A questionnaire-based study of impersonalization in Romanian and English
With special attention to passivization

Valentin Rădulescu and Daniël Van Olmen
Lancaster University

1 Introduction

The study of impersonalization has a well-established tradition. The phenomenon, exemplified by *one* in (1), is often subsumed under the functional domain of agent defocusing, which covers a variety of constructions across languages that shift attention away from the agent in the clause (e.g. Shibatani 1985; Myhill 1997). They include, inter alia, middle constructions and passives like (2a) and (2b) respectively. Agent defocusing is motivated by a variety of factors. An agent may, for instance, be demoted to an oblique when the patient is more topical (see Myhill 1997: 805), which appears to be the case in (2b) with its demonstrative-marked patient-subject. It is also possible that an agent is omitted altogether when its reference is obvious from the context (see Shibatani 1985: 831). Importantly, as argued by Sansò (2006) among others, such factors affect the choice of construction in a language and languages can differ in the construction that they use under similar conditions.

(1) *One* shouldn't say things like that. (Myhill 1997: 802)

(2) a. The book sells well. (Sansò 2006: 242)
b. That book was written by Mary. (Myhill 1997: 803)

The precise conditions of agent defocusing that constitute the topic of this article are when the agent is generic (e.g. Sansò 2006: 242-245) – more or less paraphrasable as 'everybody' or, in the case of (1), 'nobody' – or when the agent is a particular but unknown (set of) individual(s) (e.g. Myhill 1997: 806-822) – roughly equivalent to 'someone' or 'some people', as in (3). The constructions under investigation are *one* in (1) and the third person plural in (3), as well as the second person singular in (4a) and the agentless passive in (4b), intended here as the respective counterparts of (1) and (3).[1] And the languages of study, finally, are English and Romanian.

(3) *They*'ve taken the radio. (Myhill 1997: 802)

(4) a. *You* shouldn't say things like that.
b. The radio has *been taken*.

In the following sections, we will first describe our approach to the particular agent defocusing conditions at issue, then introduce the four constructions mentioned above and, lastly, discuss our interest in the two languages within this context.

1.1 Impersonalization

The phenomena in (1), (3) and (4) are not always grouped together in the literature. Kitagawa and Lehrer (1990: 742), for example, reserve the label "impersonal" for generic cases and classify cases with "specific individuals [that] … are not identified, or identifiable, by the speaker" as "vague". We, however, will follow Gast and van der Auwera (2013), who observe that many

[1] Unless specified otherwise, the examples are constructed by the authors.

languages employ the same construction for (1) and (3). Dutch *men* ‘one, they’ in (5) is a case in point (see also Giacalone Ramat and Sansò 2007 on so-called ‘man’-pronouns).

(5) a. Dat soort ding-en moet *men* niet zeg-gen.
DIST.DEM.SG sort thing-PL must.SG.PRS MEN NEG say-INF
‘One shouldn’t say things like that.’
b. *Men* heeft de radio gestolen.
MEN have.3SG.PRS DEF.SG.M radio steal.PST.PTCP
‘They’ve taken the radio.’

The different agent defocusing conditions can thus, in their view, be considered instances of a single domain of “impersonalization”, which they define as “the process of filling an argument position of a predicate with a variable ranging over sets of human participants without establishing a referential link to any entity from the universe of discourse” (Gast and van der Auwera 2013: 124).

In their map of impersonal uses, given in Figure 1, Gast and van der Auwera (2013) call contexts paraphrasable as ‘everyone’ “universal” (UNI) and contexts paraphrasable as ‘someone, some people’ “existential” (EXI). Based on a survey of impersonal constructions in European languages, the map makes additional distinctions, though.

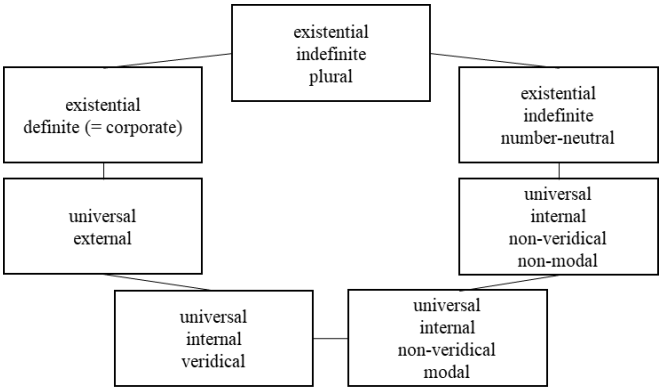

Figure 1: Adapted version of Gast and van der Auwera’s (2013: 141) semantic map of impersonalization

In the vein of, among others, Kitagawa and Lehrer (1990: 749) and Sansò (2006: 268), it sets apart universal uses with an “internal” (INT) perspective, like (4a), from those with an “external” (EXT) one, like (6a). The former is characterized by Gast and van der Auwera (2013: 139) as follows: “A ‘center of consciousness’ (e.g. the speaker or hearer) identifies, or is identified, with the set of referents under discussion.” In (6b), for instance, the statement pertains to anyone, the interlocutors included, who happens to be or can imagine themselves in Bali. By contrast, (6a) is about the set of inhabitants of the country, typically excluding the speaker and the addressee.

(6) a. In Bali, *they* eat dragonflies. (Gast and van der Auwera 2013: 144)
b. In Bali, *you* eat dragonflies.

Further differentiation within the universal-internal domain in Figure 1 rests on the parameters of veridicality – i.e. the state of affairs is (not) assumed to be true – and of modality – i.e. the non-veridical state of affairs has an/no explicit marker of possibility or necessity (see Gast and van der Auwera 2013: 137). The first parameter distinguishes (7a)[2] from (7b) and (7c): only

[2] Another example of the use presented in (7a) is (6b).

living once is presented as an unqualified fact, making this context "veridical" (VER). The second one separates (7b) and (7c). The "non-veridicality" (NVER) of (7b) is due to the occurrence of the modal (MOD) *shouldn't*, which puts the state of affairs of drinking and driving in the realm of the (un)necessary. In (7c), there may be no modal marker (NMOD) but the conditional clause still renders drinking sour milk hypothetical and hence non-veridical. Crucially, each distinction is motivated by at least one impersonal construction in the languages of Europe that can occur in one use but not in another (see Gast and van der Auwera 2013: 140-154). The parameters have also been shown to be relevant for language-specific descriptions, though. Laberge and Sankoff (1979: 432), for example, indicate that French prefers *tu* 'you' to *on* 'one' in conditional contexts.

(7) a. *You* only live once.
b. *You* shouldn't drink and drive.
c. What happens if *one* drinks sour milk?
(all from Gast and van der Auwera 2013: 138)

The existential domain in Figure 1 displays more fine-grained uses too. Gast and van der Auwera (2013: 140) distinguish definite or "semi-impersonal" contexts like (8a) from indefinite or "truly impersonal" ones like (8b) and (8c). In (8a), the referents may not be explicitly identified but the action of raising taxes itself makes them partly accessible, since only the "corporate" (COR) entity of the government can undertake it. Examples (8b) and (8c), however, are intended as applying to a particular set of human participants whose referents *are* inaccessible or, in short, unknown. The difference between (8b) and (8c) lies in number: the set of individuals in (8b) is "necessarily plural" (PL) whereas that in (8c) is "neutral with regard to number" (NN) and may consist of one person or more than one person.

(8) a. *They*'ve raised the taxes again.
b. *They*'ve surrounded us.
c. *They*'re knocking on the door.
(all from Gast and van der Auwera 2013: 152)

Cross-linguistic support for the distinctions comes from languages where the third person plural can be employed impersonally for (8a) but not (8b) and (8c) or for (8a) as well as (8b) but not (8c). Unlike *they*, German *sie* 'they', for instance, can only really occur with a plural reading (Gast and van der Auwera 2013: 144; see also Siewierska and Papastathi 2011: 596).

The third person plural has actually been the focus of a second map of impersonal uses, given in Figure 2. In their study of nine European languages, the authors, Siewierska and Papastathi (2011), recognize the same universal-external and existential-corporate uses, as in (6a) and (8a), but they add a "speech act verb" (SAV) use and make distinctions in the indefinite or truly impersonal existential domain based not on number – (8b) versus (8c) – but on a parameter characterizable as "(un)knownness" (note that "anaphoric" in Figure 2 stands for personal 'they').

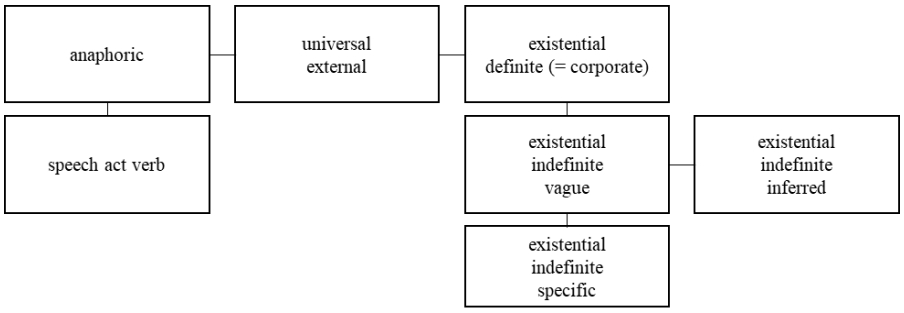

Figure 2: Adapted version of Siewierska and Papastathi's (2011: 604) semantic map for 'they'-impersonals

The extra context, involving a speech act verb, is exemplified in (9) and marks information as coming from some unidentified source (see also Myhill 1997: 811). Siewierska and Papastathi (2011) regard it as a direct off-shoot from anaphoric 'they' in their map, unrelated to the other impersonal uses.[3] This decision is driven by languages like French and Finnish: the third person plural in French can serve a variety of impersonal purposes but is not acceptable in (9) whereas its Finnish counterpart cannot be employed universal-externally or existentially but can occur with speech act verbs (see Siewierska and Papastathi 2011: 585, 591). This separate status of the use in (9) is the reason why Gast and van der Auwera (2013) do not include it in their map. Still, the fact that, in many languages, markers dedicated to the expression of impersonalization in general can appear with speech act verbs too, like Dutch *men* in (10), suggests that studies of impersonalization, like the present one, should take the context into account.

(9) *They* say he met vampires in the Black Forest. (Siewierska and Papastathi 2011: 585)

(10) *Men* zeg-t dat hij in het Zwarte Woud vampier-en
MEN say-3SG.PRS COMP 3SG.M.NOM in DEF.SG.N Black Forest vampire-PL
ontmoet heeft.
meet.PST.PTCP have.3SG.PRS
'They say he met vampires in the Black Forest.'

For the indefinite existential realm, Siewierska and Papastathi (2011) follow Cabredo Hofherr (2006), who distinguishes three uses that vary in the type or level of (un)knownness (see Creissels 2008a: 5-6 too). In a "vague" (VAG) context like (11a), the speaker knows that a particular (group of) individual(s) carried out some action but is unable or unwilling to characterize them (the term "vague" will henceforth be used in this sense and not in Kitagawa and Lehrer's 1990 sense, mentioned earlier). In an "inferred" (INF) context like (11b), the speaker deduces from the situation (a smell here) that an essentially unknown event took place and that some unspecified person/people is/are responsible for it. In a "specific" (SPE) one like (11c), finally, something happens at an exact place and time, typically the speaker's here and now – in this case, a knock on the door. They may therefore be aware of who is causing it but abstain from identifying them.

(11) a. *They*'ve found his bike in the back of a barn.
b. *They*'ve been frying chips here. [I can smell it.]
c. *They*'re knocking on the door. [It might be Mary.]
(all from Siewierska and Papastathi 2011: 581)

[3] As Siewierska and Papastathi (2011: 585) point out, it differs from the existential ones in that it is "typically not episodic, i.e. do[es] not refer to a specific event" and it "typically cannot be substituted by *someone*, as the referent of *they* clearly corresponds if not to the whole human race then to some group of people at a given time and place".

These distinctions are again motivated by cross-linguistic differences, in the distribution of the third person plural: in some languages, 'they' is restricted to universal-external and existential-corporate uses; in others, it may also be existential-vague but not -inferred or -specific; and in yet other languages, it can occur in existential-vague and -inferred but not -specific contexts or in existential-vague and -specific but not -inferred ones. Syrian Arabic would be an example of the last type of language and French an example of a language with universal-external, existential-corporate and existential-vague uses only for its third person plural (Cabredo Hofherr 2006: 245; see also Siewierska and Papastathi 2011: 596).

Figures 1 and 2 differ in the inclusion of the universal-internal and speech act verb uses and in the differentiation within the indefinite existential ones. The map in Figure 3, proposed by Van Olmen and Breed (2018a, 2018b), merges the two, to capture the "complete" range of impersonal contexts. It appends Siewierska and Papastathi's (2011) speech act verb use to Gast and van der Auwera's (2013) map and combines the former's parameter of (un)knownness and the latter's parameter of number, resulting in existential-vague/inferred/specific-plural and existential-vague/inferred/specific-number-neutral uses. The reason for bringing the parameters together is that, at least for the third person plural in West Germanic, they have been shown to interact. In Van Olmen and Breed's (2018b: 822-838) study of Afrikaans, Dutch and English, the acceptability of 'they', as well as its usage (in a completion task), decreases from vague/inferred/specific-plural to vague/inferred/specific-number-neutral contexts and from vague-plural/number-neutral to inferred-plural/number-neutral and specific-plural/number-neutral ones.

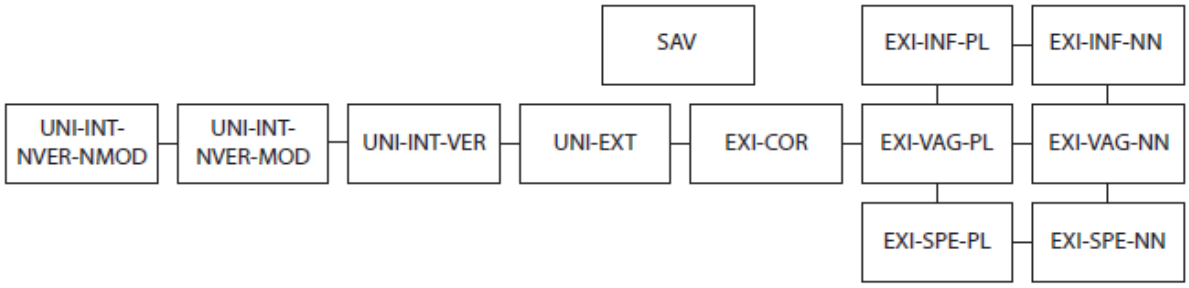

Figure 3: Van Olmen and Breed's (2018b: 840) combined semantic map of impersonalization

The twelve impersonal uses in Figure 3 will be central to the rest of the present article, which will hereafter employ their abbreviations.

1.2 Impersonal constructions

The constructions under examination in this study are *one*, the second person singular, the third person plural and the passive. 'One'-pronouns can be found in a variety of languages. Spanish, with *uno* 'one', can serve as an example (e.g. Fernández 2013: 93-94) and so can English, as (1) – repeated here as (12) – shows. Romanian has no such pronoun, though.

(12) *One* shouldn't say things like that. (Myhill 1997: 802)

In general, 'one'-pronouns are known to be more or less restricted to UNI-INT uses (e.g. Gast and van der Auwera 2013: 149; Fenger 2018: 298). Yet, they can also be employed "for reference to the first person, especially if the speaker wants to represent his/her behaviour as a result of general rules" (van der Auwera et al. 2012: 47), like the politician in (13). This extended use is, however, beyond the scope of the present article.

(13) *One* doesn't want to set quotas. *One* doesn't want to set diktats, but *one* does want to maintain a dialogue. (van der Auwera et al. 2012: 48)

English *one* is typical of formal language and has been said to invite the addressee "to identify with the group at issue" (Haas 2018a: 174) less than *you*. It has also been suggested that *one* tends to occur in NVER contexts (e.g. Moltmann 2010: 463-466; Rudolf 2016: 115), although this correlation has become less pronounced over time (Haas 2018b).

The impersonal use of the second person singular is a common phenomenon in language (e.g. Laberge and Sankoff 1979 for French; De Hoop and Tarenskeen 2015 for Dutch; Jensen and Gregersen 2016 for Danish) and is attested in English and Romanian too, as *you* in (7a) – repeated here as (14) – and *ai* 'you have' in (15) make clear (the overt subject pronoun *tu* 'you' only appears for, say, emphasis, contrast or disambiguation in Romanian and does not normally allow an impersonal reading; see Pană Dindelegan 2013: 104-111).

(14) *You* only live once. (Gast and van der Auwera 2013: 138)

(15) *Ai* doar o viață.
have.2SG.PRS just INDF.F.SG life
'You only live once.'

In the literature, considerable attention has been paid to the way in which personal 'you' turns impersonal and to the question whether the difference in interpretation is semantic or pragmatic in nature (e.g. Malamud 2012; Gast et al. 2015). What is important here is that the central role of the addressee persists in the impersonal second person singular in that it has been argued to essentially be limited to UNI-INT uses across languages (Gast and van der Auwera 2013: 146). Such uses typically include the addressee among the set of referents or at least encourage them to imagine themselves as belonging to the set. The "lingering" reference to the speaker's interlocutor gives 'you' – compared to alternative impersonal constructions – a "sense of informal camaraderie" (Kitagawa and Lehrer 1990: 752) and of "recruiting involvement, here empathetic, from the addressee" (Stirling and Manderson 2011: 1583). Much has been written about other possible effects (e.g. O'Connor 1994: 45 on distancing; Kluge 2016: 515-518 on politeness) but, though interesting, they fall outside of the present article's scope. We will also not consider the usage of the second person singular to describe personal experiences, as in (16). For a range of potential reasons (see Stirling and Manderson 2011: 1590-1598), the speaker chooses *you* here instead of *I* to somehow generalize her own specific sensation.

(16) I knew the bandage was there, *you* could feel it and everything.
(Stirling and Manderson 2011: 1593)

What the UNI-INT part of the map in Figure 3 *is* concerned with for 'you' is different contexts where "structural knowledge" (see Kitagawa and Lehrer 1990: 748) is described. The speaker conveys not what happens to themselves but what holds for anyone or what commonly befalls (or can/should befall) anyone in a situation. The fact that the parameters differentiating these contexts, i.e. (N)VER and (N)MOD, may indeed affect the impersonal second person singular is evident from research mentioned earlier (e.g. Laberge and Sankoff 1979 in Section 1.1).

The third person plural functions as an impersonal construction in numerous languages (e.g. Fernández 2013: 101-104 on Spanish; Data-Bukowska 2018 on Swedish; Van Olmen and Breed 2018a: 17-19 on Afrikaans). English and Romanian are among those languages, as *they* in (3) – repeated here as (17) – and *au* 'they have' in (18) make clear.

(17) *They*'ve taken the radio. (Myhill 1997: 802)

(18) *Au* furat radio-ul.
have.3PL.PRS steal.PST.PTCP radio-DEF.SG.M
'They've taken the radio.'

Impersonal 'they' is often characterized as a feature of informal language (e.g. Kitagawa and Lehrer 1990: 746) and is generally found in non-UNI-INT contexts only (e.g. Siewierska and Papastathi 2011; Gast and van der Auwera 2013: 144). Exceptions exist, however. According to Cabredo Hofherr (2006: 246), in Russian and Hebrew, for instance, the third person plural may also receive a UNI-INT interpretation. There is variation within the non-UNI-INT domain too, of course, as amply shown in the discussion in Section 1.1. Languages diverge not only in (not) employing 'they' for SAV but also in its (non-)restriction to the semi-impersonal UNI-EXT and EXI-COR uses or – if it can occur in truly impersonal EXI contexts – its (in)compatibility with NN readings or different forms of (un)knownness. The explanation for such differences is argued, by Siewierska and Papastathi (2011: 600) among others, to lie in the degree to which "the definite plural prototype" of 'they' as a personal pronoun has bleached in languages. UNI-EXT and EXI-COR, for instance, can be regarded as close to the prototype as the set of referents is plural and quite definite (see Siewierska and Papastathi 2011: 583). In (6a), here as (19a), *they* refers to an entire group of people that *in Bali* makes accessible, i.e. the country's inhabitants. In (8a), repeated as (19b), the state of affairs itself renders the collective entity of referents partially accessible, i.e. the government. EXI-INF and EXI-SPE, by contrast, may be considered as further removed from the prototype. Not only is "referent identification … purely situational" (Siewierska and Papastathi 2011: 584) in (11b) and (11c), here as (19c) and (19d), i.e. based on some smell in the former and the speaker's here and now in the latter; *they* seems NN in the examples in that it may refer to one or more than one person.

(19) a. In Bali, *they* eat dragonflies. (Gast and van der Auwera 2013: 144)
b. *They*'ve raised the taxes again. (Gast and van der Auwera 2013: 152)
c. *They*'ve been frying chips here. (Siewierska and Papastathi 2011: 581)
d. *They*'re knocking on the door. (Siewierska and Papastathi 2011: 581)

Siewierska and Papastathi (2011) also observe that, on the whole, languages with an overt third person plural pronoun (e.g. French) exhibit a narrower range of uses, more in keeping with the prototype, than languages without one or, put differently, with a pro-dropped one (e.g. Italian).[4] To account for these facts, they appeal to the established "relationship between morphophonological reduction and semantic bleaching" (Siewierska and Papastathi 2011: 600): less reduced, overt 'they' is more resistant to appearing in contexts that are not definite or plural. There are exceptions, however. Polish, for one, is found to pattern "with the non-pro-drop languages in apparently lacking inferred and specific 3pl IMP[ersonal]s and disfavoring vague ones" (Siewierska and Papastathi 2011: 602).

The (agentless) passive is known to serve as an impersonal construction in countless languages (e.g. Shibatani 1985; Creissels 2008a: 9-10). The English one can be used in this way too and so can the Romanian one, as in (4b) – repeated here as (20) – and (21).

(20) The radio has *been taken*.

[4] Henceforth, we will use "pro-drop" when an impersonal pronoun is conveyed solely through verb agreement, as in (18), and "overt" for its realization as a separate word, as in (19).

(21) Radio-ul a *fost* *furat*.
radio-DEF.SG.M have.3SG.PRS be.PST.PTCP steal.PST.PTCP
'The radio has been taken.'

The passive with (or without) an impersonal interpretation is often said to be a relatively formal phenomenon (e.g. Kitagawa and Lehrer 1990: 746). Moreover, Sansò's (2006: 245-263) in-depth study of five European languages reveals that it may occur in UNI contexts but is clearly a minority option there. Interestingly, Van Olmen and Breed (2018b: 819) independently find that "H[uman]I[mpersonal]P[ronoun]s are very strongly preferred to other forms of impersonalization in universal" uses in West Germanic. No such obvious preference appears to exist for the other contexts, however, which raises the question whether the passive is perhaps favored more in that domain. According to Siewierska (2008: 12), this is indeed true for SAV in British English: she claims that *it is said that* is more common than *they say that*. The passive is also argued to be preferred more generally to the third person plural when an unidentified agent is almost definitely a single person (Siewierska 2008: 16; cf. the NN distinction in Section 1.1.). Yet, to our knowledge, the matter of how the passive tends to fare in languages in more particular uses such as EXI-COR or EXI-VAG-PL versus EXI-INF/SPE-PL has so far not been examined in much detail.[5]

Languages may possess other impersonal constructions than those described here. 'Man'-pronouns, deriving from a noun meaning 'man, human being' and exemplified in (5) and (10), are a case in point. Both English and Romanian lack such a pronoun, however (see Siewierska 2011). The first person plural, as in (22), is occasionally mentioned too (e.g. Siewierska 2004: 211) but seems to have a very low frequency in languages (see Van Olmen and Breed 2018b: 818 on Afrikaans, Dutch and English). It is therefore not included in the present study (see also Kitagawa and Lehrer 1990: 744-745 on the ambiguity between the different interpretations of *we*).

(22) *Avem* doar o viață.
have.1PL.PRS just INDF.F.SG life
'We only live once.'

Another construction found in many languages is the reflexive impersonal (e.g. D'Allessandro 2007; Creissels 2008a: 8-9; Siewierska 2008: 18-21). It exists in Romanian, as the UNI-EXT context with reflexive *se* in (23) shows, and is a conventional way of expressing impersonalization in the language (e.g. Dobrovie-Sorin 1998; Pană Dindelegan 2008).

(23) În Anglia, *se* conduce pe stâng-a.
in England REFL drive.3SG.PRS on left-DEF.F.SG
'In England, they drive on the left.'

It will nevertheless not be considered here (and neither will any other possible impersonal constructions), for the contrastive reason that it has no genuine or even approximate equivalent in English, unlike the passive.[6]

[5] The factors discussed here are the ones relevant for the present study but they are not the only ones affecting the passive's potential use as an impersonal constructions. Languages where passivization requires the promotion of an object to subject, for example, simply do not have the passive as an option for intransitive clauses (cf. Creissels 2008a: 9-10).

[6] Note, for transparency's sake, that reflexive impersonals were included in the questionnaire given to our Romanian participants but that we leave the analysis of this data for further research.

1.3 Languages

Impersonalization has attracted significant interest the last decade. Especially impersonal pronouns have been the subject of much research. Recent work on English, for instance, includes Rudolf (2016), Haas (2018a) and Hall (2020). Other Germanic languages that have been studied are Afrikaans, Danish, Dutch, German and Swedish (e.g. Coussé and van der Auwera 2012; De Hoop and Tarenskeen 2015; Gast 2015; Jensen and Gregersen 2016; Fenger 2018; Van Olmen and Breed 2018a). Romance has received considerable attention too, with publications on French (Creissels 2008b), Portuguese (Posio and Vilkuna 2013) and Spanish (Fernández 2013) among others. Romanian, however, seems to have escaped notice.

No comprehensive account of the usage of impersonal pronouns in Romanian exists, to our knowledge. Reference grammars such as Cojocaru (2003) and Gönczöl-Davies (2008) do not even discuss them. Manea (2012: 70) does mention a "generic-passive expression" that was used in an unspecified "older Romanian", illustrated in (24), but does not expand on it.

(24) *Spun* că (el) ar fi zis… (Manea 2012: 70)
say.3PL.PRS COMP 3SG 3.COND be.INF claim.PST.PTCP
'They say that he (would have) claimed…'

Yet, *spun* 'they say' is not actually a passive. Rather, it is an instance of Siewierska and Papastathi's (2011) SAV function of the third person plural. The questions thus arise as to whether it is still acceptable nowadays and whether Romanian 'they' can appear in the other UNI-EXT and EXI uses identified for the third person plural in the cross-linguistic research. In the same vein, Pană Dindelegan's (2008: 147) chapter in the Romanian Academy's grammar, about passivization and impersonalization, points out the existence of a generic second person singular and a generic first person plural but only gives examples of the latter. It remains to be seen how acceptable such pronouns are in the different contexts distinguished for the UNI-INT realm in the literature on other languages.

In the present article, we will map the functional potential, i.e. the degree of (in)compatibility with various uses, of Romanian 'they' and 'you' in particular (for 'we', see Section 1.2) and contrast it to that of the impersonal pronouns in English. The inclusion of English allows a comparison of Romanian to a widely known language with a typical division of labor between a UNI-INT second person singular and a non-UNI-INT third person plural. It will also enable us to check the findings of a previous, more limited study of the English pronouns (Van Olmen and Breed 2018b). Besides pronominal forms of impersonalization, this article will investigate the passive, whose usage – we have argued – has not yet been thoroughly investigated in light of all the contexts in Figure 3. In short, we primarily aim to answer the following questions: (i) what is the functional potential of the impersonal pronouns as well as the passive in Romanian and English?; and (ii) how do the impersonal constructions and the languages compare to each other? In addition, our study may shed more light on two other issues raised above. On the one hand, Romanian can function as a non-Germanic testing ground for the interaction attested in Afrikaans, Dutch and English between number and (un)knownness in the truly impersonal EXI domain (see Section 1.1). On the other hand, comparing Romanian and English 'they' can serve to check the hypothesized impact of the third person plural's pro-dropped versus overt nature on its usage (see Section 1.2). In other words, we also seek to answer the following secondary questions: (i) do the parameters of number and (un)knownness interact with one another (in the expected way) in Romanian?; and (ii) does the pro-dropped third person plural exhibit a(n expected) wider range of less definite/plural impersonal uses than its overt counterpart?

The remainder of the article is structured as follows. Section 2 will introduce our methodology: the choice for an acceptability judgment questionnaire approach (2.1), the design of

our questionnaire (2.2), details about its distribution and participants (2.3) and some statistical background (2.4). In Section 3, we will present and discuss our findings, first for the UNI-INT uses (3.1) and then for the non-UNI-INT ones (3.2), and subsequently summarize them (3.3). Section 4, finally, is our conclusion.

2 Methodology[7]

2.1 Acceptability judgment questionnaires

Acceptability judgments are usually regarded as signs of grammaticality. Yet, Bard et al. (1996: 33) argue that they "need not be one-to-one reflections of grammaticality" since they "may be based, for example, on estimated frequency of usage, on conformity to a prescriptive norm or a prestigious register, or on degree of semantic and pragmatic plausibility". This suitability in specific semantic and pragmatic contexts is what is at play in previous acceptability judgment questionnaires for impersonalization, such as Siewierska and Papastathi (2011) and Van Olmen and Breed (2018a, 2018b). They ask native speakers to rate the acceptability of different constructions in sentences that conclude short scenarios triggering particular impersonal interpretations.

This approach has a number of advantages, despite the constructed nature of the language (a drawback that a corpus study would not suffer from). If formulated carefully, such scenarios allow researchers to exclude personal readings of constructions like 'you' in advance (cf. Siewierska and Papastathi 2011: 587-588 on some of the issues in distinguishing personal and impersonal 'they' in corpora). Consider the UNI-INT-NVER-NMOD context in (25): the fact that the speaker is replying to their brother and it is their cousin who is going away makes a personal interpretation of *you* rather unlikely here.

(25) You and your brother are helping your female cousin pack her suitcase for her trip to England. When you hand her the brand new raincoat that you bought her the day before, your brother asks you why you bought that in particular. You reply: "If **you** go to England, it is necessary to pack a raincoat."

(Van Olmen and Breed 2018a: 7)

The set-up of these questionnaires also provides more control over the parameters differentiating the various uses in Figures 1 to 3. The distinction between PL and NN contexts, for instance, may not always be so clear in authentic language data. By contrast, one can quite easily develop scenarios that prompt an unambiguous PL or NN reading (see Section 2.2), making it possible to test the impact of the parameter of number accurately. Moreover, the approach enables researchers to check, in a straightforward way, a construction's potential to occur in contexts that may not actually arise in, say, a corpus due to their general infrequency. For example, EXI uses are known to be much less common than UNI-ones, for 'man'-pronouns at least (e.g. Zifonun 2001 on German; Fonesca-Greber and Waugh 2003 on French). Similarly, Siewierska and Papastathi's (2011: 590) corpus study reveals that EXI-INF and EXI-SPE cases of 'they' are extremely rare.

The present questionnaire differs from the earlier ones in several regards. Siewierska and Papastathi's (2011) is primarily concerned with the third person plural and pays little attention to other impersonal constructions. This focus means that it only contains scenarios for the UNI-EXT, SAV and EXI uses in Figure 2. It also does not feature contexts that combine the parameters of number and (un)knownness (e.g. EXI-VAG-PL, EXI-INF-NN; see Section 2.2) to test

[7] The methodology described here closely resembles that of Van Olmen and Breed (2018a, 2018b).

their interaction. Furthermore, Romanian is not amongst the languages under investigation and the findings for English are based on "just" ten participants. Van Olmen and Breed's (2018a, 2018b) questionnaires do examine all the uses in Figure 3 but only look at impersonal pronouns in Afrikaans, Dutch and English. The present one brings in Romanian, adds the passive as one of the constructions to be assessed and has many more English-speaking participants than Van Olmen and Breed's (2018b) "mere" 23.

2.2 Questionnaire design

Every part of the questionnaire starts with a short scenario that prompts a particular impersonal reading of the test items in its final sentence. The scenario in (26) can serve as an example. It is also representative of the other scenarios in describing fairly informal situations and in often inviting the reader to take the perspective of the speaker. The reason for this type of scenario is that the use of especially impersonal pronouns appears to be more typical of everyday language (see Siewierska and Papastathi 2011: 593).

(26) UNI-EXT
Your friend is making a list of the official languages in every country. Based on your recent traveling experience, you tell her:

"In Brazil, one speaks Portuguese."	1 2 3 4 5
"In Brazil, you speak Portuguese."	1 2 3 4 5
"In Brazil, they speak Portuguese."	1 2 3 4 5
"In Brazil, Portuguese is spoken."	1 2 3 4 5

Participants are asked to go through each scenario and rate the different options for ending it on a scale from 1 to 5, with 1 standing for 'very unacceptable' and 5 for 'very acceptable'. We are aware that the choice of a Likert scale is not uncontroversial. For one thing, it assumes that there are (only) five clear-cut levels of acceptability. Still, a supposedly better alternative like magnitude estimation (see Bard et al. 1996) has been criticized too in recent years (e.g. Sprouse and Almeida 2012) and has even been shown to produce basically the same results as the Likert scale (see Van Olmen and Breed 2018b: 803 for details). The latter method also has the advantage that participants may be expected to be familiar with it.

As (26) makes clear, the options provided for English in each scenario are the dedicated impersonal pronoun *one*, the second person, the third person plural and the passive. For Romanian, every scenario includes the second person singular and the third person plural marked by verbal inflection only, the overt third person plural and the passive, as in (27).

(27) UNI-INT-NVER-MOD
Prietena ta a luat lecții de spaniolă de doar două săptămâni, când tu îi ceri să traducă un text pentru tine. Ea e un pic supărată și îți răspunde:
'Your friend has only been taking Spanish classes for two weeks when you ask her if she could translate a text for you. She is a little annoyed and says:'

"Nu se poate învăța o limbă străină în două săptămâni."	1 2 3 4 5
'"You cannot learn a foreign language in two weeks."'	
"Nu pot învăţa o limbă străină în două săptămâni." [cu sensul de ei, nu eu]	1 2 3 4 5
'"They cannot learn a foreign language in two weeks."' [meaning they, not I]	
"Ei nu pot învăţa o limbă străină în două săptămâni."	1 2 3 4 5
'"*They* cannot learn a foreign language in two weeks."'	

"Nu poți învăța o limbă străină în două săptămâni." 1 2 3 4 5
'"A foreign language cannot be learned in two weeks."'

Three comments are in order. First, the third person plural can have the same verb form as the first or the third person singular in Romanian. When this syncretism makes a sentence ambiguous, the intended meaning is indicated in brackets at the end, as with *[cu sensul de ei, nu eu]* '[meaning they, not I]' in (27). Second, as previously mentioned, overt subject pronouns only occur in a pro-drop language like Romanian for, say, emphasis, contrast or disambiguation and do not ordinarily tolerate impersonal interpretations (see Pană Dindelegan 2013: 104-111). The rationale behind nevertheless inserting a clause with overt *ei* 'they' in (27) – and not one with overt *tu* 'you (singular)' – has to do with the comparison of the third person plural's functional potential in a pro-drop versus a non-pro-drop language (see Sections 1.2 and 1.3). If Romanian and English are found to differ in this respect in the hypothesized manner, the expected low acceptability of overt *ei* will highlight that the dissimilarities are due to the pro-drop character of 'they' as an impersonal pronoun in Romanian. Third, since the present questionnaire looks at the passive as well as at impersonal pronouns, all active sentences with impersonal pronouns need to be and are transitive in both languages, as (26) and (27) show. This requirement ensures that an equivalent passive can be added as an option.

Our questionnaire contains twelve scenarios in total, one for each impersonal use distinguished in Van Olmen and Breed's (2018b) semantic map in Figure 3, repeated here as Figure 4 (see Section 1.1 for its make-up and abbreviations).

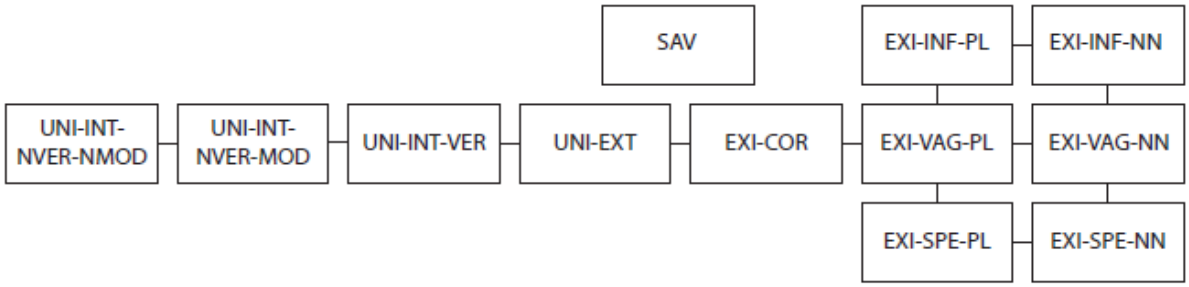

Figure 4: Van Olmen and Breed's (2018b: 840) combined semantic map of impersonalization

The three leftmost uses are exemplified in (27) to (29).[8] They all involve UNI quantification in that they are about people in general. Impersonal pronouns in these contexts can thus be paraphrased as 'everyone' or 'anyone'. The three uses also share an INT perspective, which typically means that speaker and/or addressee are among the people in general. In (27), for instance, the statement about not being able to learn a language in two weeks is intended as applying to anyone, including the speaker and their friend. The differences between (27) to (29) lie in veridicality – i.e. the proposition is (not) assumed to be true – and modality – i.e. the non-veridical proposition has an/no explicit marker of possibility or necessity. The former distinguishes (29) from (27) and (28): not wearing trainers at a formal party is presented as an unqualified fact of life here, making the sentence VER. The NVER nature of (27) is due to the occurrence of MOD *can't*, which sets learning a language in two weeks in the realm of (im)possibility. In NMOD (28), there may be no modal marker but the conditional subordinate clause still renders eating rotten eggs hypothetical and therefore NVER.

[8] For reasons of space, the remaining examples in this section will be in English only. Their Romanian equivalents can be found in the online questionnaire: https://forms.gle/cB7okX7Jpoe9UdkJ8. To access the English questionnaire, see: https://forms.gle/9Y2bd7xVrYD1yQZY9.

(28) UNI-INT-NVER-NMOD
Your sister cracks open two eggs and pours their contents into a bowl to make an omelette. A bad smell is coming from the bowl. Your sister frowns at the eggs and then says to you:
"What happens if one/you/they eat/eats rotten eggs?"
"What happens if rotten eggs are eaten?"

(29) UNI-INT-VER
Your partner tells you they want to buy a new pair of trainers for your son to wear at his graduation ball. You completely disagree and retort:
"One/you/they never wear/wears trainers at a formal party."
"Trainers are never worn at a formal party."

Like (27) to (29), the scenario in (26) introduces a UNI context: people at large in Brazil are said to speak Portuguese. Unlike (27) to (29), it takes an EXT rather than an INT point of view. As their friend is listing the official languages of countries, the speaker's contribution makes sense only as a statement about the inhabitants of Brazil and not as one about anyone, including the interlocutors, who happens to be or can imagine themselves in the country. In short, (26) is an instance of the UNI-EXT use in Figure 4.

The next two uses in Figure 4, from left to right, are tested with (30) and (31). The SAV context in (30) features the speech act verb *say* and marks information as originating from some unidentified source. The context in (31) involves EXI quantification, i.e. presuming the existence of a distinct (set of) individual(s) responsible for the event described: the rising prices are attributed not to people in general but to a particular, though unidentified, entity.

(30) SAV
Your colleague is about to go on holiday and visit a renowned fortress. You have heard that the fortress is crawling with ghosts, and you tell your colleague:
"One/you/they say/says that the fortress is haunted."
"It is said that the fortress is haunted."

(31) EXI-COR
You go to the petrol station and notice the higher price for petrol. When you get home, you tell your daughter:
"One/you/they have/has raised the petrol prices again."
"The petrol prices have been raised again."

What differentiates (31) from the other EXI uses – in (32) to (37) – is that the referents may be deemed partly known here (as in UNI-EXT). The event itself hints at the typically institutional or, in Figure 4's terms, COR(porate) plural-like entity behind it. To be more precise, the rising petrol prices can only really be the fault of the government or the oil companies/suppliers.

The last six uses combine the parameters of (un)knownness and number. The former has three values – i.e. VAG(ue), INF(erred) and SPE(cific) – and the latter two – i.e. PL(ural) and N(umber-)N(eutral). To explain their combinations, let us consider (32) to (34). In (32) as well as (33), the speaker knows that the event took place but cannot or will not identify the people responsible. Both cases are, in other words, VAG. They differ in number, however, as lynching can never be done alone while finding a wallet can, making (32) EXI-VAG-PL and (33) EXI-VAG-NN. The context in (34) has its number value in common with (32): one can eat pizza on one's own. It is not VAG, though, but INF: the speaker does not actually know that someone or some people ate pizza. The smell in the air just makes them hypothesize that this event

happened. Scenario (34) thus triggers an EXI-INF-NN interpretation.

(32) EXI-VAG-PL
Your friend asks you about the latest film you watched: "What happened to the protagonist in the end?". You reply:
"One/you/they lynched him."
"He was lynched."

(33) EXI-VAG-NN
You ring up your friend to tell them that the wallet that they lost the day before has been found. You say:
"One/you/they have/has found your wallet in the park."
"Your wallet has been found in the park."

(34) EXI-INF-NN
After finishing lectures for the day, you and your course friend are walking back to the common kitchen to have a late dinner. When you get there, you notice a particular smell in the air and tell your friend:
"One/you/they have/has eaten pizza in here."
"Pizza has been eaten in here."

In (35) to (37), finally, the scenarios for the three remaining uses are presented. The EXI-INF-PL context in (35) shares with (34) that the event, its actual occurrence and the existence of the people responsible for it are inferred from certain signs – the litter in the park here. Unlike the action of eating pizza in (34), however, that of having a party normally involves more than one person. In scenarios (36) and (37), the events of sending money and making a call are happening in the here and now of the speaker. The individuals performing these actions are nevertheless not explicitly identified, making both contexts EXI-SPE. They differ in number, though: (36) is PL as only multiple people can send money from all around the world while (37) is NN since a phone call may but need not involve more than one person.

(35) EXI-INF-PL
Your friend comes to visit your hometown and you take him for a walk in the local park. There, you see many empty bottles on the grass, together with some leftover food and used napkins. You tell your friend:
"One/you/they have/has thrown a party here."
"A party has been thrown here."

(36) EXI-SPE-PL
You and your friend have created an online donations page for the environmental organization you're working for. When you check the page the next day, you exclaim to your friend: "Wow! Come and have a look! …
One/you/they are sending us money from all around the world."
Money is being sent to us from all around the world."

(37) EXI-SPE-NN
Your sister is in the shower when her phone starts ringing at the same time. You tell your sister: "Hurry up! ...
One/you/they is/are calling you on the phone!"
You are being called on the phone!"

2.3 Implementation

The same scenarios were presented in the same random order in the Romanian and the English questionnaire. Both were created in Google Forms and checked by a(nother) native speaker of the language before distribution, resulting in a few minor corrections with no significant effect on the comparability of the two questionnaires. After receiving ethics approval from Lancaster University, we promoted links to our Google Forms via Facebook and Whatsapp, emphasizing that participation was voluntary and anonymous. It was also made clear at the start of the questionnaire that some demographical data would be requested (i.e. age, gender, native language(s) and language variety) and that participants could ask to have their answers removed within two weeks of completion (based on their demographical information). In the instructions, we highlighted that: (i) there are no right or wrong answers; (ii) participants should rely on their own linguistic intuition and not on what might be (in)correct according to certain norms; (iii) there was no time limit; and (iv) participants should make full use of the five-point scale.

The English questionnaire was filled out in December 2019 – January 2020 by 62 native speakers, of which two were left out because they were over 35 years old.[9] For Romanian, in the same period, 77 native speakers completed the questionnaire, of which 16 were excluded because of age considerations and one because they did not follow the instructions of making full use of the scale. About two thirds of the 18- to 35-year-old Romanian participants identified as female and one third as male while, among the English ones, there was essentially a fifty-fifty split between male and female, with one participant identifying as "other". We have, however, no reason to believe that gender had an impact on our results. Furthermore, most participants did not identify as speaking any particular variety of English or Romanian. For English, the majority of participants who did answer the question wrote "British" or "British English", which was to be expected as the questionnaire was primarily distributed among the first author's contacts at Lancaster University. For Romanian, there were not enough answers to determine a majority. Since the link to the questionnaire was mainly disseminated to the first author's contacts in Romania, it is not unreasonable to assume that the non-respondents to the question about language variety consider themselves speakers of "standard" Romanian. The same holds for the English non-respondents, of course.

2.4 Statistics

We will give, for each impersonal construction in each scenario, the following descriptive statistics: (i) the mean, i.e. the average acceptability score of the construction in question in the impersonal use in question; (ii) the standard deviation (std), i.e. the extent to which the scores for the construction in question in the use in question vary between all participants.

To compare one impersonal construction in two contexts (e.g. the Romanian passive in EXI-VAG-PL versus EXI-VAG-NN) or two constructions in the same context (e.g. *you* versus *they* in UNI-EXT), two-sided t-tests will be used – in order to avoid making any assumptions about the direction of any difference (see Baayen 2008: 8). These t-tests will be of the so-called dependent kind for findings from the same group of participants, i.e. when contrasting impersonal constructions in the same language (e.g. *you* versus *one* in UNI-INT-VER). But they will be of the independent kind for results from different groups of people, i.e. impersonal constructions in English versus Romanian (e.g. *you* versus the second person singular in Romanian in UNI-INT-VER) (see Rasinger 2013: 20). If data is compared multiple times with these tests,

[9] We are not aware of any generational differences in the use of impersonal constructions. We simply wanted to ensure that our Romanian and English participants are as comparable as possible.

we will apply a Bonferroni correction to prevent any type I errors (see Baayen 2008: 114). Put differently, our standard threshold for significance of 0.05 – used for the "simple" comparison of two results (e.g. the Romanian passive versus pro-dropped 'they' in SAV) – will be divided by the number of tests undertaken, to ensure that we do not give too much weight to the outcome of one particular test. Examining *one* across the UNI-INT domain, for instance, requires three contrasts (i.e. NVER-NMOD versus NVER-MOD, NVER-MOD versus VER and VER versus NVER-NMOD) and, accordingly, our threshold for significance will be 0.017, which is 0.05 divided by three. In the remainder of the article, we will always refer to the threshold that is required for the type of comparison at issue.

Finally, to evaluate the influence of one or two variables on an impersonal construction's acceptability, we will compute one- or two-way analyses of variance (ANOVAs; see Rasinger 2013: 210-217), for which our threshold for significance is 0.05. The impact of the parameters of number and (un)knownness on the acceptability of the third person plural, for example, will be tested with a two-way ANOVA and, to be more precise, one with repeated measures as the data is from a single group of participants (see Baayen 2008: 264). Afterward, post-hoc Bonferroni-corrected t-tests will be used to ascertain significant differences between specific contexts (e.g. *they* in EXI-SPE-PL versus EXI-SPE-NN or pro-dropped *ei* in EXI-VAG-PL versus EXI-SPE-PL).

3 Results

As mentioned in Section 1.3, the present section will first look at the UNI-INT uses (3.1), then discuss the non-UNI-INT ones (3.2) and, lastly, give a summary of the results (3.3).

3.1 Universal-internal uses

3.1.1 In general

Table 1 gives the means and standard deviations of the impersonal pronouns and the passive in Romanian and English in all UNI-INT contexts. Note that, for convenience's sake, *tu* here and elsewhere stands for the second person singular expressed by verbal inflection, not for the overt pronoun.

Table 1: Descriptive statistics of the UNI-INT uses

Use		Romanian				English			
		tu	*ei* (pro-drop)	*ei* (overt)	Passive	*one*	*You*	*they*	Passive
UNI-INT-NVER-NMOD	mean	4.90	1.58	1.45	1.63	2.80	4.73	1.63	3.05
	std	*0.35*	*0.94*	*0.83*	*0.88*	*1.26*	*0.61*	*0.80*	*1.05*
UNI-INT-NVER-MOD	mean	4.75	1.37	1.18	4.50	2.90	4.43	1.40	3.97
	std	*0.54*	*0.74*	*0.62*	*0.83*	*1.30*	*0.79*	*0.72*	*0.84*
UNI-INT-VER	mean	4.13	1.32	1.15	3.87	2.93	4.30	2.23	4.00
	std	*1.21*	*0.70*	*0.40*	*1.14*	*1.35*	*1.05*	*1.14*	*0.84*

Let us look at the third person plural in both languages first. The numbers for *they*, pro-dropped *ei* and overt *ei* are so low in all three uses in Table 1 that we can probably safely reject them as UNI-INT impersonal constructions without any statistical analysis. *They*'s average score across these contexts is 1.75, pro-dropped *ei*'s 1.42 and overt *ei*'s 1.26, suggesting that the third person plural is, essentially, unacceptable in the UNI-INT domain. Romanian 'they' is, in this regard,

in keeping with what we know from most other languages (see Section 1.2, where only Russian and Hebrew are mentioned as exceptions to the third person plural's characteristic incompatibility with UNI-INT uses). 'They' will therefore not be considered any further in this section (but see Section 3.2).

The remaining data is presented in a more accessible way in Figure 5. Green is used for Romanian and orange for English whereas full lines with round markers indicate 'you', dashed lines with square markers the passive and the double line with a triangular marker 'one'.

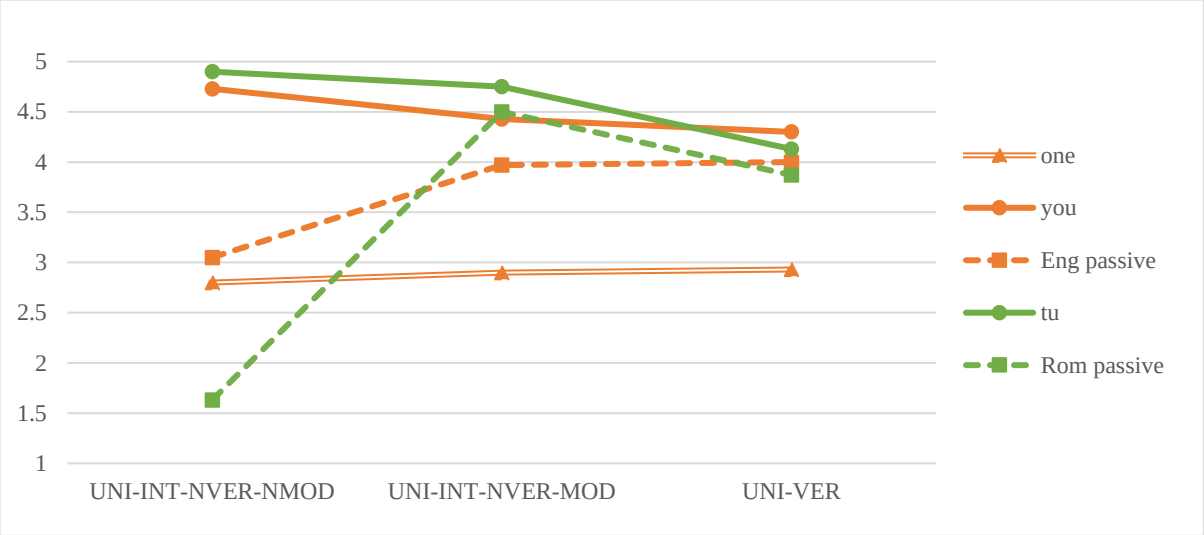

Figure 5: Distribution of the means of 'you', 'one' and the passive for the UNI-INT uses

3.1.2 'You' and 'one'

'You' is, in absolute terms, the most acceptable impersonal construction in both Romanian and English. *Tu* and *you* also do not differ much from one another: only in UNI-INT-NVER-MOD is there a small but significant difference in acceptability between the two (4.75 versus 4.43; $p<0.05$). The way in which they vary across the three uses is very similar too: both are slightly but substantially less acceptable in UNI-VER than in at least one of the other universal-internal contexts (4.13 versus UNI-INT-NVER-MOD 4.75 and UNI-INT-NVER-NMOD 4.90 for Romanian; 4.30 versus UNI-INT-NVER-NMOD 4.73 for English; $p<0.017$ in each case). The reasons for these minor dissimilarities are not entirely clear to us. Still, the fact that both *tu* and *you* score highest for UNI-INT-NVER-NMOD could be argued to be in line with Laberge and Sankoff's (1979: 432) findings for French, where *tu* is preferred (to *on*) in conditional contexts (see Section 1.1). Notwithstanding the small differences, the overall picture of the second person (singular) as an all-purpose UNI-INT impersonal construction is consistent with previous research (see Section 1.2). One explanation for its suitability for this domain of impersonalization lies in its features as a personal pronoun. 'You' is normally used to refer to the addressee. When it occurs in a generic sentence not immediately or solely relevant for that person, they will still be invited to transpose themselves into the situation through mental simulation, producing an INT perspective (see Gast et al. 2015).

Of the constructions that can appear in UNI-INT contexts in English, *one* can be argued to be the least acceptable. Its scores, averaging 2.87, are significantly lower than those of *you* and the passive, averaging 4.49 and 3.67 respectively, in all uses ($p<0.017$ in all cases), except for the latter's UNI-INT-NVER-NMOD score (see Section 3.1.3). This result confirms an earlier one about *one* compared to *you* by Van Olmen and Breed (2018b: 815), who attribute it to its "formal character … clash[ing] with the colloquial nature of the scenarios". Note also that *one*'s acceptability is unaffected by the type of UNI-INT context, as a one-way ANOVA shows

($p>0.05$). This fact supports Haas's (2018b) claim that any correlation between the use of *one* and non-veridicality has weakened/disappeared (see Section 1.2).

3.1.3 The passive, versus 'you'

For the passive, we will first examine the UNI-INT-NVER-MOD and UNI-INT-VER contexts. The construction scores relatively high in these uses in both languages, with a small but significant drop in acceptability from UNI-INT-NVER-MOD to UNI-INT-VER in Romanian (from 4.50 to 3.87; $p<0.017$) but not in English (3.97 and 4.00 respectively; $p>0.017$). Furthermore, in UNI-INT-NVER-MOD, both passives are less acceptable than 'you' ($p<0.05$ in both cases) whereas, in UNI-INT-VER, they are actually on par with the second person (singular) ($p>0.05$ in both cases).[10] This finding indicates that the supposed preference for impersonal pronouns and relative aversion to passives in the UNI-INT domain (see Section 1.2) may at least partially be affected by veridicality or modality. This idea seems to be confirmed by our data for UNI-INT-NVER-NMOD. The passive is significantly less acceptable here than in the other two UNI-INT uses ($p<0.017$ in all cases) and than 'you' ($p<0.05$ in both cases) in Romanian and English, with much lower scores of 1.63 and 3.05 respectively. Interestingly, Gast (2015) too finds, in his corpus study of the English translated and original equivalents of German *man*, that combining veridicality and modality interacting with quantification (alongside the original or translated status of *man*) predicts the choice of *man*'s English counterpart the best. Unfortunately, the concrete impact of these variables is not discussed in detail. Still, Gast's (2015: 22) association plot for (non-)veridicality suggests, for one, that, in NVER (versus VER) contexts, there are fewer passives and more instances of *you* than expected in view of the overall distribution of impersonal constructions. Our data may be pointing in the same direction, in that the acceptability of English and Romanian 'you' is significantly higher than that of the passive in the two NVER uses, providing some supporting evidence for the aforementioned dispreference for passives in this domain (see also Sansò 2006: 245-263), but not in the UNI-INT-VER one. Our figures would then also signal a potentially even stronger preference for the second person (singular) to the passive in UNI-INT-NVER-NMOD (versus -MOD) contexts.

A closer look at the UNI-INT-NVER-NMOD scenario reveals that veridicality and modality are of course not the only factors at play in the competition between 'you' and the passive. Our intuition is that the latter's scores for this use would be much higher if its subject, the active object, was more animate and/or more definite, i.e. if the scenario tested a sentence like 'what happens if animals are treated inhumanely?' instead of 'what happens if rotten eggs are eaten?' (see Section 2.2). The inanimate and indefinite nature of the active object here seems especially problematic in Romanian, given the fairly large difference in the passive's acceptability between this language (1.63) and English (3.05) ($p<0.05$). It is important to add, however, that the UNI-INT-NVER-MOD and UNI-INT-VER scenarios contain inanimate and indefinite passive subjects too – i.e. 'a foreign language' in 'a foreign language cannot be learned in two weeks' and 'trainers' in 'trainers are never worn at a formal party' respectively – and that, in these contexts, Romanian and English on the one hand and the passive and 'you' on the other are not too dissimilar. It is unclear to us why such passive subjects would be less problematic here and we leave the question for further research.

3.2 Non-universal-internal uses

3.2.1 In general

[10] The English passive does outperform *one* in both UNI-INT-NVER-MOD (3.97 versus 2.90) and UNI-INT-VER (4.00 versus 2.93) ($p<0.05$ in either case).

Table 2 provides the means and standard deviations of the impersonal pronouns and the passive in Romanian and English in all non-UNI-INT uses.

Table 2: Descriptive statistics of the non-UNI-INT uses

Use		Romanian				English			
		tu	*ei* (pro-drop)	*ei* (overt)	passive	*one*	*you*	*they*	passive
UNI-EXT	mean	2.10	3.53	2.60	2.40	2.57	3.10	4.82	3.57
	std	*1.19*	*1.27*	*1.17*	*1.17*	*1.13*	*1.24*	*0.43*	*1.01*
EXI-COR	mean	1.05	4.37	2.27	4.30	1.48	1.22	4.77	4.55
	std	*0.29*	*0.92*	*1.02*	*0.87*	*0.75*	*0.67*	*0.46*	*0.65*
EXI-VAG-PL	mean	1.03	4.63	2.50	4.67	1.32	1.08	4.25	4.67
	std	*0.18*	*0.64*	*1.28*	*0.68*	*0.62*	*0.33*	*0.95*	*0.73*
EXI-VAG-NN	mean	1.17	4.45	2.47	4.55	1.47	1.03	3.18	4.78
	std	*0.59*	*0.75*	*1.20*	*0.77*	*0.75*	*0.18*	*1.20*	*0.45*
EXI-INF-PL	mean	1.20	3.73	2.35	3.27	1.40	1.20	2.83	4.37
	std	*0.82*	*1.13*	*1.18*	*1.38*	*0.72*	*0.58*	*1.21*	*0.76*
EXI-INF-NN	mean	1.48	3.78	2.18	2.37	1.40	1.33	3.53	4.12
	std	*1.13*	*1.26*	*1.21*	*1.25*	*0.76*	*0.73*	*1.03*	*0.85*
EXI-SPE-PL	mean	1.03	4.30	2.05	3.15	1.27	1.08	4.02	4.57
	std	*0.18*	*0.98*	*1.06*	*1.27*	*0.58*	*0.28*	*0.98*	*0.65*
EXI-SPE-NN	mean	1.00	2.87	1.75	3.78	1.20	1.03	2.37	4.10
	std	*0.00*	*1.33*	*0.91*	*1.25*	*0.44*	*0.18*	*1.23*	*1.12*
SAV	mean	1.45	1.98	2.53	2.22	2.20	1.85	4.55	3.97
	std	*1.10*	*1.05*	*1.26*	*1.22*	*1.13*	*1.04*	*0.70*	*0.99*

It is clear from the numbers that 'you' and 'one' do not function as non-UNI-INT impersonal constructions. *Tu*'s scores range from 1.00 to 2.10, *you*'s from 1.03 to 3.10 and *one*'s from 1.20 to 2.57. These results are not surprising as the second person (singular) and 'one'-pronouns are known to be restricted to the UNI-INT domain (see Section 1.2). In fact, their highest scores in Table 2 are all for UNI-EXT and the likely reason for this incongruity is that certain participants read an INT perspective into (26), repeated as (38).

(38) UNI-EXT
Your friend is making a list of the official languages in every country. Based on your recent traveling experience, you tell her:
"In Brazil, one/you/they speak/speaks Portuguese."
"In Brazil, Portuguese is spoken."

Although the scenario strongly favors an EXT point of view (see Section 2.2), it is not inconceivable that, when actually seeing the sentences with 'you' and 'one', some participants somehow tried to include themselves and people at large in the interpretation (e.g. 'if one is in Brazil, one is forced to communicate in Portuguese') (see also Van Olmen and Breed 2018b: 823). In short, notwithstanding the slightly higher scores for UNI-EXT, the overall numbers justify not considering 'you' and 'one' any further here.

The same holds for overt *ei*, whose scores vary from 1.75 to 2.60 in Table 2. It is found to be substantially less acceptable than pro-dropped *ei*, averaging 3.76 compared to 2.26, in all non-UNI-INT contexts ($p<0.05$ for every t-test). This result shows that, in a pro-drop language like Romanian, it is the pro-dropped third person plural that serves as an impersonal pronoun and that impersonal *they* needs to be compared to pro-dropped and not overt *ei*. The only

exception is SAV: overt *ei* is significantly more acceptable in this use than its pro-dropped counterpart ($p<0.05$). It is tempting to attribute this finding to the "controversial" status of SAV as an impersonal context. As discussed in Section 1.1, in some languages, 'they' cannot be employed impersonally but may occur in SAV while, in other languages, it has a range of impersonal uses but does not appear in SAV. Siewierska and Papastathi (2011) therefore regard SAV as a direct off-shoot from personal 'they' in their semantic map and as unrelated to UNI-EXT and the EXI uses (see Figure 2). Our Romanian data could be seen as supporting evidence for this decision. However, both overt and pro-dropped *ei* score relatively low, with 2.53 and 1.98 respectively, suggesting that neither option is actually particularly acceptable in SAV.

Figure 6 shows the remaining data in a more accessible manner. Green, orange and the dashed lines with square markers again stand for Romanian, English and the passive respectively while the dotted lines with diamond markers represent the third person plural.

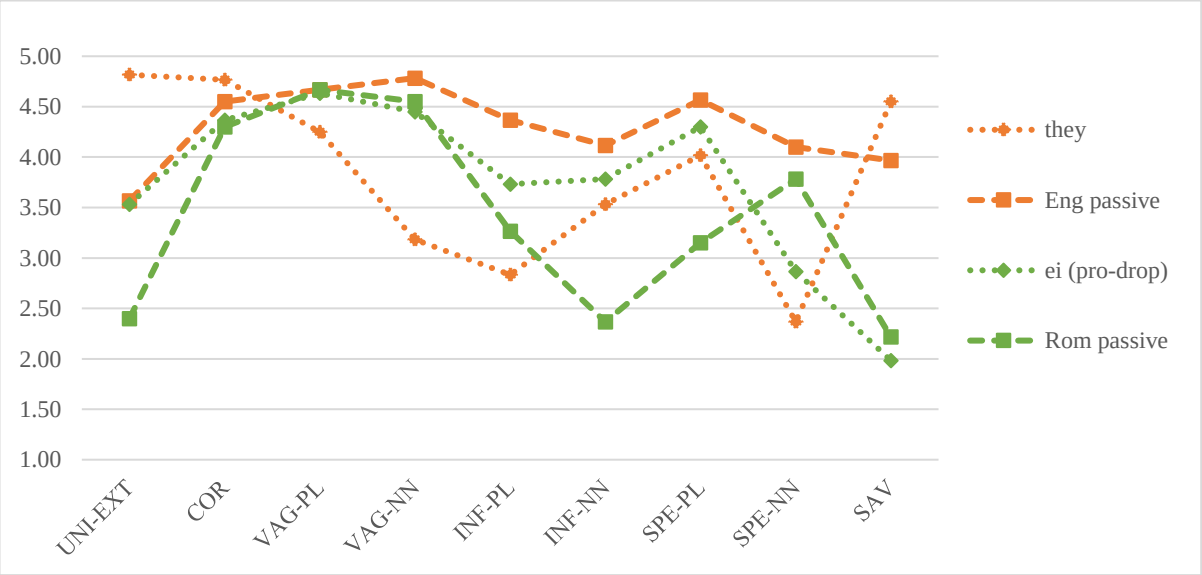

Figure 6: Distribution of the means of 'they' and the passive for the non-UNI-INT uses

3.2.2 'They'

Let us start with the third person plural. For English, a one-way ANOVA confirms what is clear from Figure 6: the acceptability of *they* varies significantly in the non-UNI-INT domain ($p<0.05$). Post-hoc Bonferroni-corrected t-tests reveal that it scores much higher for UNI-EXT (4.82) and EXI-COR (4.77) than for the other existential uses ($p<0.006$ in all cases). These two uses do not differ from SAV (4.55), though, whose score in turn is similar to that of EXI-VAG-PL (4.23) ($p>0.006$ for every test) but dissimilar to those of the remaining EXI uses ($p<0.006$ in each case). Van Olmen and Breed (2018b: 827) too find that *they* is most acceptable in SAV and in the two contexts where the referents are partly known (see Section 1.1). Our results thus corroborate that the third person plural's "overall speaker preference for [the so-called] semi-impersonal uses" (Siewierska and Papastathi 2011: 599), where its personal pronoun traits of definiteness and plurality can still be felt (see Section 1.2), manifests itself in English.

For the "purely" impersonal EXI contexts, a two-way ANOVA demonstrates that number, (un)knownness and their interaction all have a significant impact ($p<0.05$ in each case) on the acceptability of *they*. Moreover, post-hoc Bonferroni-corrected t-tests show that the significant differences between particular uses follow the logic of the map in Figure 4: if differences exist, 'they' will be more acceptable in PL than in NN contexts – as the former are closer to its personal trait of plurality – and more acceptable in VAG than in INF or SPE contexts – as the

former are arguably more definite than the latter two types with their situational identification. To be more precise, *they* scores significantly higher: (i) in EXI-VAG-PL (4.25) than in EXI-VAG-NN (3.18) and in EXI-INF-PL (2.83); (ii) in EXI-VAG-NN (3.18) than in EXI-SPE-NN (2.37); and (iii) in EXI-SPE-PL (4.02) than in EXI-SPE-NN (2.37) ($p<0.006$ in all cases).[11] On the whole, our results therefore confirm Van Olmen and Breed's (2018b) finding that number and (un)knownness interact in the way expected for *they* (see Section 1.1). The only exception here, for which we have no ready explanation, is that *they* is substantially more acceptable in EXI-INF-NN (3.53) than in EXI-INF-PL (2.83) ($p<0.006$).[12]

Number and (un)knownness, as well as their interaction, affect Romanian 'they' too, as a two-way ANOVA makes clear ($p<0.05$ for all), and our post-hoc t-tests show that they do so entirely in the expected way. Pro-dropped *ei* has a significantly higher score: (i) in EXI-VAG-PL (4.63) than in EXI-INF-PL (3.73); (ii) in EXI-VAG-NN (4.45) than in EXI-INF-NN (3.78) and EXI-SPE-NN (2.87); and (iii) in EXI-SPE-PL (4.30) than in EXI-SPE-NN (2.87) ($p<0.006$ in all cases).[13] A comparison of these purely impersonal EXI uses to the three remaining non-UNI-INT ones produces some remarkable results, though. EXI-COR (4.37) is one of four contexts which exhibit no substantial variation in the third person plural's acceptability ($p>0.006$ in all cases) but which all score higher than any other context ($p<0.006$ in all cases). The uses with which EXI-COR shares the top spot are EXI-VAG-PL (4.63), EXI-VAG-NN (4.45) and EXI-SPE-PL (4.30) – and not UNI-EXT (3.53) and SAV (1.98), like in many other languages (e.g. Siewierska and Papastathi 2011: 596; Van Olmen and Breed 2018b: 838; see also Section 1.2). The fact that pro-dropped *ei* is less acceptable, though not unacceptable, in UNI-EXT means that there is no general preference for semi-impersonal uses of 'they' in Romanian, unlike in English. It also makes one wonder whether any other form of impersonalization is favored in this context. The same question arises for SAV. The fact that the third person plural's acceptability in this use is very low (and significantly lower than in any other use; $p<0.006$ for every t-test) suggests that Romanian is like French in having a range of impersonal contexts for 'they' but not employing it to mark information as coming from an unidentified source (see Section 1.1).[14]

A number of (dis)similarities between pro-dropped *ei* and *they* are evident from the preceding paragraphs. They obviously differ in their acceptability in UNI-EXT and SAV relative to the other uses but EXI-VAG-NN too occupies another position in Romanian than in English. It yields one of the highest scores for pro-dropped *ei*, yet one of the lowest ones for *they*. What the two forms share is relatively high scores in EXI-COR, EXI-VAG-PL and EXI-SPE-PL and considerably lower ones in EXI-INF-PL, EXI-INF-NN and EXI-SPE-NN. However, to answer the question raised in Section 1.3 – does a pro-dropped third person plural fare better than an overt one in less definite and less plural contexts? – a more direct comparison is required. The results of such an undertaking seem to confirm Siewierska and Papastathi (2011: 600) hypothesis. *They* only scores significantly higher than pro-dropped *ei* in the two semi-impersonal uses,

[11] EXI-VAG-PL (4.25) versus EXI-SPE-PL (4.02) on the one hand and EXI-VAG-NN (3.18) versus EXI-INF-NN (3.53) on the other hand do not differ in English.

[12] We considered the possibility of participants just finding the EXI-INF-PL scenario of having a party in the park somewhat strange, leading them to give lower scores. Their much higher scores for the passive in this context are, however, at odds with this explanation. We also gave thought to the possibility of participants regarding the EXI-INF-NN scenario of eating pizza in a common kitchen as an activity almost necessarily involving more than one person, prompting them to rate *they* higher than perhaps expected. A party is still a more clearly PL activity in our view, though.

[13] EXI-VAG-PL (4.63) versus EXI-VAG-NN (4.45) on the one hand and versus EXI-SPE-PL (4.30) on the other and EXI-INF-PL (3.73) versus EXI-INF-NN (3.78) do not differ in Romanian.

[14] Manea (2012) was quoted in Section 1.3 as saying that the SAV use of 'they' was attested in "older Romanian". We are not in any position to evaluate this claim. Still, if it is true, the data from our 18- to 35-year-old participants indicates that it has disappeared from the language.

i.e. UNI-EXT and UNI-COR, and in the inherently plural SAV use ($p<0.05$ in every case). Pro-dropped *ei*, by contrast, has a substantially higher acceptability than *they* in the "less plural" contexts of EXI-VAG-NN and EXI-SPE-NN and the "less definite" plural contexts of EXI-VAG-PL and EXI-INF-PL ($p<0.05$ in all cases). No differences exist in EXI-SPE-PL and EXI-INF-NN. Overall, the more reduced form of 'they' can therefore be said to be more bleached semantically, in faring better in uses that are further removed from its personal pronoun features.

3.2.3 The passive, versus 'they'

Let us now turn to the passive. A one-way ANOVA for English indicates that its acceptability varies substantially in the non-UNI-INT domain ($p<0.05$). Post-hoc t-tests point out that two uses in particular stand out, in a negative sense. The passive scores significantly lower ($p<0.006$ in all cases) in UNI-EXT (3.57) than in all other contexts, except for EXI-SPE-NN (4.1) and SAV (3.97), and in SAV than in the other contexts, excluding the two aforementioned ones and EXI-INF-NN (4.12). A plausible explanation for UNI-EXT's status is that the passive is essentially vague with respect to perspective, which has already been argued to be somewhat of a problem with this use (see the discussion of the scores of 'one' and 'you' in Table 2). The original subject might be taken to be *you*/*one*, in which case the passive receives an INT interpretation, or *they*, in which case it gets an EXT one. Speakers may prefer to avoid this ambiguity by employing an alternative impersonal construction. SAV's status is interesting given Siewierska's (2008: 12) assertion (see Section 1.2) that the passive is favored in this context in British English. A comparison with *they* is needed, though, to assess this claim properly. Another observation for which such an undertaking seems relevant is that the passive exhibits no clear preference for more definite and/or more plural uses. EXI-VAG-NN (4.78) is its top use, for example, and semi-impersonal EXI-COR (4.55) only comes fourth ($p>0.006$, however). More generally, despite other minor statistically significant differences, the passive can be said to score consistently, and perhaps surprisingly, high in all contexts other than UNI-EXT and SAV, ranging from 4.78 to 4.10 – while *they*'s scores vary from 4.77 to 2.37 in these uses.

A first finding of the two strategies' direct comparison is that UNI-EXT and SAV are two contexts where *they* scores substantially higher ($p<0.05$). In the former (4.82 versus 3.57), the pronoun has the advantage of making the EXT perspective explicit. Its higher acceptability in the latter (4.55 versus 3.97) contradicts Siewierska (2008: 12): it is actually the third person plural that is (slightly) favored in SAV. A possible explanation for this difference is the scenario's informality, which may be more compatible with the pronoun than with the passive (see Section 1.2). A second outcome concerns EXI-COR, where the statistically significant difference between *they* (4.77) and the passive (4.55) is fairly minimal. It suits the pronoun because of its semi-impersonal and plural character but it also fits well with the passive. The referents can be identified through the event itself here, making the inclusion of a separate marker pointing at them perhaps less essential. A third result is that, in all remaining contexts, the passive is significantly more acceptable than *they* ($p<0.05$ in every case), despite the former's assumed formal character (see Section 1.2). Thus, as suggested by Van Olmen and Breed (2018b: 839), pronouns may not be a preferential construction in the purely impersonal EXI domain in English, unlike in the other areas of impersonalization (see also Section 3.1.3). A fourth and final finding is that the differences between the passive and *they* are much larger in NN than in PL contexts in EXI-VAG (NN 4.78 and 3.18; PL 4.67 and 4.25) and EXI-SPE (NN 4.10 and 2.37; PL 4.57 and 4.02). This fact could be regarded as supporting Siewierska's (2008: 16) idea (see Section 1.2) that passives are favored over 'they' when the unidentified agent is almost certainly a single person.

Like its English equivalent, the Romanian passive shows substantial variation in the non-

UNI-INT domain, as a one-way ANOVA shows ($p<0.05$), and post-hoc t-tests reveal the "negative" prominence of the same two contexts. The passive scores significantly lower in UNI-EXT (2.40) and SAV (2.22) than in all other uses ($p<0.006$ in all cases), except for EXI-INF-NN (2.37). The passive's low acceptability in the former can probably again be explained by its perspectival ambiguity. Its score for the latter, together with our knowledge of *ei* in this context, suggests that a construction other than the passive or third person plural – the reflexive impersonal (see Section 1.2) in our intuition – is preferred for SAV in Romanian. As regards the other uses, the passive is most acceptable not just in semi-impersonal EXI-COR (4.30) but also in purely impersonal EXI-VAG-PL (4.67) and EXI-VAG-NN (4.45) (no substantial differences exist between them). Its score in EXI-SPE-NN, though significantly lower than in the preceding contexts ($p<0.006$ in all cases), is quite high too (3.78). However, it trails to such low levels in EXI-INF-PL (3.27), EXI-SPE-PL (3.15) and EXI-INF-NN (2.37) that one cannot but wonder whether pro-dropped – and semantically bleached – *ei* is preferred in these uses.

The explicit comparison needed to answer the question produces a fairly complex picture. As expected, the third person plural is significantly more acceptable in UNI-EXT than the passive (3.53 and 2.40; $p<0.05$) because of its unambiguously EXT reading and, in SAV, the two strategies seem equally unacceptable (1.98 and 2.22; $p>0.05$). For EXI-COR, pro-dropped *ei* and the passive score similarly high (4.37 and 4.30; $p>0.05$), which may be attributed to their comparable suitability for the use (see the argument for English). Moreover, they do not differ in EXI-VAG-PL (4.63 and 4.67) and EXI-VAG-NN (4.45 and 4.55) either. In three of the four remaining contexts, however, the third person plural has a substantially higher score than the passive – in EXI-INF-PL (3.73 and 3.27), EXI-INF-NN (3.78 and 2.37) and EXI-SPE-PL (4.30 and 3.15), to be precise ($p<0.05$ in each case). In other words, unlike in English, the pronoun actually seems to be favored in these purely impersonal EXI uses – a preference that is likely facilitated by the less definite and less plural character of impersonal pro-dropped *ei*. Still, one would probably assume the third person plural to be more acceptable in EXI-SPE-NN too then but, for some reason, the passive scores significantly higher here (3.78 and 2.87; $p<0.05$). Note, finally, that the above discussion features no evidence for a passive bias in number-neutral or single contexts, unlike in English, and that, arguably, no treatment of the competition between impersonal constructions in Romanian is complete without the reflexive impersonal (see Pană Dindelegan 2008: 138, who states that it is generally preferred to the passive).

To conclude the present section, we will briefly compare the passive in the two languages. It will not come as a surprise that it scores significantly higher in English than in Romanian in nearly every single use ($p<0.05$ in all cases). Contrasting the English passive to *they* has already shown that it is a more generally accepted construction in this area, particularly in the purely impersonal EXI domain and its NN contexts. The Romanian passive, by contrast, tends to fare worse than pro-dropped *ei* in such uses. One possible explanation for the differences in the passive between the two languages is therefore that its acceptability (at least partially) depends on the competition that it faces as an impersonal construction and thus on the degree to which the third person plural is bleached as an impersonal pronoun in the language. It is important to stress again in this regard that 'they' is not the only established alternative to the passive in Romanian, unlike in English. The existence of the reflexive impersonal may well play a role in the Romanian passive's overall lower scores too, averaging 3.41 (compared to 4.30 for its English equivalent). It almost certainly accounts, for instance, for the difference in SAV (2.22 versus 3.97), where the reflexive impersonal is probably the only option in our view. An assessment of this construction's impact in general will, however, have to be left for future research.

3.3 Summary

As a way to sum up our results, Figures 7 to 12 present the varying acceptability of the impersonal pronouns examined here and of the passive on Van Olmen and Breed's (2018b) semantic map (see Figure 4) in such a way that the darker a context is, the more acceptable the construction is there (*one* is not included here as, for our purposes, it can just be assumed to be a less acceptable version of *you*).

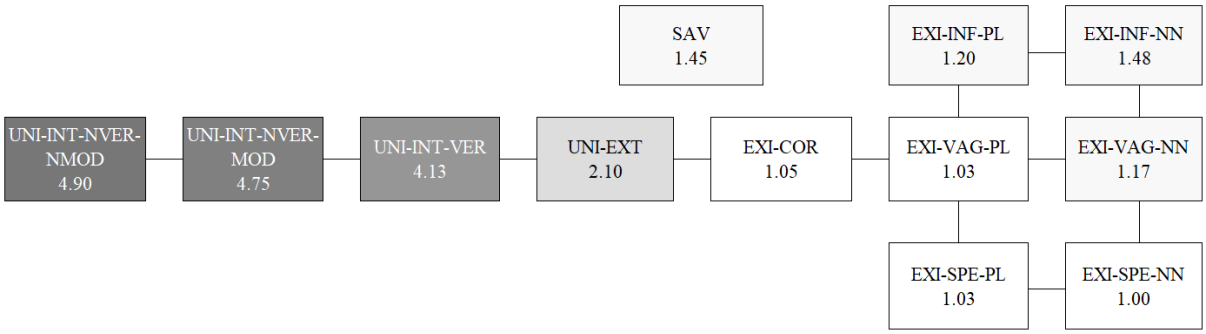

Figure 7: Romanian 'you'

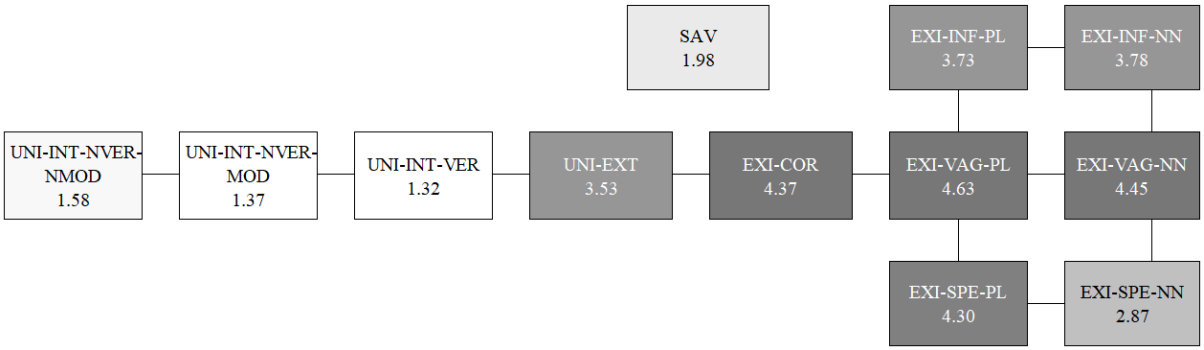

Figure 8: Romanian 'they'

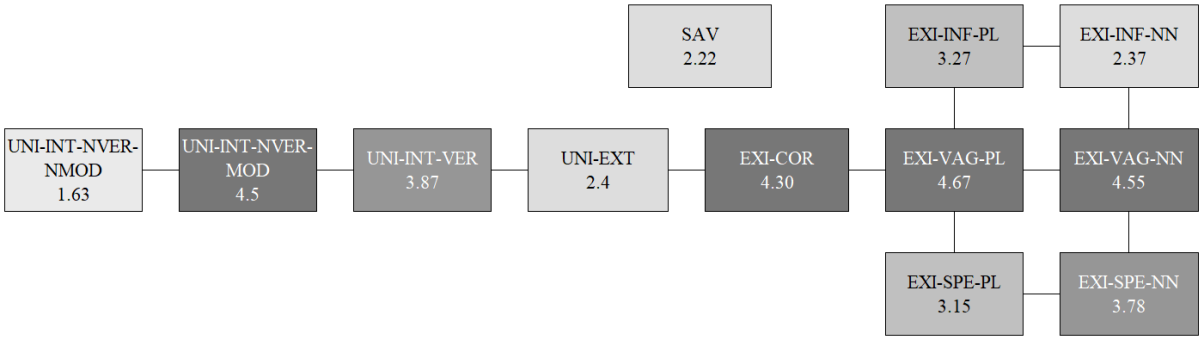

Figure 9: The Romanian passive

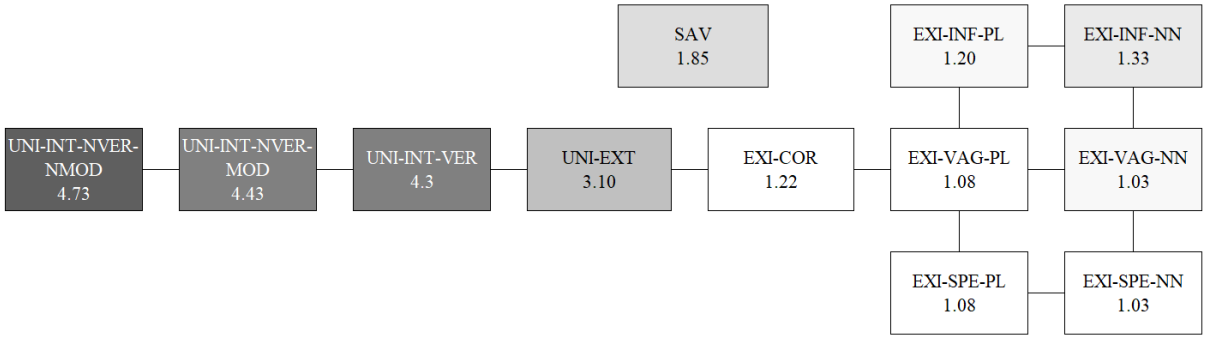

Figure 10: English 'you'

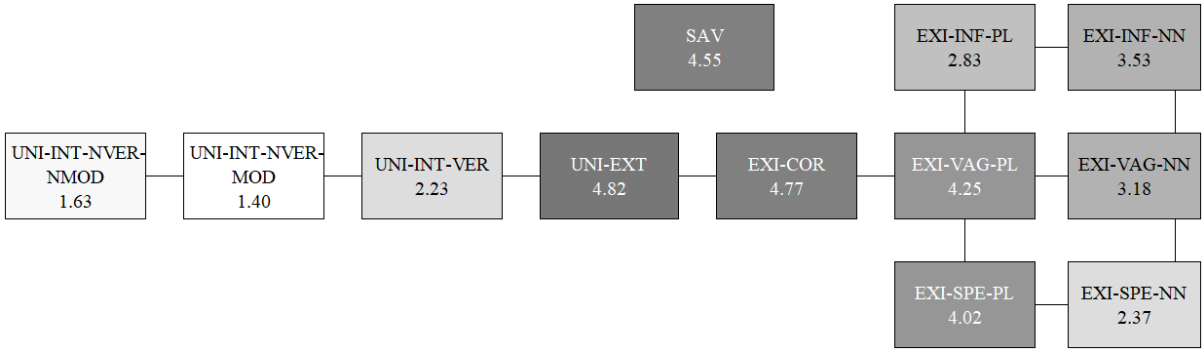

Figure 11: English 'they'

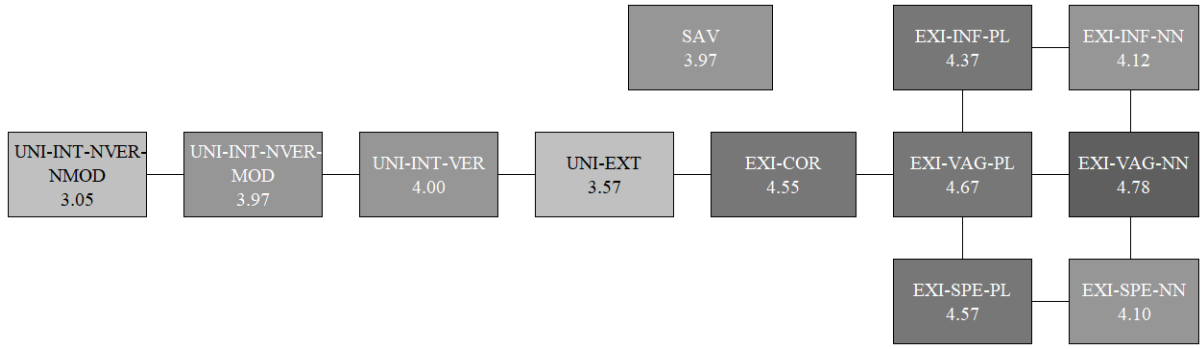

Figure 12: The English passive

4 Conclusion

As mentioned in Section 1.3, the main questions that this article aims to answer are: (i) what is the functional potential of the impersonal pronouns and the passive in Romanian and English?; (ii) how do the impersonal constructions and the languages compare to each other?; (iii) do the parameters of number and (un)knownness interact with one another (in the expected way) in Romanian?; and (iv) does the pro-dropped third person plural exhibit a(n expected) wider range of less definite/plural impersonal uses than its overt equivalent? We will now discuss these four questions one by one.

Let us start with (iv). Our results clearly show that the answer to this question is positive. Unlike its pro-dropped counterpart, overt *ei* is, in essence, unacceptable as an impersonal pronoun in Romanian. Pro-dropped *ei* is also more acceptable than *they* in contexts that are further removed from the third person plural's definite and plural personal pronoun prototype (compare Figures 8 and 11). Romanian can thus be added to the list of languages supporting Siewierska and Papastathi's (2011) hypothesis – based on the relation between formal reduction and semantic bleaching – that the more reduced 'they' is, the more likely it is to be able to occur in less definite or less plural impersonal contexts. This fact makes Polish, the only pro-drop language in their sample of nine with a limited range of impersonal uses for the third person plural, look even more exceptional. An investigation into the reason why it occupies this special position thus seems warranted. Still, the hypothesis itself would probably also benefit from more data from pro-drop versus non-pro-drop languages beyond Europe.

As to question (iii), our findings indicate that the answer is yes too. (Un)knownness and number combine to influence the acceptability of both English and Romanian 'they' in the truly impersonal EXI domain in the predicted way on the whole (see Figures 8 and 11). Pro-dropped *ei* in EXI-VAG/SPE-PL/NN can serve as an example of an interaction between the parameters that meets expectations. The VAG-SPE contrast itself is not sufficient to affect the pronoun's

acceptability, as there is no difference between EXI-VAG-PL and EXI-SPE-PL, and neither is the PL-NN contrast, as no difference exists between EXI-VAG-PL and EXI-VAG-NN. However, the combination of SPE and NN does produce a significantly lower acceptability of pro-dropped *ei*, compared to: (i) EXI-SPE-PL – in line with the prediction that 'they' fares worse in NN contexts due to possible persistence of its "personal" sense of plurality; and (ii) EXI-VAG-NN (and EXI-VAG-PL) – in keeping with the prediction that it fares worse in contexts relying on situational identification due to potential persistence of its "personal" sense of definiteness. Our Romanian and English data therefore reinforce Van Olmen and Breed's (2018b) argument – based on evidence from Afrikaans and Dutch as well as English – that, to capture the third person plural's behavior in the truly impersonal EXI domain, the separate parameters of (un)knownness and number in, respectively, Siewierska and Papastathi's (2011) and Gast and van der Auwera's (2013) maps need to be brought together. In future research, this requirement for a combined semantic map should be tested for other languages of the Romance family and others.

To answer question (ii), we first have to address (i). For Romanian, our findings reveal a clear division of labor between the impersonal pronouns 'you' and 'they'.[15] The former's functional potential lies in the UNI-INT domain and the latter's in the non-UNI-INT domain (compare Figures 7 and 8). Their respective incompatibility with the other sphere is, of course, not surprising (e.g. Kitagawa and Lehrer 1990) and can be traced back to their features as personal pronouns. The second person singular's addressee-oriented nature clashes with uses excluding the addressee as a possible referent; the third person plural's focus on parties other than speaker and addressee is at odds with contexts that include the speech participants as potential referents. Within the UNI-INT domain, there is, all in all, little variation. The fact that the second person singular's highest level of acceptability is found in UNI-INT-NVER-NMOD can nevertheless be said to be in line with Laberge and Sankoff's (1979: 432) findings for French, where *tu* is preferred (to *on*) in conditional contexts. Within the non-UNI-INT domain, Romanian 'they' seems unable to occur in SAV, which confirms Siewierska and Papastathi's (2011) treatment of this context as separate from the other ones. In these remaining uses (see Figure 8), the pronoun may exhibit the effects of (un)knownness and number – as mentioned earlier and especially in lower levels of acceptability in the EXI-INF and EXI-SPE-NN contexts – but scores rather consistently high, with no bias toward semi-impersonal UNI-EXT and EXI-COR (found in other languages; see below). This fact can be attributed to its pro-dropped character, as discussed before. The passive's functional potential, lastly, is not limited to the UNI-INT or non-UNI-INT sphere in Romanian (see Figure 9). It does exhibit substantial variation, though. The passive scores particularly low in UNI-EXT and SAV (and in UNI-INT-NVER-NMOD; see below). The likely reason for the former fact is that passives are essentially vague with respect to INT versus EXT point of view and speakers may prefer the perspectival clarity of the third person plural. The latter fact, together with the equally low acceptability of 'they' in this use, suggests that Romanian favors yet another construction for SAV, the reflexive impersonal. This construction deserves to be studied in more detail in future research. Its functional range and competition with the other impersonal constructions in the language could help explain some of the passive's other relatively lower acceptability levels in the non-UNI-INT domain.

Let us now answer question (i) for English and, at the same time, start answering (ii) too. Our findings confirm the division of labor between UNI-INT *you* and *one* and non-UNI-INT *they* (see Figures 10 and 11; cf. Van Olmen and Breed 2018b: 842). In this regard, English and Romanian are thus similar. Within the UNI-INT domain, *you* is consistently more acceptable

[15] The lack of an impersonal pronoun covering the complete range of uses makes Romanian similar to its Slavic neighbors and Romance languages like Italian and Spanish but different from, for instance, French with *on.* The absence of such a pronoun in English renders the language quite unique, though, among its closest neighbors and relatives (e.g. German with *man*) (see Siewierska 2011: 71).

than *one*. This result is in line with earlier findings (see Van Olmen and Breed 2018b: 815) and can be attributed to a clash between *one*'s formal character and the informality of our scenarios. Contrary to certain claims in the literature (e.g. Moltmann 2010), *one*'s acceptability does not seem to be affected by (non-)veridicality. This fact is, however, in keeping with Haas's (2018b) assertion that its impact on *one* has faded over time. *You* does not differ much from its Romanian equivalent in acceptability terms (compare Figures 7 and 10). Like *tu*, it also scores highest in UNI-INT-NVER-NMOD and can therefore be argued to support the aforementioned link – which merits further scrutiny – between the impersonal second person singular and conditional contexts. Within the non-UNI-INT domain, *they* displays substantial variation. Like its Romanian counterpart and as discussed before, the pronoun shows the effects of the interacting parameters of (un)knownness and number in, inter alia, its low level of acceptability in EXI-SPE-NN. Yet, it is typically less acceptable in these truly impersonal EXI uses than pro-dropped *ei*, as a result of its overt status (see above). *They* is also generally less acceptable in these contexts than in UNI-EXT and EXI-COR, where it actually fares better than pro-dropped *ei* (compare Figures 8 and 11). So unlike Romanian 'they', it does have the "overall speaker preference for semi-impersonal uses" that Siewierska and Papastathi (2011: 599) observe for the third person plural in many languages. *They* also differs from pro-dropped *ei* in being very much acceptable in SAV. The English passive, lastly, clearly has the functional potential to appear in UNI-INT contexts but, on the whole, its scores for the non-UNI-INT domain are higher (see Figure 12). The only real exception there is UNI-EXT, a feature that it shares with its Romanian equivalent and can be attributed to its perspectival ambiguity. Unlike the Romanian passive, however, the English one is an acceptable construction in SAV and scores fairly consistently high in the truly impersonal EXI sphere (compare Figures 9 and 12). Part of the explanation for the latter fact may lie in the stronger competition that the Romanian passive faces from the more bleached third person plural in the language and probably also from the reflexive impersonal.

The previous paragraph has compared each impersonal construction in English to its Romanian counterpart. To answer question (ii) fully, the present paragraph will discuss the intralinguistic relationships between the constructions. Within the UNI-INT sphere, the passive has been said to score fairly high in both Romanian and English, with the exception of UNI-INT-NVER-NMOD – though, generally, still lower than 'you' in the two languages, except in UNI-VER (compare Figures 7 and 9 and 10 and 12). These findings support the idea of a UNI-INT preference for pronouns (e.g. Sansò 2006; Van Olmen and Breed 2018b) but show, at the same time, the passive's evident UNI-INT potential. They also suggest, consistent with Gast's (2015) exploratory findings, that the choice between 'you' and the passive is at least partially affected by veridicality and/or modality. In the non-UNI-INT domain, UNI-EXT stands out as the use with one of the lowest levels of acceptability for the passive in Romanian and English, as mentioned earlier. Our argument that its perspectival vagueness is the reason is reinforced by the much higher acceptability of 'they' in the use (compare Figures 8 and 9 and 11 and 12). SAV is of interest too in the two languages. The passive's similarly low score to the third person plural in Romanian indicates that the language favors yet another construction for the context. The fact that, in English, the passive is less acceptable than 'they' here challenges Siewierska's (2008: 12) claim that the language prefers the passive for SAV. The informality of our scenario has perhaps played a role in the results, though. More generally, the English passive is found to score consistently high, not only in the semi-impersonal contexts but also in the truly impersonal EXI ones. In fact, in the latter and their NN versions in particular, it outperforms *they*. This finding lends support to Van Olmen and Breed's (2018b) suggestion that pronouns may not be the preferred construction in the non-UNI-INT domain and to Siewierska's (2008: 16) assertion that passives are favored in NN (or single) contexts, at least in English. The Romanian passive's acceptability varies much more in the non-UNI-INT domain. As pointed out before, part of the explanation could be that it just has more competition in these contexts from the

more bleached third person plural (and the reflexive impersonal) than its English equivalent.

This article is by no means the final word on impersonal constructions in Romanian versus English. The constructions analyzed here should, for example, be compared to the reflexive impersonal in Romanian. It would be a good idea to examine their actual usage in corpora too. Such a study would also make it possible, for instance, to investigate the ways that the second person singular is employed in contexts beyond "structural knowledge" (Kitagawa and Lehrer 1990: 748; see Section 1.2).

Abbreviations

2	second person
3	third person
COMP	complementizer
COND	conditional
DEF	definite
DEM	demonstrative
DIST	distal
F	feminine
INDF	indefinite
INF	infinitive
M	masculine
N	neuter
NEG	negation
PL	plural
PRS	present
PST	past
PTCP	participle
REFL	reflexive
SG	singular

References

Baayen, R.H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Bard, E.G., D. Robertson and A. Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72(10, 32-68.

Cabredo Hofherr, P. 2006. "Arbitrary" pro and the theory of pro-drop. In P. Ackema, P. Brandt, M. Schoorlemmer and F. Weerman (eds.), *Agreement and Arguments*. Oxford: Oxford University Press, 230-257.

Cojocaru, D. 2003. *Romanian Grammar*. Durham, NC: Slavic and East European Language Research Center, Duke University.

Coussé, E. and J. van der Auwera. 2012. Human impersonal pronouns in Swedish and Dutch: A contrastive study of *man* and *men*. *Languages in Contrast* 12(2), 121-138.

Creissels, D. 2008a. Impersonal and related constructions: A typological approach. Manuscript. Lyon: Lumière University Lyon 2.

Creissels, D. 2008b. Impersonal pronouns and coreference: The case of French *on*. Manuscript. Lyon: Lumière University Lyon 2.

D'Alessandro, R. 2007. *Impersonal* Si *Constructions: Agreement and Interpretation*. Berlin: De Gruyter.

Data-Bukowska, E. 2018. The third person plural impersonal in Swedish: A typological

account. *Linguistica Copernicana* 15, 163-193.
De Hoop, H. and S. Tarenskeen. 2015. It's all about *you* in Dutch. *Journal of Pragmatics* 88, 163-175.
Dobrovie-Sorin, C. 1998. Impersonal *se* constructions in Romance and the passivization of unergatives. *Linguistic Inquiry* 29(3), 399-437.
Fenger, P. 2018. How impersonal does *one* get? A study of *man*-pronouns in Germanic. *Journal of Comparative Germanic Linguistics* 21(3), 291-325.
Fernández, S.S. 2013. Impersonality in Spanish personal pronouns. In K.J. Kragh and J. Lindschouw (eds.), *Deixis and Pronouns in Romance Languages*. Amsterdam: Benjamins, 87-107.
Fonesca-Greber, B. and L.R. Waugh. 2003. On the radical difference between the subject personal pronouns in written and spoken European French. In P. Leistyna and C.F. Meyer (eds.), *Corpus Analysis: Language Structure and Language Use*. Amsterdam: Rodopi, 225-240.
Gast, V. 2015. On the use of translation corpora in contrastive linguistics: A case study of impersonalization in English and German. *Languages in Contrast* 15(1), 4-33.
Gast, V., L. Deringer, F. Haas and O. Rudolf. 2015. Impersonal uses of the second person singular: A pragmatic analysis of generalization and empathy effects. *Journal of Pragmatics* 88, 148-162.
Gast, V. and J. van der Auwera. 2013. Towards a distributional typology of human impersonal pronouns, based on data from European languages. In D. Bakker and M. Haspelmath (eds.), *Languages Across Boundaries: Studies in the Memory of Anna Siewierska*. Berlin: De Gruyter, 119-158.
Giacalone Ramat, A. and A. Sansò. 2007. The spread and decline of indefinite man-constructions in European languages: An areal perspective. In P. Ramat and E. Roma (eds.), *Europe and the Mediterranean Linguistic Areas: Convergences from a Historical and Typological Perspective*. Amsterdam: Benjamins, 95-131.
Gönczöl-Davies, R. 2008. *Romanian: An Essential Grammar*. London: Routledge.
Haas, F. 2018a. "You can't control a thing like that": Genre and change in modern English human impersonal pronouns. In R.J. Whitt (ed.), *Diachronic Corpora, Genre and Language Change*. Amsterdam: Benjamins, 171-194.
Haas, F. 2018b. Changing human impersonal pronouns in English: A corpus study. Paper at Societas Linguistica Europaea 51, Tallinn.
Hall, D. 2020. The impersonal gets personal: A new pronoun in Multicultural London English. *Natural Language & Linguistic Theory* 38, 117-150.
Jensen, T.J. and F. Gregerson. 2016. What do(es) you mean? The pragmatics of generic second person pronouns in modern spoken Danish. *Pragmatics* 26(3), 417-446.
Kitagawa, C. and A. Lehrer. 1990. Impersonal uses of personal pronouns. *Journal of Pragmatics* 14(5), 739-759.
Kluge. B. 2016. Generic uses of the second person singular – How speakers deal with referential ambiguity and misunderstanding. *Pragmatics* 26(3), 501-522.
Laberge, S. and G. Sankoff. 1979. Anything *you* can do. In T. Givón (ed.), *Discourse and Syntax*. New York: Academic Press, 419-440.
Malamud, S. 2012. Impersonal indexicals: *One*, *you*, *man* and *du*. *Journal of Comparative Germanic Linguistics* 15, 1-48.
Manea, C. 2012. Remarks on the passive voice in English and Romanian. *Studii de Gramatică Contrastivă* 18, 54-73.
Moltmann, F. 2010. Generalizing detached self-reference and the semantics of generic *one*. *Mind & Language* 25(4), 440-473.
Myhill, J. 1997. Toward a functional typology of agent defocusing. *Linguistics* 35(5), 799-

844.
O'Connor, P.E. 1994. "You could feel it through the skin": Agency and positioning in prisoners' stabbing stories. *Text* 14(1), 45-75.
Pană Dindelegan, G. 2008. Construcții pasive și construcții impersonale. In V. Guțu Romalo (ed.), *Gramatica Limbii Române II: Enunțul*. Bucharest: Editura Academiei Române, 133-147.
Pană Dindelegan, G. (ed.). 2013. *The Grammar of Romanian*. Oxford: Oxford University Press.
Posio, P. and M. Vilkuna. 2013. Referential dimensions of human impersonals in dialectal European Portuguese and Finnish. *Linguistics* 51(1), 177-229.
Rasinger, S.M. 2013. *Quantitative Research in Linguistics: An Introduction*. London: Bloomsbury.
Rudolf, O. 2016. *Human Impersonal Strategies in English and Russian: A Comparative Corpus Study*. Jena: University of Jena PhD dissertation.
Sansò, A. 2006. 'Agent defocusing' revisited: Passive and impersonal constructions in some European languages. In W. Abraham and L. Leisiö (eds.), *Passivization and Typology: Form and Function*. Amsterdam: Benjamins, 232-273.
Shibatani, M. 1985. Passives and related constructions: A prototype analysis. *Language* 61(4), 821-848.
Siewierska, A. 2004. *Person*. Cambridge: Cambridge University Press.
Siewierska, A. 2008. Ways of impersonalizing: Pronominal vs. verbal strategies. In M.A. Gómez González, J. Lachlan Mackenzie and E.M. González Álvarez (eds.), *Current Trends in Contrastive Linguistics: Functional and Cognitive Perspectives*. Amsterdam: Benjamins, 3-26
Siewierska, A. 2011. Overlap and complementarity in reference impersonals: *Man*-constructions vs. third person plural-impersonal in the languages of Europe. In A. Malchukov and A. Siewierska (eds.), *Impersonal Constructions: A Cross-Linguistic Perspective.* Amsterdam: Benjamins, 57-90.
Siewierska, A. and M. Papastathi. 2011. Towards a typology of third personal plural impersonals. *Linguistics* 49(3), 575-610.
Sprouse, J. and D. Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's *Core Syntax*. *Journal of Linguistics* 48(3), 609-652.
Stirling, L. and L. Manderson. 2011. About *you*: Empathy, objectivity and authority. *Journal of Pragmatics* 43(6), 1581-1602.
van der Auwera, J., V. Gast and J. Vanderbiesen. 2012. Human impersonal pronoun uses in English, Dutch and German. *Leuvense Bijdragen* 98, 27-64.
Van Olmen, D. and A. Breed. 2018a. Human impersonal pronouns in Afrikaans: A double questionnaire-based study. *Language Sciences* 69, 1-29.
Van Olmen, D. and A. Breed. 2018b. Human impersonal pronouns in West Germanic: A questionnaire-based comparative study of Afrikaans, Dutch and English. *Studies in Language* 42(4), 798-846.
Zifonun, G. 2001. *Man lebt nur einmal*: Morphosyntax und Semantik des Pronomens *man*. *Deutsche Sprache* 38(3), 232-253.