIAL: 0 (= no), 1 (= yes)

It does not matter whether the temporal adverbial names either a specific point in time or temporal interval or a broad, general temporal interval. The only thing that is important is, whether the clause contains any form of temporally specifying expression.

Examples of **English** temporal adverbials are *today, tomorrow, on Sunday, every week, in 2018, last year, ...*

(34) In New Orleans, **around the turn of last century**, ONE could find strings in a marching band. (annotation: 1)

(35) **Today**, one doesn't have to be a member of any political movement. (annotation: 1)

A corpus item is annotated with “0” if the sentence containing the dedicated impersonal pronoun does not also contain a temporal adverbial.

3.1.6 M6 – Presence of a spatial adverbial

Is there are an adverbial expressions in the containing sentence that specifies the place for the situation that is described in the clause containing the dedicated impersonal pronoun?

(36) SPATIAL ADVERBIAL: 0(= no), 1 (= yes)

The place can be a small, specific location or a general localization in an area, country, or on a continent. What is important, though, is that the entire situation (and action) is localized by the adverbial and not just an individual that is part of the situation/action (see [Maienborn 2001](#)).

Examples of **English** spatial adverbials are *here, there, in the UK, next to the church, ...*

(37) ONE doesn't eat fox **in England**? (annotation: 1)

(38) But **in Norway** ONE can never be sure, even in mid-July or August, that it will equally warm and sunny every day. (annotation: 1)

A corpus item is annotated with “0” if it does not contain any spatial adverbial or if the spatial adverbial is not used to localize the situation.

3.2 Step 2: semantic properties of the containing clause

The following four properties are annotated for the sentence/clause containing the first (or only) occurrence of the dedicated impersonal pronoun:

- S1 situation type
- S2 interpretation of an impersonally usable personal pronoun
- S3 inclusion of speaker/author
- S4 for particular cases: intuitions about who is meant

Of these four properties, only situation type can be strictly determined. Both “interpretation of an impersonally usable personal pronoun” and “inclusion of speaker/author” depend on the annotators’ intuition and how they interpret the sentence in the given context. This is also true—and even more so—for the fourth property. So, if time is scarce, the **annotation can be restricted to the first two properties** (see Section 4).

3.2.1 S1 – Situation type

Which type of situation is described by the clause containing the dedicated impersonal pronoun?

This annotation property is adopted from the work by the “Situation Entities” project at the Universität des Saarlandes (Friedrich et al. 2015b, Friedrich 2017). The annotation property SITUATION TYPE below corresponds to their annotation property “Habituality”, which is described in detail in their annotation manual (see Friedrich et al. 2015a:22–27).

The annotation guide of the “Situation Entities” project uses the traditional episodic/ habitual distinction (Carlson 2005), but extends it with one further type of situation, static. So, it distinguishes **three types of situations**:

(39) SITUATION TYPE: epi (episodic), hab (habitual), stat (static)

The following descriptions of the three situation types are short summaries of the descriptions in Friedrich et al. 2015a.

EPI Episodic situations involve a particular event or sum of events with physical or mental activities. A clause describes an episodic situation if the verb phrase describes such an event or sum of events:

- (40) Peter is eating a cookie. (single eating-event)
- (41) My bike was stolen. (single stealing-event)
- (42) At 10am, the students of this school joined the demonstration. (sum of joining-events)

From a clause describing an episodic situation, we learn that the particular event (or sum of particular events) happens/happened once.

Test suggested by Friedrich et al: Can a temporal adverbial describing bounded span of time be added that specifies when the event happens/happened?

⇒ (42) already contains such a temporal adverbial (i.e., *at 10am*)

⇒ to (40) and (41), we can add *now* and *last night*

HAB A clause describes a **habitual situation** if it describes a regularity for a type of event with physical or mental actions:

(43) In Spain, one eats dinner late. (regularity of eating-events)

(44) Mary swims. (regularity of swimming-events)

(45) Glass breaks easily. (regularity of breaking-events)

Test suggested by Friedrich et al: Can *typically* or *usually* be added to the clause without a significant change in meaning?

⇒ To (43) and (44), *usually* can be added.

⇒ (45) allows the addition of *typically*

Note: the regularity does not have to comprise regular occurrences of a type of event for a particular individual; the event that occurs regularly or typically can involve (possibly different) members of the same class of individuals, as in (46).

(46) Spiders die in autumn after producing an egg sac.

STAT The situation described by a clause is a **static situation** if the verb describes a state or property.⁹

(47) In Norway, children are well respected. (being-respected state)

(48) Mary owns three cats. (owning state)

If **modals** or **sentence negation** (*not*, *never*) combine with verbs describing events with physical or mental actions, they stativize. That is, in these cases, the clauses describe states of having an obligation or the possibility to do something (for the modals) or the state that a potential event did not occur (for negation).

(49) One **has to** respect one's parents. (obligation to act)

(50) Mary **can** swim. (possibility to act)

(51) John did **not** win the lottery. (no potential winning)

⁹Friedrich et al. 2015b list the following English verbs that describe states/properties as examples: *be*, *have*, *like*, *live*, *hate*, *prefer*, *mean*, *consist of*, *mind*, *recognize*, *seem*, *prefer*, *doubt*, *need*, *belong*, *believe*, *cost*, *get*, *impress*, *know*, *taste*, *love*, *own*.

3.2.2 S2 – Interpretation of an impersonally usable personal pronoun

If the dedicated impersonal pronoun is substituted with an impersonally usable personal pronoun, does the personal pronoun only have its regular reference?

- (56) PERSONAL PRONOUN: ref (regular reference only), imp (impersonal interpretation possible)

To make this more explicit: the impersonally usable personal pronoun I have in mind for this property is the second person singular pronoun (e.g., English *you*, German/Norwegian/Swedish *du*, French *tu*), which cross-linguistically and across many different language families allows for an impersonal use (see Siewierska 2011). Other personal pronouns with an impersonal use could be used, as well, provided that the dedicated impersonal pronoun does not have a referential use in which it has a comparable interpretation to the personal pronoun in its primary, non-impersonal use (i.e., for some languages, 3rd person plural pronouns cannot be used).¹⁰ The following description assumes the use of an impersonally usable **second person singular pronoun**.

If a second person singular pronoun has its regular reference, it strictly refers to a single addressee. In its impersonal use, a second person singular pronoun can be used to talk about “people (of a certain kind) in general”. This is illustrated for **English *you*** in (57).

- (57) a. But she was very discreet, never mentioned him if I didn’t directly ask; **you**’d have believed I, not she, was in love with him.
 b. **You** seldom see dogs fighting, a dog is content with advertising his strength and his anger.
 c. Besides, he thought **you** should complete what **you** set out to do.

For this annotation, the dedicated impersonal pronoun is substituted with the impersonally usable personal pronoun and the possible interpretations of the personal pronoun are determined:

- *Matching impersonal use is available*: imp (impersonal)

In case the personal pronoun can be interpreted in its impersonal use, the resulting statement should also be more or less the same statement as the original statement with the dedicated impersonal pronoun.

- *Only reference to the addressee possible*: ref (referential)

Dedicated impersonal pronouns never refer only to the addressee, so if the second person singular pronoun can only refer to the addressee after the substitution, the resulting sentence will be a different statement.

Application of the test for English:

- (58) a. When all is said and done, **ONE** doesn’t need to hide all that much from a gecko.

¹⁰What if such a pronoun does not exist in the language of interest? Alternatively, it can be tested which interpretation arises for nominal expressions that can occur inside generic sentences of that language (e.g., bare plurals). The specifics of this alternative will have to be worked out further, though.

- b. When all is said and done, **you** don't need to hide all that much from a gecko.
⇒ can have the same general interpretation as the sentence with *one*
(annotation: imp)

Since English *one* is said to only ever occur in generic sentences (e.g., Fenger 2018, Malamud 2012, Moltmann 2012), it is in principle always substitutable by impersonally used *you*. Hence, the annotation with “ref” cannot be illustrated with English *one* and is, therefore, illustrated using a German and a Norwegian example.

(59) (German)

- a. **Man** sprach gestern in der Oper viel von einem Duell zwischen dem Gatten einer berühmten Sängerin und Hrn. G.
'Yesterday at the opera, MAN talked a lot about a duel between the husband of a famous singer and Mr. G.'
- b. **Du** sprachst gestern in der Oper viel von einem Duell zwischen dem Gatten einer berühmten Sängerin und Hrn. G.
⇒ *du* can only refer to the addressee
⇒ the sentence has a different meaning from the original (annotation: ref)

- (60) (Norwegian)
- a. Deretter satte **man** ut reker i bur i og rundt et oppdrettsanlegg i Rogaland.
'After that, MAN set out shrimps in cages in and around a fish farm in Rogaland.'
 - b. Deretter satte **du** ut reker i bur i og rundt et oppdrettsanlegg i Rogaland.
⇒ *du* can only refer to the addressee
⇒ the sentence has a different meaning from the original (annotation: ref)

Important:

- When the possible interpretations of the personal pronoun are checked, the personal pronoun must not be stressed (see Gruber 2011, Zobel 2014).
- The “impersonal use” of a second person singular pronoun with the interpretation “people in general” does not exclude that the sentence is true for the addressee. That is, the addressee can be one of those “people in general”. The question is whether the sentence talks only about the addressee and no one else.

3.2.3 S3 – Inclusion of the speaker/author

Is the statement formulated in such a way that the speaker/author is definitely included or excluded from the group of individuals that the dedicated impersonal pronoun talks about?

- (61) SPEAKER INCLUSION: nec (necessarily/definitely included), pot (potentially included), no (not included / definitely excluded)

This annotation should be done after SITUATION TYPE and PERSONAL PRONOUN have been determined because the annotations lead to **different cases**:

Case 1) SITUATION TYPE is “habitual” or “stative (+ modal or conditional)” and PERSONAL PRONOUN is “impersonal”

In this case, the sentence containing the dedicated impersonal pronoun plausibly expresses a rule/regularity. For the dedicated impersonal pronoun in such a clause, the question can be precisified to:

Is the rule/regularity presented in such a way that it is valid for the speaker/author or would be valid for him if he had the *relevant properties*?

The *relevant properties* are determined by the rule/regularity: each rule/regularity links two properties (see 3.2.1) and says “if the first property applies then the second property (normally) applies, as well”. The “first property” is the “relevant property”.

- (62) One has to respect one’s parents.
first property: being human
second property: having to respect the parents

(63) In Norway, children sleep in their strollers outdoors.

first property: being a child in Norway

second property: sleeping in strollers outdoors

For rules/regularities, the annotation runs as follows:

NEC This applies if the author has the relevant property and, therefore, communicates that the rule/regularity is also true for them.

⇒ The author is human.

⇒ If we read (62), we assume that the author would agree that he has to respect his parents.

⇒ (62) is annotated as “nec”

POT This applies if the rule/regularity is presented in such a way that we can assume that it would be valid for the author if they had the relevant properties.

⇒ If we read (63), we assume that the author is not a child.

⇒ However, the rule would apply to the author if they were a child in Norway (i.e., they would sleep in their strollers) – there is no reason to assume otherwise.

⇒ (63) is annotated as “pot”

NO This applies if the author/speaker makes it explicit that the rule/regularity does not hold for them even if they have/had the relevant property.

Case 2) SITUATION TYPE is “episodic”, “habitual”, or “static” (+ modal or conditional) and PERSONAL PRONOUN is “referential”

In this case, the question can be precisified as follows:

Are the individuals that the dedicated impersonal pronoun is used to talk about “the group of the author” or some other group? In other words: is *man* interpreted more like *they* or more like *we*?

Given the restrictions of English *one* (see Section 3.2.2), this case is expected to never arise for it. Hence, the annotation is illustrated with German examples.

NEC the dedicated impersonal pronoun contributes the “group of the speaker” ($\approx we$).

(64) **Man** hörte, da wir alle unwillkürlich den Atem anhielten, mit einem mal das Ticken der Uhr, die man zur Feststellung der Zugzeit auf den Tisch gelegt hatte.

‘MAN suddenly heard the ticking of the clock that had been put on the table to measure the time of each turn, because we all unintentionally held our breaths.’

⇒ the narrator and the people around him heard the ticking $\approx we$

POT the contribution of the dedicated impersonal pronoun could be either the “group of the speaker” or some other group.

- (65) **Man** sprach gestern in der Oper viel von einem Duell zwischen dem Gatten einer berühmten Sängerin und Hrn. G.
 ‘Yesterday at the opera, MAN talked a lot about a duel between the husband of a famous singer and Mr. G.’
 ⇒ this could be the writer and the people around him at the opera (\approx *we*), or just the people at the opera (\approx *they*)

NO the dedicated impersonal pronoun is used to talk about a group excluding the speaker (\approx *they*).

- (66) **Man** tat uns nichts – **man** stellte uns nur in das vollkommene Nichts.
 ‘MAN didn’t harm us – MAN only put us into complete nothingness.’
 ⇒ another individual or group of individuals not containing the speaker or his group (\approx *they*)

- (67) **Man** mußte mich verwundet haben, oder ich hatte mich selbst an der Hand verletzt.
 ‘MAN must have injured me, or I had injured my own hand.’
 ⇒ another individual or group of individuals not containing the speaker (\approx *they*)

Case 3) SITUATION TYPE “static” (without a modal or conditional) and combinations not covered by Cases 1 & 2

For these examples, it has to be decided on a case by case basis, whether they express a rule/regularity for certain people (“people of a certain kind in general”) or information about a particular group of individuals (“they” or “we”):

- rule/regularity → **Case 1**
- single, particular situation → **Case 2**

3.2.4 S4 – Particular cases: intuitions about who is meant

Looking at the preceding context and the material in the sentence containing the dedicated impersonal pronoun, which individuals/group of individuals is the pronoun used to talk about?

- (68) **CONTEXTUAL SPECIFICATION:** note down who is meant if it can be determined from the context; otherwise use “0” for “no specification”

This property is helpful if the person/people working with the annotated data is not a native speaker of the language that is investigated. The individuals/groups that are identified by the annotators are based completely on their personal interpretation of the corpus item in its context and therefore has no

This annotation should be done after SITUATION TYPE and PERSONAL PRONOUN have been determined. It could in principle be done for all corpus items, but for these guidelines only those corpus items are looked at for which the other annotations could suggest that we are

dealing with a particular sentence: corpus items that have the annotation “episodic” or “static” (without a modal or conditional) for SITUATION TYPE and/or the annotation “referential” for PERSONAL PRONOUN.

Specification in the context and/or sentence is illustrated with a shortened version of an attested example of German *man* in (69), which talks about an influential group of artists.

- (69) Die ersten Mitglieder waren X, Y und Z, in dessen Atelier **man** bevorzugt zusammenkam.
 ‘The first members were X, Y, and Z, in whose workshop MAN preferred to meet.’

In example (69), the relative clause “in whose workshop...” describes a habitual situation (“habitual”), but second person singular *du* can only refer to the addressee (“referential”). The sentence provides information about the particular individuals that *man* talks about via the context and the subject of the main clause (“the first members”): they are (a subset of) members of that group of artists that is specified in the context and via the subject. The item will be annotated with this information.

This contrasts with cases like (70), where no contextual information is provided. The clause describes a particular event (a setting-the-university-on-fire) and is, therefore, “episodic”, and second person singular *du* can only refer to the addressee (“referential”). The clause itself does not provide any restricting or identifying information on the particular individuals who set the university on fire:

- (70) **Man** hat gestern die Uni angezündet.
 ‘Yesterday, MAN set the university on fire.’

Therefore, this corpus item will be annotated with “0”.

4 Possible conclusions for the range of interpretation and uses

In this section, the possible conclusions are discussed that can be drawn from the annotated data regarding the range of interpretation and uses available for the dedicated impersonal pronoun under consideration.

The most important point to keep in mind is that the data generated from a corpus sample that has been annotated according to this manual can only be used to **support claims of existence for predefined interpretations/uses**. Such a corpus sample can never conclusively show that an interpretation or use is unavailable or impossible.

Given a set of predefined interpretations/uses, the expected values for the annotated properties can be checked against the data generated by the annotation in order to confirm the existence of these interpretations/uses. In addition, this process can reveal data that does not conform to any of the interpretations/uses that are assumed to exist; any combination of property values that does not fit any of the predefined interpretations/uses will need further detailed investigation.

In what follows, I summarize the **combinations of property values for four uses** described in the literature that are distinguishable by the annotation. The descriptions are based on the classifications presented in [Gast and van der Auwera \(2013\)](#) and [Cabredo Hofherr \(2015\)](#) (see the appendix for summaries).

	internal generic	external generic	anchored existential	unanchored existential
M1	all	all	decl, (other)	decl, (other)
M2	modal, 0	0	0	0
M3	1, 0	0	0	0
M4	(non- progressive) present/past	(non- progressive) present/past	(progressive) present	past
M5	1, 0	1, 0	1, 0	1, 0
M6	1, 0	1	1, 0	1, 0
S1	habitual, static	habitual, static	episodic, static	episodic, static
S2	imp, (ref)	imp, (ref)	ref	ref
S3	nec, (pot)	no, (pot)	no, (pot)	no, (pot)

Table 1: Expected values for the properties of the uses described in the literature

- **Internal generic use:**¹¹ In this use, the dedicated impersonal pronoun is roughly paraphrasable as “people in general” expressing necessary speaker inclusion. The situation type is either habitual or static. The clause containing the dedicated impersonal pronoun may contain a modal and be part of a conditional structure. It is compatible with any sentence type. The temporal verb form is expected to be predominantly non-progressive/imperfective present tense and rarely past tense. Temporal and spatial PPs may occur that do not restrict the main predication to a specific time and place. A 2nd person singular pronoun that is substituted for the dedicated impersonal pronoun is expected to have an impersonal use.
- **External generic use:** In this use, the dedicated impersonal pronoun is roughly paraphrasable as “people in X”, with X providing a restriction on the generality; it does not express necessary speaker inclusion. The situation type is either habitual or static. The clause containing the dedicated impersonal pronoun must not contain a modal or be part of a conditional structure. It is compatible with any sentence type. The temporal verb form is expected to be predominantly non-progressive/imperfective present tense and rarely past tense. A spatial PP has to occur in the clause; temporal PPs may occur but may not restrict the main predication to a specific time. A 2nd person singular pronoun that is substituted for the dedicated impersonal pronoun is expected to have an impersonal use.
- **Anchored existential use:** In this use, the dedicated impersonal pronoun is roughly paraphrasable as “someone” without necessary speaker inclusion. The situation type is either episodic or static. The clause containing the dedicated impersonal pronoun may not contain a modal or be part of a conditional structure. The sentence type is expected to be declarative. The temporal verb form is expected to be (progressive) present tense. Temporal and spatial

¹¹The internal generic use could be split into three uses based on the values of M1–M3, following [Gast and van der Auwera 2013](#). Currently, I do not see the benefit of doing so.

PPs may occur that restrict the main predication to the utterance time and a specific place. A 2nd person singular pronoun that is substituted for the dedicated impersonal pronoun is expected to only have a referential use.

- **Unanchored existential use:**¹² In this use, the dedicated impersonal pronoun is roughly paraphrasable as “someone” expressing no necessary speaker inclusion. The situation type is either episodic or static. The clause containing the dedicated impersonal pronoun must not contain a modal or be part of a conditional structure. The sentence type is expected to be declarative. The temporal verb form is expected to be predominantly past tense. Temporal and spatial PPs may occur that restrict the main predication to a specific time and place. A 2nd person singular pronoun that is substituted for the dedicated impersonal pronoun is expected to only have a referential use.

Table 1 summarizes the value combinations for the four uses; possible but unexpected values are given in parentheses. S4 is not included since the values depend entirely on the intuitions of the annotators.

As stated above, any other combinations of values are unexpected given the uses described in the literature and require further detailed investigation. For instance, the combination of an episodic situation type (S1 episodic) and necessary speaker inclusion (S3 nec) does not fit into any of the five uses and points towards a 1st-person-like use.

¹²Given the properties chosen for the annotation, the *corporate use* and the *inferred existential use*, which are distinguished by Cabredo Hofherr (2015), are grouped with the unanchored existential use (see Appendix C.1). This is the case because the corporate use is centrally characterized by lexical-semantic properties of the main predicate, which are annotated, and the inferred existential reading is characterized by the epistemic state of the speaker, which cannot be determined for corpus data.

References

- Alonso-Ovalle, Luis. 2002. Arbitrary Pronouns Are Not That Indefinite. In *Romance Languages and Linguistic Theory 2000*, ed. Claire Beyssade, Reineke Bok-Bennema, Frank Drijkoningen, and Paola Monachesi. John Benjamins.
- Cabredo Hofherr, Patricia. 2006. “Arbitrary” pro and the theory of pro-drop. In *Agreement and Arguments*, ed. Peter Ackema, Patrick Brandt, Maaïke Schoorlemmer, and Fred Weerman, 230–257. Oxford: Oxford University Press.
- Cabredo Hofherr, Patricia. 2015. Les pronoms impersonnels humains - Syntaxe, sémantique, morphologie. Habilitation Thesis, Université Paris 8.
- Carlson, Gregory. 2005. Generics, Habituals and Iteratives. In *The Encyclopedia of Language and Linguistics*, ed. Keith Brown, 18–21. Elsevier.
- Chierchia, Gennaro. 1995. The Variability of Impersonal Subjects. In *Quantification in Natural Languages*, ed. Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee, 107–143. Kluwer.
- Cinque, Guglielmo. 1988. On *si* constructions and the theory of *arb*. *Linguistic Inquiry* 19:521–582.
- Condoravdi, Cleo. 1989. Indefinite and generic pronouns. In *Proceedings of the eighth west coast conference on formal linguistics*, ed. E. Jane Fee and Katherine Hunt.
- Egerland, Verner. 2018. First Person Readings of MAN: On semantic and pragmatic restrictions on an impersonal pronoun. In *Order and structure in syntax II: Subjecthood and argument structure*, ed. Michelle Sheehan and Laura R. Bailey, 179–195. Berlin: Language Science Press.
- Faarlund, Jan Terje, Svein Lie, and Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Universitetsforl.
- Fenger, Paula. 2018. How impersonal does one get? A study of man-pronouns in Germanic. *Journal of Comparative Germanic Linguistics* 21:291–325.
- Friedrich, Annemarie. 2017. *States, events, and generics: computational modeling of situation entity types*. Doctoral Dissertation, Universität des Saarlandes.
- Friedrich, Annemarie, Kleio-Isidora Mavridou, and Alexis Palmer. 2015a. *Annotation Manual: Situation Entity Types*. Department of Computational Linguistics, Saarland University.
- Friedrich, Annemarie, Chris Palmer, Melissa Peate Sørensen, and Manfred Pinkal. 2015b. Annotating genericity: a survey, a scheme, and a corpus. In *Proceedings of LAW IX – The 9th Linguistic Annotation Workshop*, 21–30.
- Gast, Volker, and Johan van der Auwera. 2013. Towards a distributional typology of human impersonal pronouns, based on data from European languages. In *Languages Across Boundaries: Studies in Memory of Anna Siewierska*, ed. Dik Bakker and Martin Haspelmath, 119–158.

- Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers*. Doctoral Dissertation, University of Amsterdam.
- Gruber, Bettina. 2011. Indexical pronouns: generic uses as clues to their structure. *Poznań Studies in Contemporary Linguistics* 47:331–360.
- Karttunen, Lauri. 2016. Presuppositions: What went wrong. In *Proceedings of SALT 26*, 705–731.
- Kaufmann, Magdalena. 2012. *Interpreting Imperatives*. Springer.
- Kratzer, Angelika. 1997. German impersonal pronouns and logophoricity. Presentation at Sinn und Bedeutung II, Berlin, Germany.
- Kratzer, Angelika. 2012. *Modals and Conditionals: New and Revised Perspectives*. Oxford University Press.
- Krifka, Manfred, Francis Jeffrey Pelletier, Gregory N. Carlson, Alice ter Meulen, Gennaro Chierchia, and Godehard Link. 1995. *The Generic Book*, chapter Genericity: An Introduction, 1–124. University of Chicago Press.
- Maienborn, Claudia. 2001. On the Position and Interpretation of Locative Modifiers. *Natural Language Semantics* 9:191–240.
- Malamud, Sophia A. 2006. *Semantics and pragmatics of arbitrariness*. Doctoral Dissertation, University of Pennsylvania.
- Malamud, Sophia A. 2012. Impersonal indexicals: *one, you, man, and du*. *Journal of Comparative Germanic Linguistics* (Online First):1–48.
- Malamud, Sophia A. 2013. (In)definiteness-driven typology of arbitrary items. *Lingua* 126:1–31.
- Moltmann, Friederike. 2006. Generic one, arbitrary PRO, and the first person. *NLS* 14:257–281.
- Moltmann, Friederike. 2010. Generalising Detached Self-Reference and the Semantics of Generic One. *Mind and Language* 25:440–473.
- Moltmann, Friederike. 2012. Two Kinds of First-Person-Oriented Content. *Synthese, special issue edited by Philippe de Brabanter and Mikhail Kissine* 184:157–177.
- Siewierska, Anna. 2011. Overlap and complementarity in reference impersonals : Man-constructions vs. third person plural-impersonals in the languages of Europe. In *Impersonal Constructions. A cross-linguistic perspective*, ed. Andre Malchukov and Anna Siewierska, 57–90. Amsterdam/Philadelphia: John Benjamins.
- Siewierska, Anna, and Maria Papastathi. 2011. Towards a typology of third person plural impersonals. *Linguistics* 49:575–610.
- Zifonun, Gisela. 2000. “*Man lebt nur einmal.*” Morphosyntax und Semantik des Pronomens *man*. *Deutsche Sprache* 3/00:232–253.
- Zobel, Sarah. 2014. *Impersonally Interpreted Personal Pronouns*. Doctoral Dissertation, Universität Göttingen, Staats- und Universitätsbibliothek Göttingen. URL <http://hdl.handle.net/11858/00-1735-0000-0023-991B-2>.

A Annotation specifics and examples for Norwegian

This appendix provides the additional, Norwegian-specific information used during the annotation process with the Norwegian annotators of IPG_CORE (two linguistics students, one with a finished BA, one with a finished MA) and summarizes the inter-annotator agreements for the annotated properties.

The **Norwegian dedicated impersonal pronoun of interest** is *man*.

The Norwegian examples are taken from the main corpus study conducted for IPG_CORE using the *Leksikografisk bokmålskorpus* and a support study using the *Oslo Multilingual Corpus*, both of which have been accessed via the *Glossa* interface of the *Tekstlaboratoriet* at the University of Oslo: <https://www.hf.uio.no/iln/om/organisasjon/tekstlab/>

A.1 Morphosyntactic properties

A.1.1 SENTENCE TYPE

- (71) SENTENCE TYPE: decl (declarative), int (interrogative), imp (imperative)
- (72) Man begynte igjen å være på vakt for dverger. (Norwegian)
‘MAN again started to look out for dwarves.’
- (73) Var det sånt man snakket om? (Norwegian)
‘Was this the type of thing MAN talked about?’

So far, none of the corpus items investigated so far had an imperative containing *man*, though occurrences embedded inside imperatives should also be possible.

A.1.2 CO-OCCURRING MODAL

- (74) CO-OCCURRING MODAL: note down the exact form of the modal or use 0 (“no modal”)

Norwegian has the present tense modal forms *kan*, *vil*, *skal*, *må*, *bør* with the past tense forms *kunne*, *ville*, *skulle*, *måtte*, *burde*.

- (75) Man **kan** ikke annet enn klappe i hendene. (annotation: kan)
- (76) Man **skal** dessuten ære sine barn. (annotation: skal)
- (77) Denslags var ikke noe man **kunne** diskutere seg fram til, her var svaret overlatt til den enkeltes fantasi. (annotation: kunne)
- (78) Man **må** spørre seg om noe annet menneske i moderne tid har hatt slik en enorm innflytelse på den tenkende verden. (annotation: må)

A.1.3 IN CONDITIONAL

- (79) IN CONDITIONAL: 0 (= no), 1 (= yes)

Norwegian has the introductory expressions *hvis*, *om*, and *når*.

- (80) Hvis man var heldig, ville man merke gasslukt. (annotation: 1)
 ‘If MAN was lucky, MAN would notice a gas smell.’
- (81) Det går utmerket godt an å nære en faglig interesse for bakterier og virus selv, om man ikke akkurat trakter etter en ubeskyttet nærkontakt med de samme organismene. (annotation: 1)
 ‘It’ s perfectly possible to nurture a professional interest in bacteria or viruses even though MAN doesn’t exactly long for a close and unprotected encounter with them.’
- (82) Det er ikke helt trygt å ha dem bak seg, når man ikke har øyne i nakken. (annotation: 1)
 ‘It is not entirely safe to have them behind SELF, if MAN doesn’t have eyes in the back of your head.’

Like English, Norwegian allows for conditional clauses without an introductory expression, but unlike English, this possibility is not constrained to certain types of conditionals.

- (83) Fikk man tre skattekister på rad der, bar det rett inn på Indiana Jones’ eventyrsti. (annotation: 1)
 ‘If MAN got three treasure chests in a row, it went straight onto Indiana Jones’ adventure trail.’

A.1.4 TEMPORAL VERBAL FORM

Norwegian has more or less the same inventory of temporal verbal forms as English without the progressive forms:

- (84) TEMPORAL VERBAL FORM: pres (present), past, perf (perfect), plup (pluperfect) fut (future) fut.perf (future perfect)
- (85) Ensom lever man og ensom dør man. (annotation: pres)
 ‘MAN lives alone and MAN dies alone.’
- (86) Og man var enige om at det hadde vært en fin dag. (annotation: past)
 ‘And MAN agreed that it had been a good day.’
- (87) Man har oppnevnt en offentlig forsvarer for Dem. (annotation: perf)
 ‘MAN has appointed a public defender for you.’
- (88) På en slik sledetur viste det seg en gang at man hadde glemt å ta med seg skytevåpen. (annotation: plup)
 ‘On such a sled trip, it once became apparent that MAN had forgotten to bring firearms.’
- (89) Man skal bare drikke colaen først. (annotation: pres)
 ‘MAN just has to drink the coke first.’

A.1.5 TEMPORAL ADVERBIAL

- (90) TEMPORAL ADVERBIAL: 0 (= no), 1 (= yes)

Examples of Norwegian temporal adverbials: *i dag, nå, klokka to, deretter,*

- (91) Mulighetene for “misbruk” av for eksempel bioteknologi, kjemi og våpen teknologi som opprinnelig ble utviklet med andre mål for øyet er stor, og er noe av det man **i dag** frykter mest. (annotation: 1)
 ‘The possibilities for “abuse” of, for instance, biotechnology, chemistr, or weapons technology that originally were developed with another goal in mind are large and are something that MAN fears most today.’
- (92) **Deretter** satte man ut reker i bur i og rundt et oppdrettsanlegg i Rogaland. (annotation: 1)
 ‘After that, MAN set out shrimps in cages in and around a fish farm in Rogaland.’

A.1.6 SPATIAL ADVERBIAL

- (93) SPATIAL ADVERBIAL: 0 (= no), 1 (= yes)

Examples of Norwegian spatial adverbials: *i Oslo, her,*

- (94) Deretter satte man ut reker i bur **i og rundt et oppdrettsanlegg i Rogaland**. (annotation: 1)
 ‘After that, MAN set out shrimps in cages in and around a fish farm in Rogaland.’
- (95) Friheten man drømmer om **fra celler eller slott**, har den noen fellestrekk, kan den lett forstås fra menneske til menneske? (annotation: 1)
 ‘The freedom MAN dreams of from prison cells or castles, does it have any common features, can it be easily understood from person to person?’

A.2 Semantic properties

A.2.1 SITUATION TYPE

- (96) SITUATION TYPE: epi (episodic), hab (habitual), stat (static)

Norwegian versions of the tests proposed by [Friedrich et al. \(2015a\)](#):

- **Episodicity test:** Can *nå* or *i går* or any other specific temporal adverbial be added?
- **Habituality test:** Can *typisk* or *vanligvis* be added?

The following Norwegian examples list the situation types of all finite clauses in more complex corpus examples to give the annotators more complex examples to compare to.

- (97) I spørsmålet om overgangsordninger for innføringen av EUs matvarepriser i Norge, ga han klart uttrykk for at han støtter det norske synet.
Main clause: han klart uttrykk. . . (→ episodic)
Embedded clause: han støtter det norske synet (→ static)
- (98) Vi må være villige til å bruke de handelspolitiske virkemidlene vi har til rådighet.
Main clause: vi må være villige til å. . . (→ static)
Relative clause: (midlene) vi har til rådighet (→ static)

- (99) I tillegg ser mellom anslagsvis 150000 mennesker TVNorges mest sette nyhetssending hver dag.
Main clause: ... mennesker ser ... nyhetssending hver dag (→ habitual)
- (100) Folk jeg traff i Havanna påsto at den karibiske haien får seg en kubaner til middag hver dag.
Main clause: Folk. . . påsto at. . . (→ episodic)
Relative clause: (folk) jeg traff i Havanna (→ episodic)
Embedded clause: den karibiske haien får seg. . . hver dag (→ habitual)
- (101) Hvis det blir flertall for en slik beslutning, er veien også åpen for en privatisering om det er politisk ønskelig.
Main clause: veien er også åpen for en privatisering. . . (→ static)
hvis-clause: det blir flertall for en slik beslutning (→ static)
om-clause: det er politisk ønskelig (→ static)

A.2.2 PERSONAL PRONOUN

The Norwegian 2nd person singular pronoun *du* has an *impersonal use* with a generic interpretation (Faarlund et al. 1997), so the instructions in the main guidelines can be applied to Norwegian without problems.

- (102) PERSONAL PRONOUN: ref (regular reference only), imp (impersonal interpretation possible)

Two examples based on the intuitions of the two annotators:

- (103) a. I vår globale tid er **man** jo mindre nasjonal og mer kosmopolitisk.
 ‘In our global time, MAN is less national and more cosmopolitan.’
 b. I vår globale tid er **du** jo mindre nasjonal og mer kosmopolitisk.
 ⇒ can have the same general interpretation as the sentence with *man*
 (annotation: imp)
- (104) a. Deretter satte **man** ut reker i bur i og rundt et oppdrettsanlegg i Rogaland.
 ‘After that, MAN set out shrimps in cages in and around a fish farm in Rogaland.’
 b. Deretter satte **du** ut reker i bur i og rundt et oppdrettsanlegg i Rogaland.
 ⇒ *du* can only refer to the addressee
 ⇒ the sentence has a different meaning from the original (annotation: ref)

A.2.3 SPEAKER INCLUSION

- (105) SPEAKER INCLUSION: nec (necessarily/definitely included), pot (potentially included), no (not included / definitely excluded)

No language specific information applies.

A.2.4 CONTEXTUAL SPECIFICATION

- (106) **CONTEXTUAL SPECIFICATION:** note down who is meant if it can be determined from the context; otherwise use “0” for “no specification”

No language specific information applies.

B Annotation specifics and examples for German

This appendix provides the additional, German-specific information used by two annotators (the author and a German MA/PhD student) in the course of a corpus study that was started in preparation for IPG_CORE and was completed in parallel to the corpus study on Norwegian *man*, and it summarizes the inter-annotator agreements for the annotated properties.

The **German dedicated impersonal pronoun of interest** is *man*.

The German examples are taken from different corpora of written German available online via the *Digitales Wörterbuch der Deutschen Sprache* (DWDS): <https://www.dwds.de/>

B.1 Morphosyntactic properties

B.1.1 SENTENCE TYPE

- (107) **SENTENCE TYPE:** decl (declarative), int (interrogative), imp (imperative)

For German, no further sentence types are distinguished.

- (108) Man trifft sich hier auf einer Party, dort bei einem Vortrag. (annotation: decl)
 ‘MAN meets here, at a party, and there, at a talk.’
- (109) Hat man da nicht schon mal Sorgen? (annotation: int)
 ‘Doesn’t MAN worry in that case?’

The example for *man* occurring in an embedded clause inside an imperative is repeated here:

- (110) Gib zu, dass man da nichts machen kann! (annotation: imp)
 ‘Admit that MAN can’t do anything about that!’

Note that this is an attested example from the internet; no comparable example has been found in the corpus sample.

B.1.2 CO-OCCURRING MODAL

- (111) **CO-OCCURRING MODAL:** note down the infinitive of the modal or use 0 (“no modal”)

German provides the modal verbs *dürfen*, *können*, *mögen*, *müssen*, *sollen*, and *wollen*, which have a complete inflection paradigm. Hence, the infinitive instead of the exact form is annotated for German.

- (112) Schließlich **will** man ja nichts versäumen. (annotation: *wollen*)
 ‘After all, MAN does not want to miss anything.’
- (113) Aber man **müsse** solche Aussagen häufig wiederholen. (annotation: *müssen*)
 ‘But MAN has to frequently repeat such statements.’
- (114) Trotz dieser grundlegenden Bedeutung von Schlaf gibt es über seine Funktion nur Hypothesen [10], die man grob zwei Kategorien zuordnen kann und die sich nicht gegenseitig ausschließen. (annotation: *können*)
 ‘Despite the fundamental significance of sleep, there are only hypotheses regarding its function, which MAN can roughly be put into two categories and which do not exclude each other.’

B.1.3 IN CONDITIONAL

- (115) IN CONDITIONAL: 0 (= no), 1 (= yes)

German has the introductory expressions *wenn* and *falls*.

- (116) Wenn man das Licht ausmacht,
glaubt er, um ihn herum verschwindet die Welt! (annotation: 1)
 ‘When MAN turns off the light, he believes that the world around him vanishes.’
- (117) Wie man das macht, wenn man nicht selbst Mitglied der Kontrollkommission ist,
hat Herr Neumann leider nicht verraten. (annotation: 1)
 ‘How MAN does that if MAN is not a member of the control committee, Herr Neumann unfortunately did not reveal.’
- (118) Vielleicht könntest du dein Liebesleben, falls man das so nennen kann, mal auf deine eigenen vier Wände beschränken. (annotation: 1)
 ‘Maybe you could restrict your love life, if MAN can call it that, to your own home.’

Like English and Norwegian, German allows for conditional clauses without an introductory expression, and like Norwegian, this possibility is not constrained to specific types of conditionals.

- (119) Nimmt man die aus früheren Zeiten übernommenen Fabriken dazu sowie jene, die nach dem Abschluß der Aufbau- und Modernisierungsphase von der Substanz lebten,
kann man dem Zeitgeschichtler Lutz Niethammer nur beipflichten. (annotation: 1)
 ‘If MAN includes the factories that were adopted from earlier times and those that [...], MAN can only agree with the historian Lutz Niethammer.’

The annotation is “0” if *man* does not occur inside a conditional:

- (120) **Man** muss dabei sein. (annotation: 0)
 ‘MAN has to be there.’

B.1.4 TEMPORAL VERBAL FORM

German has more or less the same inventory of temporal verbal forms as Norwegian. However, German also has two “subjunctive” forms for present and past: Konjunktiv I and Konjunktiv II, which are added for German, since the forms are distinct from the indicative forms.

- (121) TEMPORAL VERBAL FORM: pres (present), past, perf (perfect), plup (pluperfect) fut (future) fut.perf (future perfect), kon1 (Konjunktiv I), kon2 (Konjunktiv II)
- (122) Man **verzichtet** damit auf die Idee, Komplexität transparent und einsichtig (intelligibel) zu machen. (annotation: pres)
 ‘By that, MAN sets aside the idea to make complexity transparent and intelligible.’
- (123) Wenn Probleme auftauchten, dann **gab** man denen die Nummer, und sie riefen dort an. (annotation: prt)
 ‘If problems arose, then MAN gave them the number, and they called there.’
- (124) Man **könnte** das ja als Eingeständnis werten, an den Gerüchten sei gar nichts dran. (annotation: kon2)
 ‘MAN could take this as a confession that there was nothing to the rumors.’

In case the inflected verb is a modal verb, certain verbal forms may have an idiosyncratic interpretation. This should be ignored. Only the given verbal form is of interest.

B.1.5 TEMPORAL ADVERBIAL

- (125) TEMPORAL ADVERBIAL: 0 (= no), 1 (= yes)

Illustrative examples of German temporal adverbials: *heute, am Sonntag, letztes Jahr, vor drei Tagen, ...*

- (126) **Wenig später** entschied man sich jedoch wieder für eine eigenständige Ausbildungsstätte. (annotation: 1)
 ‘Soon after, MAN again decided in favor of an independent training post.’
- (127) Auf der anderen Seite will man **gleich zu Beginn** sich Notizen machen. (annotation: 1)
 ‘On the other hand, MAN wants to take notes from the beginning.’

B.1.6 SPATIAL ADVERBIAL

- (128) SPATIAL ADVERBIAL: 0 (= no), 1 (= yes)

Illustrative examples of German spatial adverbials: *in Berlin, hier, daneben, ...*

- (129) Man beobachtete **von allen Terrassen** die DAN-Maschine auf dem Rollfeld. (annotation: 1)
 ‘From all terraces, MAN observed the DAN-plane on the runway.’
- (130) Naja, ich wußte nicht, ob man **in der Situation, die ich beschrieben habe**, handeln muß. (annotation: 1)
 ‘Well, I didn’t know whether MAN has to act in the type of situation that I described.’

B.2 Semantic properties

B.2.1 SITUATION TYPE

(131) SITUATION TYPE: epi (episodic), hab (habitual), stat (static)

The treatment of modal predicates and clauses that are part of a conditional is extended to also cover occurrences of **Konjunktiv I and Konjunktiv II** (Friedrich et al. 2015a:45). Both types of Konjunktiv-marking are connected to a modal-like meaning. Hence, the presence of Konjunktiv I or Konjunktiv II signals that the situation that is described is static.

(132) Und wie schnell man wieder draußen sei. (annotation: stat)
 ‘And how fast MAN is outside again.’

(133) Man wäre einige Tage mit dem Schiff unterwegs. (annotation: stat)
 ‘MAN would be traveling by ship for a number of days.’

Similarly, *lassen* in connection with an infinitive (\approx ‘to allow’) and *haben* in connection with a *zu*-infinitive (\approx ‘must’) have modal-like meaning and also describe a static situation (Friedrich et al. 2015a:46).

(134) Man ließ uns gewähren. (annotation: stat)
 ‘MAN let us do that.’

(135) Man hat immer zu repräsentieren. (annotation: stat)
 ‘MAN always has to represent the company.’

German versions of the tests proposed by Friedrich et al. (2015a):

- **Episodicity test:** Can *jetzt* or *gestern* or any other specific temporal adverbial be added?
- **Habituality test:** Can *typischerweise* or *normalerweise* be added?

B.2.2 PERSONAL PRONOUN

The German 2nd person singular pronoun *du* has an *impersonal use* with a generic interpretation (Gruber 2011, Malamud 2006, 2012, Zobel 2014), so the instructions in the main guidelines can be applied to German without problems.

(136) PERSONAL PRONOUN: ref (regular reference only), imp (impersonal interpretation possible)

Two examples based on the author’s native intuitions:

(137) a. Heutzutage muss **man** ja nicht mehr physisch präsent sein.
 ‘Nowadays, MAN doesn’t have to be physically present.’
 b. Heutzutage musst **du** ja nicht mehr physisch präsent sein.
 \Rightarrow can have the same general meaning as the sentence with *man*
 (annotation: imp)

(138) a. Gestern ist **man** mit Waffen gegen die unbewehrte Volksmasse angestürmt.
 ‘Yesterday, MAN attacked the unarmed masses with weapons.’

- b. Gestern bist **du** mit Waffen gegen die unbewehrte Volksmasse angestürmt.
 ⇒ *du* can only refer to the addressee
 ⇒ the sentence has a different meaning from the original (annotation: ref)

B.2.3 SPEAKER INCLUSION

- (139) SPEAKER INCLUSION: nec (necessarily/definitely included), pot (potentially included), no (not included / definitely excluded)

No language specific information applies.

B.2.4 CONTEXTUAL SPECIFICATION

- (140) CONTEXTUAL SPECIFICATION: note down who is meant if it can be determined from the context; otherwise use “0” for “no specification”

No language specific information applies.

C Recent classifications of the interpretations/uses of human impersonal pronouns

Note: the term *human impersonal pronouns* not only covers dedicated impersonal pronouns but also impersonally used personal pronouns (e.g., impersonally used English 2nd person singular *you* or 3rd person plural *they*).

C.1 The classification in Cabredo Hofherr 2015

Cabredo Hofherr 2015 bases her classification of the uses on earlier work (e.g., Cabredo Hofherr 2006) and the results by Siewierska and Papastathi (2011). She distinguishes **seven potential readings** for human impersonal pronouns, which she illustrates with examples taken from Spanish and French (Cabredo Hofherr 2015:16).

1. specific existential reading (temporally anchored)

- (141) Tocan a la puerta. (Spanish)
 ‘*pro*.3pl are knocking at the door.’ (pro.3pl ≈ someone)

2. vague existential reading (not temporally anchored)

- (142) Ils ont trouvé une moto dans la cour. (French)
 ‘ILS found a motorbike in the yard.’ (ILS ≈ someone)

3. inferred existential reading (inferred from a result)

- (143) Aquí han comido mariscos. (Spanish)
 ‘*pro*.3pl have eaten sea food here.’ (pro.3pl ≈ someone)

4. corporate reading (predicates with a designated subject)

- (144) Ils ont encore augmenté la TVA. (French)
 ‘ILS increased the taxes again.’ (ILS ≈ the authorized people)

5. restricted universal reading (licensed by a locative)

- (145) En España hablan español. (Spanish)
 ‘In Spain, *pro*.3pl speak Spanish.’ (*pro*.3pl ≈ people in Spain)

6. evidential reading (with verbs of saying)

- (146) Dicen que en esa casa vivió Darwin. (Spanish)
 ‘*pro*.3pl say that Darwin lived in this house.’ (*pro*.3pl ≈ ??)

7. unrestricted universal reading

- (147) On ne mange pas avec les doigts. (French)
 ‘ON doesn’t use the fingers to eat.’ (ON ≈ people in general)

The evidential reading listed by Cabredo Hofherr is special since in this use, the human impersonal pronoun cannot be paraphrased by either “people in general” or “someone”, and no specific group of individuals can be inferred from the verbs of saying that it co-occurs with (see also [Siewierska and Papastathi 2011](#)).

Dedicated impersonal pronouns (i.e., “R-impersonnels spécialisés”) from different languages differ with respect to the available readings. English *one* only allows for generic readings, while German *man* and French *on* in principle allow for generic and existential readings. German *man* and French *on*, however, seem to differ with respect to the subtypes of existential readings: German *man*, unlike French *on*, is degraded in the specific existential reading, see (148).

- (148) a. #**Man** klopft an der Tür. (German)
 b. **On** frappe à la porte. (French)
 ‘IMP is knocking on the door.’ (IMP ≈ someone)

C.2 The typology in [Gast and van der Auwera 2013](#)

[Gast and van der Auwera \(2013\)](#) distinguish **seven potential uses** for human impersonal pronouns. These seven uses differ with respect to two sets of parameters of variation (see [Gast and van der Auwera 2013:149–153](#)): (i) the properties of the state of affairs described by the sentence and (ii) the properties of the set of individuals described by the pronominal expression.

On the highest level, **properties of the state of affairs** distinguish whether the proposition expressed by the ‘bare clause’ (i.e., the lexical verb including its arguments and event-modifying adverbials) is veridical or non-veridical. Such a proposition is non-veridical if it occurs in the scope of a non-veridical (intensional) operator, and it can be assumed that the speaker does not take the proposition to be true in the context of utterance. Otherwise, the proposition is said to be veridical. These two main types are further divided into the following **four cases**:

- **veridical/episodic**: a particular, temporally and spatially specific state of affairs
- **veridical/generic**: a generalization over time and space
- **non-veridical/modal**: the clause contains a modal expression inducing non-veridicality

- **non-veridical/non-modal:** the clause does not contain any overt expression to which the non-veridicality can be attributed

The **properties of the set of individuals** concern the interpretation of the impersonal expression. On the highest level, Gast and van der Auwera distinguish universal (i.e., generic) interpretations from existential interpretations, which correspond to the two uses with impersonal meaning that are traditionally distinguished (see the Introduction). The universal and existential cases each have further subcases distinguishing **five possible cases** for the individuals that are talked about:

- **universal/internal:** a general set of individuals that the speaker identifies with¹³
- **universal/external:** a general set of individuals that the speaker does not identify with
- **existential/definite:** a particular set of individuals that are “accessible in the discourse environment”
- **existential/indefinite/vague:** a particular set of one or more individuals that are “inaccessible in the discourse environment”
- **existential/indefinite/plural:** a particular set of multiple individuals that are “inaccessible in the discourse environment”

If the two sets of parameters were completely independent, we would expect to find 20 possible cases (= 4 × 5). Based on the data from impersonal human pronouns from European languages, [Gast and van der Auwera \(2013:152–153\)](#) distinguish **seven contexts**; the corresponding readings in [Cabredo Hofherr 2015](#) are indicated to the right (see above). The contexts are illustrated by seven “(generalized) diagnostic sentences”.

- | | |
|---|-----------|
| 1. veridical/episodic + existential/indefinite/vague | (CH: 1–3) |
| (149) They ’re knocking on the door. | |
| 2. veridical/episodic + existential/indefinite/plural | (CH: 1–3) |
| (150) They ’ve surrounded us. | |
| 3. veridical/episodic + existential/definite | (CH: 4) |
| (151) They ’ve raised the taxes again. | |
| 4. veridical/generic + universal/external | (CH: 5) |
| (152) They eat dragonflies in Bali. | |
| 5. veridical/generic + universal/internal | (CH: 7) |
| (153) One only lives once. | |

¹³[Gast and van der Auwera \(2013\)](#) adopt the notion of speaker identification from the articles on English *one* by [Moltmann \(2006, 2010, 2012\)](#); see also Section 2.1.

6. non-veridical/modal + universal/internal (CH: 7)

(154) **One** should never give up.

7. non-veridical/non-modal + universal/internal (CH: 7)

(155) What happens if **one** drinks sour milk?

These seven contexts reflect the **correlation between episodic sentence contexts and existential interpretations**, on the one hand, and **generic/modal/non-modal non-veridical sentence contexts and universal/generic interpretations**, on the other hand, but incorporate some of the subdivisions proposed in [Cabredo Hofherr 2006](#) and [Siewierska and Papastathi 2011](#).

According to [Gast and van der Auwera \(2013:155–161\)](#), all dedicated impersonal pronouns that developed from the noun for *human* (e.g., German/Norwegian/Swedish *man*, Dutch/Frisian *men*, French *on*) can be used in all seven contexts (*pace* [Cabredo Hofherr \(2015\)](#) on German *man*). In contrast, Italian *si* is claimed to be usable only in the contexts 4–7.