‘Human impersonal pronouns in West Germanic
A questionnaire-based comparative study of Afrikaans, Dutch and English[1]

Daniël Van Olmen (Lancaster University)
Adri Breed (North-West University Potchefstroom)

Abstract

In this article, we examine and compare the main human impersonal pronouns in Afrikaans, Dutch and English. The second person singular, the third person plural and the ‘man’- and ‘one’-pronouns are studied by means of an acceptability judgment questionnaire and a completion questionnaire. The combination of the two methods reveals interesting descriptive facts about the three West Germanic languages. They include, among other things, the ‘man’-prominence of Afrikaans versus the ‘you’-prominence of Dutch and English for expressing the universal meaning ‘anyone’ and the more prominent position of ‘they’ in Dutch than in the other languages for conveying the existential meaning ‘someone, some people’. Our findings have a number of more theoretical implications too. The two existing semantic maps for human impersonal pronouns make different distinctions in the existential domain, based on type/level of (un)knownness on the one hand and number on the other. Our study tests both sets of distinctions and shows that the two dimensions interact with each other in Afrikaans, Dutch and English. The results thus support a recent proposal in the literature for a combined semantic map. The data from the completion questionnaire, finally, also indicates that existential uses prefer alternative forms of impersonalization to human impersonal pronouns in all three languages.

Keywords

acceptability, Afrikaans, completion, Dutch, English, impersonal pronoun, questionnaire, semantic map

## 1 Introduction

The last decade has seen an increasing interest in human impersonal pronouns (henceforth HIPs), i.e. the pronominal expression 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 & van der Auwera 2013: 124). The research has paid particular attention to West Germanic (e.g. Hoekstra 2010, Coussé & van der Auwera 2012, Malamud 2013, Deringer et al. 2015, Zobel 2016, Van Olmen et al. 2018). Few studies have, however, examined the actual usage of the HIPs in these languages in detail – probably partially because of the problem of disentangling the impersonal uses of, for example, ‘you’ and ‘they’ from their personal ones in large amounts of data (cf. Siewierska & Papastathi 2011: 587-588 on this issue). Van der Auwera, Gast & Vanderbiesen (2012), for instance, give us an idea of the frequencies of the HIPs in Dutch, English and German but their results are based on a one-text parallel corpus (see Gast 2015 for

[1] An earlier version of this paper, which builds on Van Olmen & Breed (2018), was presented at the 6th Germanic Sandwich Conference in Münster in March 2017. Thanks are due to Frank Brisard, Timothy Colleman, Thys Human, Janien Linde, Reinhild Vandekerckhove, Ben Verhoeven and two anonymous reviewers.

another parallel corpus study of English and German impersonalization). Moreover, they do not analyze the usage contexts of the HIPs in terms of the subtle distinctions typically made in the literature – likely partly on account of the inevitable difficulties in applying such fine-grained classifications to attestations in authentic, intrinsically fuzzy discourse (cf. Siewierska & Papastathi 2011: 591-592 on this problem). In the present article, we take the existing semantic maps of the domain, namely Siewierska & Papastathi (2011) and Gast & van der Auwera (2013), as our starting point to compare the functional potential and usage of the main HIPs in English – i.e. *you*, *one* and *they* – and Dutch – i.e. *je* 'you', *men* 'man' and *ze* 'they' – to those in Afrikaans – i.e. *jy* 'you', (*'n*) *mens* '(a) human' and *hulle* 'they' (see Van Olmen & Breed 2018).[2] To avoid the aforementioned issues with corpus research, we follow Siewierska & Papastathi (2011: 592-599) and Gast (2017) and adopt a questionnaire-based approach.

With this study, we aim to address at least three issues. First, the HIPs in Dutch and English feature prominently in Gast & van der Auwera's (2013) semantic map. The judgments about their functional potential seem to be based primarily on the authors' intuitions, though. For truly impersonal existential contexts, i.e. implying the existence of a specific (set of) human being(s) with no additional cues to enable their partial identification (cf. Siewierska & Papastathi's 2011: 581 "semi-impersonal" uses), for instance, they distinguish cases with a necessarily plural interpretation, like (1a), from cases with a number-neutral reading, like (1b).[3] Their claim about Dutch and English is that, in the latter language, the third person plural can occur in both contexts while it normally only licenses a plural interpretation in the former (see Gast & van der Auwera 2013: 149). Yet, *ze* in (2) appears to allow a singular reading, though it is obviously impossible to know the speaker's true intent.

(1) a. D ***Ze*** *hebben ons omsingeld.*
'They have surrounded us.'
b. D ?****Ze*** *hebben je zak in het park gevonden.*
'They have found your bag in the park.'

(2) D *Broodjes waren lekker maar* ***ze*** *hebben mijn auto gestolen op de parking.*
'Sandwiches were good but they stole my car from the parking lot.'
(https://www.goudengids.be/bedrijf/Bonheiden/L6256483/Broodjeszaak+Mondje-Vol, accessed on 16 April 2018)

Our questionnaires, which rely on information from a considerable number of participants, can tell us to what extent certain unsubstantiated claims in the literature about the HIPs in West Germanic, like the one about 'they' in (1), are valid or not.

Second, aside from the descriptive aims mentioned in the previous paragraph, we seek to test a striking difference between Gast & van der Auwera's (2013) map and Siewierska & Papastathi's (2011). The distinction in the truly existential uses between plural and number-neutral contexts is not made in Siewierska & Papastathi (2011). Instead, in line with Cabredo Hofherr (2006) among others, they distinguish three contexts that vary in the kind or degree of (un)knownness. In (2a), a

[2] We are aware that other HIPs exist in these languages, such as the first person plural and – in Dutch and Afrikaans – the formal second person *u*. Our completion tasks confirm, however, that the informal second person singular, 'one', 'they' and the so-called 'man'-pronoun (see Giacalone Ramat & Sansò 2007) are the most common ones in all three. Note also that the translations of *men* and (*'n*) *mens* refer not to their present-day meanings but to their lexical sources.
[3] Afrikaans, Dutch and English are abbreviated to, respectively, A, D and E in the examples.

so-called "vague" context, the set of human participants is unidentifiable to the speaker but there is said to exist at least one particular individual that carried out the known act of grand theft. In the "inferred" context in (3a), the speaker concludes from the situation, i.e. the odor in the room, that the unknown action of eating popcorn must have been performed and that there must have been an unspecified (group of) eater(s). The "specific" context in (3b), finally, involves an event occurring at a precise time and place. The speaker may thus have an idea about who is making it happen but refrains from overtly identifying the person or people at the door.

(3) a. E **They** have eaten popcorn in here. (I can smell it.)
b. E **They** are knocking on the door. (It will be your brother.)

Siewierska & Papastathi's (2011: 596) questionnaire brings to light significant variation in the acceptability of 'they' along this dimension of (un)knownness in many of the languages under examination. The results lead them to postulate a semantic map configuration whereby the inferred and specific uses are separate off-shoots of the vague use, which is linked to the third person plural's other (semi-)impersonal uses. Still, maybe somewhat surprisingly, Gast & van der Auwera (2013: 143) write: "We have collapsed the distinction between 'vague', 'inferred' and 'specific' … because we lack the evidence for it in the languages investigated by us." On their semantic map, it is the number-neutral use that follows the plural one, which is connected to the semi-impersonal ones. As discussed earlier, this alternative dimension of number lacks substantial empirical support, however. It *has* been shown to be relevant for Afrikaans *hulle*, whose acceptability and usage decreases not only from plural to number-neutral contexts but also from vague to inferred and specific ones (Van Olmen & Breed 2018). The question that we want to answer here is whether this interaction of the two dimensions is actually also found in other third person plural HIPs – as well as, perhaps, in the potentially existential HIP *men* in Dutch.

Third, it is well-known (e.g. Posio & Vilkuna 2013, Gast 2015) that impersonalization can be conveyed by other means than HIPs. In the so-called speech act verb use in (4a), for instance, the third person plural can easily be replaced by the indefinite plural noun phrase *mense* 'people' in (4b) (cf. Siewierska & Papastathi's 2011: 604 'people' use). Another possible strategy in Afrikaans would be the passive in (4c).

(4) a. A ***Hulle*** *sê dat daardie huis wemel van die spoke.*
'They say that that house is crawling with ghosts.'
b. A ***Mense*** *sê dat daardie huis wemel van die spoke.*
'People say that that house is crawling with ghosts.'
c. A ***Daar word gesê*** *dat daardie huis wemel van die spoke.*
'It is said that that house is crawling with ghosts.'

What we now want to examine is whether there exist any uses for which such alternative forms of impersonalization are preferred to HIPs. This question is motivated by the fact that, in much corpus research, HIPs are found not to be employed (very often) for purposes which they are, in principle, able to serve. In Siewierska & Papastathi's (2011: 590) corpus study of 'they' in nine languages, for example, "there are no instances of either the inferred or the specific existential [use]". The so-called 'man'-pronoun is another case in point. In many languages, it can be used not only existentially – as in (5a), where it more or less has the meaning of 'someone' or 'some people' – but also universally – like in (5b), where it can be roughly paraphrased as 'anyone' (see Giacalone Ramat & Sansò 2007: 124; Siewierska 2011: 71). In this second context, the sentence does not imply the

existence of a specific (group of) individual(s) but says something about a quasi-universal set of people. Crucially, despite their functional potential, such 'man'-pronouns have been found to rarely serve existential purposes and mainly appear as a universal HIP in actual usage (e.g. Zifonun 2001 on German *man* 'man'; Fonesca-Greber & Waugh 2003 on French *on* 'human').

(5) a. D ***Men*** *heeft hier flitspalen geïnstalleerd.*
'They have installed speed cameras here.'
b. D *Als* ***men*** *naar Engeland gaat, is het essentieel om een regenjas mee te nemen.*
'If one goes to England, it is best to take a raincoat.'

This scarcity of (particular) existential uses of HIPs in corpus research may be due to the fact that this type of impersonalization simply does not come up very often in discourse and/or to some kind of aversion to pronominal strategies and preference for alternative strategies to express it. At least the second factor has been shown to have an impact in Afrikaans, by means of a completion task questionnaire (Van Olmen & Breed 2018). It remains to be seen, though, whether it plays a role in other languages – such as Dutch, which, unlike Afrikaans with just *hulle*, has not one but two HIPs that can, in principle, be used existentially, i.e. *ze* and *men*.

The rest of this article is structured as follows. Section 2 describes our methodology, including the design of our questionnaires and some statistical considerations. In Section 3, we present and discuss our results for the so-called universal-internal uses, to which the second person singular, Afrikaans (*'n*) *mens* and English *one* are restricted. Section 4 reports and examines the findings for the so-called universal-external use and the speech act verb and existential uses, to which the third person plural is limited (Dutch *men* is the only other HIP featuring in Sections 3 and 4). To easily refer to this varied set of uses, we employ the negative term "non-universal-internal" because their common feature is that they are not universal-internal. Section 5, lastly, is our conclusion.

## 2 Methodology[4]

The questionnaire-based approach adopted in this article is two-fold. In Section 2.1, we present the acceptability judgment task used to describe the functional potential of the main HIPs in Afrikaans, Dutch and English. Section 2.2 deals with the completion task used to examine their usage and test the (dis)preference for pronominal strategies to convey impersonalization. Following Greenbaum & Quirk (1970), one could say that this double approach captures, on the one hand, speakers' attitude toward the forms under investigation and, on the other, their use of them – although the completion task can obviously only be considered an approximation of speakers' actual usage preferences. In Section 2.3, we briefly compare our questionnaires to those of Siewierska & Papastathi (2011) and Gast (2017). Section 2.4 looks at the delivery of the questionnaires and at our participants. In Section 2.5, finally, we introduce the statistical tests used in the paper.

### 2.1 Acceptability judgment questionnaire

Acceptability judgments have typically been regarded as an indication of the grammaticality of the test items. Bard et al. (1996: 33) point out, however, that they "need not be one-to-one reflections of grammaticality" because they "may be based, for example, on estimated frequency of usage, on

[4] The methodology described in this section is similar to that of Van Olmen & Breed (2018: 6-10).

conformity to a prescriptive norm or a prestigious register, or on degree of semantic or pragmatic plausibility". It is this acceptability in very particular semantic/pragmatic contexts that is central to our study. As *you* and *they* in (6) make clear, most HIPs produce well-formed sentences when they are inserted in some clause and, on the face of it, such sentences are likely to be judged equally acceptable. Only *you* in (6a) has the potential to serve as a HIP, though, in the right discourse environment.

(6) a. E **You** should never give up.
b. E **They** should never give up.

For that reason, each part of our questionnaire first sketches a scenario that activates some specific impersonal interpretation of the test sentences and then asks the participant to judge their acceptability as a possible end of the scenario on a scale from one to five, i.e. from 'very unacceptable' to 'very acceptable'.[5] The questionnaire segment in (7) is a case in point.

(7) E Someone asks how your daughter is feeling after yet another unsuccessful job. You reply: "She is a little down but she still believes that …"

| | | | | | |
|---|---|---|---|---|---|
| … one should never give up." | 1 | 2 | 3 | 4 | 5 |
| … you should never give up." | 1 | 2 | 3 | 4 | 5 |
| … they should never give up." | 1 | 2 | 3 | 4 | 5 |

Like in (7), our scenarios tend to involve quite casual situations and fairly colloquial language and encourage the participant to the take the speaker's point of view. The rationale behind this set-up is that the use of HIPs is often more characteristic of informal speech (see Siewierksa & Papastathi 2011: 593). Note that the Dutch questionnaire, like the English one, offers three potential ways of completing each scenario – i.e. with *je*, *men* and *ze* – while the Afrikaans questionnaire presents four options – i.e. with *jy*, the 'man'-pronoun's less grammaticalized form *'n mens*, its more grammaticalized variant *mens*[6] and *hulle*.

The questionnaire tests twelve different uses, based on Siewierska & Papastathi's (2011) map for third person plural HIPs and Gast & van der Auwera's (2013) map for all HIPs. The first three are specific to the latter. They are all universal and take what Gast & van der Auwera (2013: 139) call an internal perspective: "A 'center of consciousness' (e.g. the speaker or hearer) identifies, or

[5] Our choice of a Likert scale, in spite of its shortcomings (e.g. the suggestion that there are just five levels of acceptability), over, for instance, the increasingly popular method of magnitude estimation (see Bard et al. 1996) is motivated by the following considerations. First, magnitude estimation has proved useful for syntactic research (e.g. Hofmeister et al. 2007) but is not very suitable for a questionnaire with lengthy scenarios necessary to trigger the correct impersonal reading. Second, Sprouse & Almedia (2012) have criticized its cognitive assumptions and shown that its outcomes do not actually differ substantially from those of more established methods. Third, our participants can be expected to be accustomed to making scale-based judgments (see Section 2.4) but would need to be trained in magnitude estimation before filling out a questionnaire using this method.

[6] For the relation between the indefinite article's disappearance and grammaticalization, see Giacalone Ramat & Sansò (2007: 102), Kirsten (2016: 192) and Van Olmen et al. (2018: §1). In a nutshell, *mens* with *'n* can function as a noun or a HIP. *'n Mens het nie vleuels nie* 'a human being does not have wings' and *'n mens mag nooit drink en bestuur nie* 'one should never drink and drive' can serve as respective examples. Without *'n*, by contrast, *mens* can basically only be a HIP, as in *mens leef net een keer* 'one only lives once'. Adding *'n* to this last example produces a sentence vague between a nominal reading and a pronominal one and thus illustrating the type of context in which the noun becomes the HIP: *'n mens leef net een keer* 'a human being/one only lives once'.

is identified, with the set of referents under discussion". The statement in (7), for instance, is about people in general and encourages the speech participants to regard it as applying to them too. The three universal-internal uses differ in the state of affairs of the clause (the cross-linguistic support for which can be found in Gast & van der Auwera 2013: 143-151). In (7), it is non-veridical and modal (abbreviated as UNI-INT-NEVER-MOD): it is not presumed to be true because of the presence of some modal operator of possibility or necessity, like *should* here. In (8a), the state of affairs is not assumed to be true either but the non-veridical operator is non-modal, i.e. the conditional subordinate clause in which the HIP is used. In (8b), by contrast, the state of affairs is veridical.

(8) a. universal-internal use in a non-veridical non-modal clause (UNI-INT-NVER-NMOD)
E 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 gave her the day before as a gift, your brother asks you why you bought it for her. You reply: "If **you/one/they** go/goes to England, it's best to take a raincoat."

b. universal-internal use in a veridical clause (UNI-INT-VER)
E Someone comments on your female friend's love for fine dining. You are irritated and reply: "She can afford it and, besides, **you/one/they** only live/lives once."

The next two uses occur in both semantic maps. The first one, in (9a), resembles (7) and (8) in that it concerns people in general. The difference is that the interlocutors are not invited to identify with the set of referents in this case: it is the inhabitants of Greece that drive unpredictably, not anyone (like, for example, the speaker or hearer) who happens to be in the country at some point. The use is therefore called "universal-external" (cf. Gast & van der Auwera 2013: 141). The second one, in (9b), is existential: it implies the existence of a specific group of people who put the speed cameras up. What (9a) and (9b) have in common is that they contain cues enabling partial identification of the referents. In (9a), it is the locative phrase *in Greece* that, in a way, delimits reference to Greek residents. In (9b), it is the predicate itself that gives an indication of the referents: there are only a few institutions, like the government and the police, that have the authority to install speed cameras. The use is often termed "(existential-)corporate" (cf. Gast & van der Auwera 2013: 128). Because of these cues, both (9a) and (9b) are considered semi-impersonal by Siewierska & Papastathi (2011: 581).

(9) a. universal-external use (UNI-EXT)
E A friend who is going to Greece for business asks you whether they should rent a car to get around. Based on your own negative experience in that country, you reply: "I don't think that is a good idea. In Greece, **you/one/they** drive/drives quite unpredictably."

b. corporate use (EXI-COR)
E A friend is driving you to work in the morning. You see signs warning people that there are speed cameras ahead. You tell your friend: "**You/one/they** have/has installed speed cameras here."

The sixth use, the speech act verb one in (10), is unique to Siewierska & Papastathi (2011). Gast & van der Auwera (2013: 142) do acknowledge its existence but do not see how or where it fits in with their map functionally. Siewierska & Papastathi (2011: 585) struggle to give cases like (10) a place as well: "They do not really fall under vague in Cabredo Hofherr's classification since they

are typically not episodic, i.e. do not refer to a specific event. Note also that they 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 or place." Still, since 'they' only has this impersonal use in some languages and, in others, a range of impersonal uses such as (9) but not this one, they regard it as a separate off-shoot from the third person plural's personal use.

(10) speech act verb use (SAV)
E Your colleague has just bought a new house. You have heard that it is haunted and tell them: "**You/one/they** say/says that that house is crawling with ghosts."

For the final six uses, the truly impersonal existential ones, we combine Siewierska & Papastathi's (2011) dimension of (un)knownness with Gast & van der Auwera's (2013) dimension of number (see Section 1). In both (11a) and (11b), for instance, the speaker gathers from the context, i.e. the beer bottles in the first scenario and the odor in the second one, that the unknown action of meeting for a party or consuming pizza must have been carried out by an unidentifiable (group of) individual(s). The two examples have an inferred interpretation, in other words. Their number reading is different, however. The act of gathering in (11a) entails a necessarily plural subject, that of eating pizza in (11b) a number-neutral one.

(11) a. inferred plural use (EXI-INF-PL)
E You and your friend are taking your kids to the playground in the morning. You see dozens of empty beer bottles there and tell your friend: "**You/one/they** have/has gathered here for a party."
b. inferred number-neutral use (EXI-INF-NN)
E You and a colleague decide to have a late lunch together in the lunchroom. On entering the room, you notice a very particular smell in the air and you tell your colleague: "**You/one/they** have/has eaten pizza in here."

In (12), scenarios are given for the vague and specific existential uses with a plural and a number-neutral interpretation.

(12) a. vague plural use (EXI-VAG-PL)
E A friend asks you: "What ever happened to that famous Latin American rebel leader?" You answer: "Oh, after he was released from prison, **you/one/they** lynched him."
b. vague number-neutral use (EXI-VAG-NN)
E You ring up your friend to tell them that the bag that they lost the day before has been found. You say: "**You/one/they** have/has found your bag in the park."
c. specific plural use (EXI-SPE-PL)
E Your friend is in the shower when their landline and their cellphone start ringing at about the same time. You tell your friend: "Hurry up! **You/one/they** are/is calling you on both phones!"
d. specific number-neutral uses (EXI-SPE-NN)
E You are in the bathroom when you hear knocking on the front door. You call out to your flat mate: "Can you run downstairs? **You/one/they** are/is knocking on the door."

All twelve uses illustrated in (7) to (12) are tested with two scenarios in our questionnaire. In the first part, one scenario for each use is presented in an arbitrary order. The second part contains the remaining twelve scenarios, again in a random order.

The instructions included in the questionnaire for judging acceptability are as follows: (i) do not hesitate to use the whole five-point scale, (ii) feel free to give equally (un)acceptable sentences the same score, (iii) rely on your own linguistic intuition for the judgments and not on what certain norms may say is (in)correct, (iv) answer in as instinctive a way as possible and do not change the scores afterward unless really necessary and (v) take all the time that is required. The questionnaire also collects some demographical information from participants (i.e. their date and place of birth, gender, native language(s) and knowledge of other languages) but emphasizes that participation is anonymous and voluntary. Moreover, participants are told that they can ask for their answers to be removed from the study within one month of taking part (on the basis of their demographical information).

Note, finally, that the English questionnaire was checked by a native speaker. The Dutch one was translated by the first author and the Afrikaans one by the second author. The two translations were checked by another native speaker, resulting in some minor alterations that made them fit the cultural context better but did not affect the different impersonal readings. The Afrikaans counterpart of (12d), for instance, does not refer to a two-story house.

## 2.2 Completion questionnaire

This questionnaire features the same twenty-four scenarios as the acceptability judgment questionnaire in the same random order. For the completion task, the HIP slot is simply left open, though. Participants are requested to fill in the blank in such a way that "the sentence is saying something about people in general or about people who are unknown to them or, in other words, whom they do not want to or are unable to identify in any specific way." This description is supposed to get them to employ an impersonalization strategy. We are aware that, in all our scenarios, the HIP slot functions as the subject of the sentence. This fact has two consequences. On the one hand, a wide range of strategies, such as the passive in (4c), are not really an option for participants. This "bias" toward HIPs is intentional, however: if the HIP-favorable blank in a particular scenario is mainly filled in with non-pronominal strategies, such as *mense* 'people' in (4b), we can safely conclude that HIPs are probably not the preferred impersonalization strategy for the use that the scenario instantiates. On the other hand, we need to provide the English and Dutch participants with all potential person-number forms of the verb so that they are free to select any impersonal-like subject to complete the sentence (this problem does not arise in Afrikaans because its verbs are not inflected for person or number). In addition, the questionnaire explicitly points out to participants that they should not let their answers depend on the verb form that happens to come first. It also moves the different forms of the verb around in the second scenario of the same use, to counterbalance any possible influence. In any case, no substantial differences are found between the answers for two scenarios of the same use.

The instructions included in the completion questionnaire are nearly identical to those for the acceptability judgment questionnaire. They too mention, among other things, relying on one's own linguistic intuition, answering as instinctively as possible and the voluntary and anonymous nature of participation. An important difference is that the completion questionnaire exemplifies the task beforehand with (13). The rationale behind including this (universal-internal-veridical) scenario is that our description of impersonalization in the previous paragraph may not be sufficiently clear.

(13) E Your father tells you that your grandmother has decided to take driving lessons and he expresses doubts about her chances of success. You reply: "…… am/are/is never too old to learn."

In the instructions, we point out to the participants that *one* would be in keeping with "people in general" or "people who are unknown to them or, in other words, whom they do not want to or are unable to identify in any specific way". We also say that *grandparents* or *she*, by contrast, would not and explain that *grandparents* would give us a sentence about a set of individuals still identifiable as displaying the particular feature of being grandparents and *she* a sentence about the specific individual of the grandmother in the scenario. Note that the Afrikaans and Dutch completion questionnaires give (13) in English, to illustrate the task with an impersonalization strategy that the two languages do not possess, i.e. 'one'. As the demographical data collected in the questionnaire confirms, our Afrikaans- and Dutch-speaking participants (see Section 2.4) can be expected to understand English. Unfortunately, we cannot avoid the use of an established impersonalization strategy in this way in the English questionnaire. Our choice of *one* (over, for instance, *you*) in the instructions may still be presumed to have only a limited impact on the participants, as the HIP is known to be very infrequent (e.g. van der Auwera et al. 2012: 57), especially in non-formal language. This assumption is confirmed by its infrequency in the results (see Section 3.2).

## 2.3 Earlier questionnaires

HIPs have been studied through questionnaires before. Siewierska & Papastathi (2011: 592-599), for one, examine the acceptability of the third person plural in nine European languages, including Dutch and English. Our study differs from theirs in several respects, though. First, we look at *all* main HIPs in Dutch, English *and* Afrikaans, add Gast & van der Auwera's (2013) universal-internal uses and their dimension of number to the mix and also test Siewierska & Papastathi's (2011: 593) speech act verb use, which they ignore because it "is so evidently tied to specific verbs rather than situations". Second, the items that we ask participants to assess for each scenario are limited to sentences with the main HIPs in the language, like in (7). The three options given by Siewierska & Papastathi (2011: 593-594) include one with 'they' and two with non-pronominal impersonalization strategies like passives, 'someone' and 'people', which can vary from scenario to scenario. This set-up suits their objective to evaluate the third person plural's acceptability in light of alternative and perhaps better strategies perfectly. Our goal, by contrast, is to get an idea of the whole functional potential of all HIPs in each impersonal use. Comparing them to other impersonalization strategies in the acceptability judgments may (partly) obscure the picture. The possible usage preference for the latter is, in fact, what the completion questionnaire is concerned with. Third, we have data for, on average, thirty-eight participants (see Section 2.4) per language. Siewierska & Papastathi's (2011) mean per language is fifteen and their results for Dutch and English in particular are based on "only" fourteen and ten people respectively. Given the large amount of languages investigated by them, these numbers are still remarkable, of course.

Another questionnaire-based project relevant here is Gast's (2017) online database for a typology of HIPs. Not unlike our completion task, it asks individual scholars – with the possible help of informants – to complete short scenarios instantiating various impersonal contexts with the appropriate HIP(s) of their (native) language (of study). Some of the scenarios even contain the same standard sentences from the literature, such as 'one only lives once' in (8b). Gast's (2017) aims are different from ours, though. The cross-linguistic information in his database is intended to be used

to uncover generalizations about the functional, as well as formal, variation in HIPs. The questionnaire is, in other words, not interested in the potentially minor fluctuations in usage which we wish to bring to light with our completion task and with considerable numbers of "naïve" participants. Furthermore, Gast (2017) focuses very much on HIPs and pays little attention to (their relation to) other impersonalization strategies. Our completion questionnaire, however, does not actually specify HIPs as its topic and thus seeks to answer the question to what extent the use of HIPs is (dis)preferred to that of alternative strategies in the range of impersonal contexts.

### 2.4 Delivery and participants

The data for Afrikaans was collected by asking undergraduate students of Afrikaans at the North-West University Potchefstroom (South Africa) to fill out the questionnaires on paper at the start of a first-, second- and third-year lecture in May 2016 (the third-year students were the second author's).[7] The acceptability judgment task and the completion task were each given to roughly half of the students in every class in a random manner. We emphasized that they were free to take part or not and that non-participation would not affect their studies in any way. The acceptability judgment questionnaire was completed by seventy-two students. We excluded two participants that identified as non-native speakers of Afrikaans and another twenty-two because they did not follow the instructions and only or mainly provided scores for the acceptable HIPs. The forty-eight students whose judgments were taken into account all had a date of birth between 1992 and 1997 and three quarters of them were female and the others male. This profile is identical to that of the seventy-five participants included in the completion questionnaire results. Of the eighty students who completed the task, four were left out because they were non-native speakers and one because his date of birth was in 1958 and he was much older than all the other participants.

To get data for Dutch, we asked lecturers at the University of Antwerp and Ghent University (Belgium) in November 2016 to send their undergraduate students of Dutch and/or English a message inviting them to participate in our study and complete an online version of the questionnaires.[8] The acceptability judgment task was disseminated to different year groups than the completion task in order to avoid the possibility that the same person filled out both questionnaires. The former was completed by forty-seven students, three of which were excluded for age or native language reasons (note that the issue of judgments only for the acceptable HIPs did not arise because the online form could deliberately not be submitted before all items were assigned scores). The remaining participants, of which roughly four fifths identified as female and the rest as male, were all born between 1989 and 1998. The latter questionnaire was filled out by thirty-nine students. Nobody was left out and, with dates of birth between 1989 and 1997 and a gender distribution of approximately seventy/thirty percent female/male, they have a very similar profile to the participants of the acceptability judgment task, as well as to our Afrikaans-speaking participants.

The data for English was obtained by inviting the first author's first-, second- and third-year

[7] For which we obtained ethical approval from the North-West University Ethics Committee and the Lancaster University Research Ethics Committee in April 2016. We want to thank Prof Thys Human and Mrs Janien Linde for sacrificing some of their lecture time for our study.

[8] This study, as well as the one for English, was approved by the Lancaster University FASS-LUMS Research Ethics Committee in November 2016. Special thanks are due to Dr Frank Brisard, who circulated the link to our acceptability judgment task among the first-year students of English at the University of Antwerp, to Dr Reinhild Vandekerckhove, who asked the second- and third-year students of Dutch at the University of Antwerp to fill out the completion task, and to Dr Timothy Colleman, who forwarded the completion questionnaire link to the first-year students of Dutch at Ghent University and the acceptability questionnaire one to the second-year students.

undergraduate students of English and/or linguistics at Lancaster University (United Kingdom) in December 2016 to complete an online version of the questionnaire on the virtual learning environment. We were able to target different people for the two tasks by making each of them visible to specific seminar groups on the virtual learning environment and we again made it clear that non-participation would have no impact on their studies. The acceptability judgment questionnaire was filled out by twenty-seven students, four of which were excluded because they were not native speakers of English. The rest of the participants had dates of birth between 1992 and 1998 and identified as female in about three quarters of the cases, as male in one sixth of the cases and as other[9] in two cases. The completion questionnaire was answered by thirty-three students. Not including the five participants whose native language was not English, they were all born between 1987 and 1998. Roughly four fifths of them were female and the others male. Our results for both tasks in all three languages thus come from people with more or less the same profile. There are, of course, substantial differences in the number of Afrikaans-, Dutch- and English-speaking participants. They can be attributed partially to the fact that the former were asked to fill out the questionnaires in class and to the low response rate among the latter (compared to the Belgian students). To assess our findings for the three languages, a statistical analysis is therefore of particular importance.

In view of the above, let us conclude this section by pointing out explicitly that, in the rest of the article, the terms "Dutch" and "English" are essentially used to refer to the established varieties of Belgian Dutch and British English. We are not aware of any recognized or supposed differences in the impersonal domain, say, between Belgian and Dutch Dutch or within British English (with the possible exception of impersonal *man* in Multicultural London English, see Cheshire 2013). In other words, our results may very well be generalizable to Dutch and English at large. Still, to be certain about this for, for instance, Dutch, we would need to replicate the present study at a university in the Netherlands.

### 2.5 Statistics[10]

For the acceptability judgment questionnaire, we give the following descriptive statistics: a mean, i.e. the average score of a particular HIP in a particular context, and a standard deviation ("std" for short), i.e. a measure of the extent to which the scores of all our participants vary for a specific HIP in a specific use. A standard deviation is low when, on the whole, the scores are close to the mean and high when they are more widely dispersed (see Rasinger 2013: 134-136). To contrast the scores of one HIP in two contexts (e.g. *one* in modal non-veridical universal-internal versus non-modal non-veridical universal-internal uses) or those of two HIPs in one context (e.g. *one* versus *you* in veridical universal-internal uses) with each other, we compute two-sided t-tests. They check whether one mean differs significantly from another one in consideration of their standard deviations without making any assumptions about the direction of the difference (see Baayen 2008: 81), which we do not really have. To compare the findings for the same language, we use dependent t-tests because the data comes from the same set of individuals. Independent t-tests are employed, for instance, to weigh the scores of *mens* against those of *men* because two different groups of people are responsible for the data (see Rasinger 2013: 200). When one result is contrasted with more than one other result in this manner (e.g. *jy* versus *je* versus *you* in the veridical universal-internal use), we Bonferroni-correct our typical level of significance of 0.05 by dividing it by the

[9] Note that the South African and Belgian students were given this third gender option as well.
[10] The statistical analysis has been carried out with SPSS (IBM Corp. 2013).

number of comparisons made (e.g. 0.017 for the second person singular in a particular context in the three languages). This procedure minimizes the risk that too much weight is given to one (or more) of a number of t-tests producing a p-value lower than 0.05 since it (or they) may merely be due to chance (see Baayen 2008: 114). To examine the impact of one or two variables on the acceptability of a HIP, finally, we calculate one- or two-way analyses of variance or ANOVAs (see Rasinger 2013: 210-217). The effects of the dimensions of number and (un)knownness on the acceptability of *ze*, for example, are checked through a two-way ANOVA and, more specifically, one with repeated measures because the information comes from just one set of participants (see Baayen 2008: 264). Post hoc Bonferroni-corrected t-tests are calculated afterward to compare particular contexts (e.g. *ze* in vague versus inferred versus specific number-neutral uses).

For the completion questionnaire, we present the absolute numbers of all the different impersonalization strategies, as well as their percentages. Fisher's exact tests are computed to assess the (dis)similarities between uses. They can reveal whether or not the varying proportions of one HIP versus another or of HIPs versus non-pronominal alternatives, i.e. the categorical dependent variables, in these uses, i.e. the independent variables, differ significantly, even with low raw numbers (see Baayen 2008: 122). When more than two contexts are compared, we once more apply a Bonferroni correction to our normal level of significance of 0.05.

## 3 Universal-internal uses

### 3.1 Acceptability judgment task

In Table 1, we give the mean scores and standard deviations of every main HIP in Afrikaans, Dutch and English (see the second row from the top) for all universal-internal uses (see the leftmost column).

Table 1: Descriptive statistics of the universal-internal uses in the acceptability judgment task

| Use | | Afrikaans | | | | Dutch | | | English | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *jy* | *mens* | *'n mens* | *hulle* | *je* | *men* | *ze* | *you* | *one* | *they* |
| UNI-INT-NVER-NMOD | mean | 4.83 | 4.05 | 4.54 | 1.88 | 4.79 | 2.97 | 1.12 | 4.78 | 3.15 | 1.43 |
| | *std* | *0.64* | *1.36* | *0.81* | *1.31* | *0.69* | *0.87* | *0.28* | *0.55* | *1.09* | *0.83* |
| UNI-INT-NVER-MOD | mean | 4.67 | 3.84 | 4.68 | 1.94 | 4.79 | 3.36 | 1.28 | 4.76 | 3.46 | 1.20 |
| | *std* | *0.87* | *1.45* | *0.59* | *1.31* | *0.70* | *1.16* | *0.70* | *0.48* | *1.19* | *0.58* |
| UNI-INT-VER | mean | 4.79 | 3.98 | 4.55 | 1.69 | 4.83 | 3.21 | 1.21 | 4.70 | 3.22 | 2.24 |
| | *std* | *0.65* | *1.31* | *0.86* | *1.08* | *0.48* | *1.01* | *0.67* | *0.47* | *1.40* | *1.34* |

No statistics are needed to conclude from Table 1 that, unsurprisingly (see Siewierska & Papastathi 2011), the third person plural is considered much less acceptable – or, in essence, unacceptable – in universal-internal uses than the other HIPs in the three languages: its score averages around 1.84 in Afrikaans, 1.20 in Dutch and 1.62 in English. For that reason, 'they' is ignored in the remainder of Section 3. The other results in Table 1 are presented in a visually more accessible way in Figure 1. Note that: (i) the Afrikaans HIPs are shown in black with round marks, the Dutch ones in dark gray with triangular marks and the English ones in light grey with square marks, (ii) the second person singular pronouns all have solid lines and the other HIPs dashed ones, (iii) the uses are

ordered in accordance with Gast & van der Auwera's (2013) semantic map,[11] (iv) the lines connecting these uses are meant to make it easier for the reader to tell partially overlapping dots apart and should not be interpreted as an indication of some sort of linear development; (v) only the upper half of the five-point scale is included here.

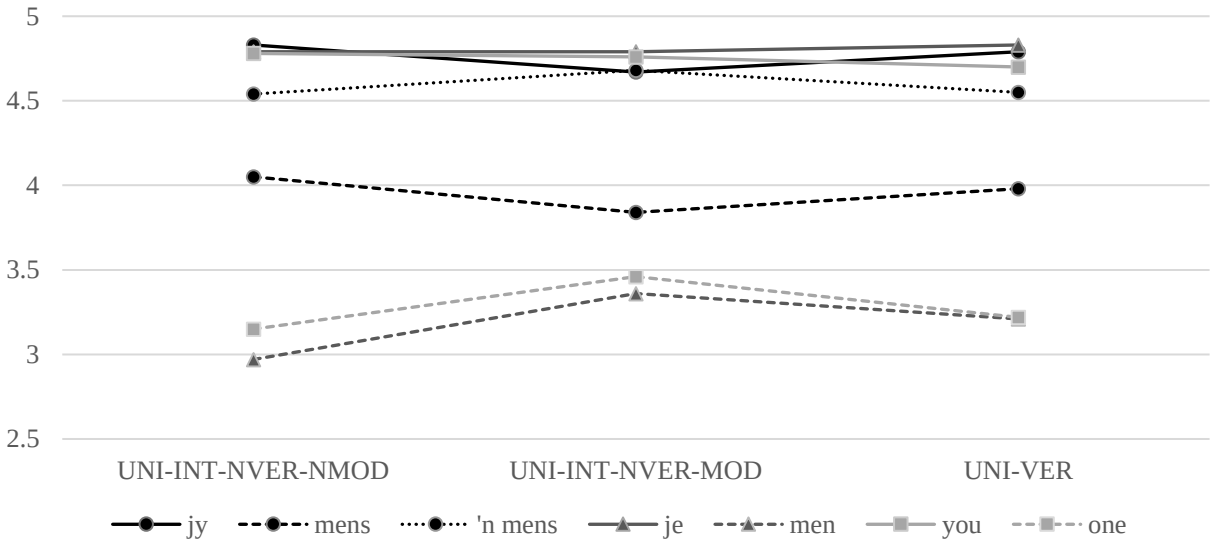


Figure 1: Distribution of the mean scores of 'you', 'man' and 'one' for the universal-internal uses

Let us now look at these findings from three points of view: what are the (dis)similarities between the HIPs within each language, how do the same types of HIP (e.g. the 'man'-pronouns) behave in different languages and, finally, is there any variation in the acceptability of the HIPs between the three universal-internal uses?[12]

For both Dutch and English, the answer to the first question is that the second person singular is regarded as more acceptable than its competitor. Its scores, averaging 4.80 in Dutch and 4.75 in English, are significantly higher than those of *men* and *one*, i.e. around 3.18 and 3.28 respectively, in all three universal-internal uses ($p < 0.05$ for all t-tests). This result should not come as a surprise: the two languages have been shown to favor the second person singular before (e.g. van der Auwera et al. 2012: 57, De Hoop & Tarenskeen 2015). Moreover, a particular feature of 'you' is that "the addressee is presupposed or invited to empathize with the (set of) protagonist(s) about which some statement is made" (Gast et al. 2015: 150). This effect could have played a role in the second person singular's higher acceptability given the familiarity between the speech participants in most of the

[11] The explanation for this arrangement of the universal-internal uses goes as follows. Some impersonalization strategies are restricted to non-veridical non-modal contexts (e.g. German *jemand* 'someone'), others can occur in the two non-veridical contexts but not in veridical ones (e.g. the Russian infinitive) and yet other strategies can be employed in all three universal-internal contexts (e.g. Spanish *uno* 'one'). Crucially for the order of the uses, no form of impersonalization seems to be able to appear solely in non-veridical non-modal contexts and veridical ones (see Gast & van der Auwera 2013: 149).

[12] The following discussion only mentions the results of our (post hoc) Bonferroni-corrected t-tests but, for each language, we also performed a two-way ANOVA testing the influence of the different HIPs and the different contexts on acceptability. In Afrikaans and English, there is no interaction between the two and the uses themselves have no significant effect either ($p > 0.05$ in both cases). Only the choice of HIP is found to have an impact ($p < 0.05$). In Dutch, by contrast, we notice that the type of context, as well as the type of HIP, has a significant effect ($p < 0.05$ across the board).

scenarios. *Men* and *one* can, of course, act as a universal-internal HIP. The former's lower acceptability may be explained by its potentially disappearing and formal character (e.g. Duinhoven 1990, Weerman 2006), which clashes with the colloquial nature of the scenarios.[13] The latter's is probably also partly due to a conflict in formality, as well as to the HIP's association in non-modal contexts with verbs of perception and cognition (see van der Auwera et al. 2012: 46), which do not appear in our questionnaire.

For Afrikaans, the picture is somewhat more complex. *Jy* and *'n mens* score significantly higher than *mens* in all universal-internal uses ($p < 0.017$ across the board). The more grammaticalized form of the 'man'-pronoun, though not a recent development at all (see Van Olmen et al. 2018: §3.2), is considered less acceptable than the less grammaticalized one. We will briefly return to this issue in Section 3.2. No substantial differences are found between *jy* and *'n mens* ($p > 0.017$), except for in non-veridical non-modal uses like (14) (cf. 8a).

(14) A *As **jy**/**'n mens** na Engeland toe gaan, is dit noodsaaklik om 'n reënjas in te pak.*
'If one goes to England, it is necessary to pack a raincoat.'

We have no real explanation for the fact that, in this specific context, the second person singular is regarded as more acceptable than the 'man'-pronoun.[14]

The comparison of 'you' in Afrikaans, Dutch and English reveals that there are no significant differences in acceptability between the languages for any universal-internal use ($p > 0.017$ in each case). The results do, in other words, not reflect the supposedly even stronger preference for the second person singular in Dutch than in English (see van der Auwera et al. 2012: 57). We will examine in Section 3.2 whether this preference is perhaps more a matter of usage than of acceptability. Contrasting the 'man'-pronouns with each other confirms what must already be evident from the language-internal comparison: *men* scores significantly lower in acceptability than not only *'n mens* but also *mens* ($p < 0.025$) or, put differently, the 'man'-pronoun seems to be more "alive" in Afrikaans than in Dutch. Maybe not unrelatedly, unlike *men*, which Afrikaans (as an off-shoot of

[13] One of the reviewers, who we assume is a speaker of Dutch Dutch, still seems to consider *men*'s levels of acceptability relatively high. He or she hypothesizes that they are at least in part the result of the language background of the participants: it would not come as a surprise to him or her if the disappearing HIP *men* persisted for longer in the more conservative variety that Belgian Dutch is said to be. We acknowledge the possibility but have to leave its examination for future research.

[14] *'n Mens* in (14) is, to some extent, vague between an interpretation as a HIP and one as the indefinite noun phrase 'a human being'. This ambiguity, which is hard to get around with a grammaticalizing 'man'-pronoun, is found in the other universal-internal uses as well, however (see footnote 6). The only way to avoid it would be to include a possessive or reflexive in the sentence (see Donaldson 1993: 139-140, Van Olmen et al. 2018: §3.1): *'n mens* takes second person singular forms as a HIP – such as *jouself* 'yourself' in *'n mens moet jouself beheers* 'one should restrain oneself' (and not 'a human being should restrain him- or herself') – but third person singular masculine ones as an indefinite noun phrase – such as *sy* 'his' in *'n mens moet sy verstandelike vermoëns gebruik om God te leer ken* 'a human being should use his or her intellectual abilities to get to know God' (and not 'one should use one's intellectual abilities to get to know God'). This solution runs the risk of favoring the HIP *jy*, though. In any case, one of the reviewers points out that the second person singular's higher acceptability in (14) might be due to its stronger disposition to bind the implicit argument of the main clause. In *if you go to England, it is necessary to pack a raincoat*, the packer is always interpreted as identical to the person traveling to England. A sentence such as *if someone goes to England, it is necessary to pack a raincoat*, by contrast, allows a second reading in which the traveler and the packer are different people. The question now is whether ambiguous *'n mens* resembles *someone* in permitting such an interpretation. The second author's intuitions suggest that it does not, though more research is obviously needed. Moreover, even if it did, we are not convinced that the rather awkward second reading would arise during the questionnaire and affect the participants' ratings.

Early Modern Dutch) lost for good in the early 20th century, (*'n*) *mens* is an innovation of the last three hundred years (see Van Olmen et al. 2018: §3.2). The only exception is the non-veridical modal use, in which, for unclear reasons, the Dutch 'man'-pronoun's lower acceptability compared to *mens* is not significant at a Bonferroni-corrected level ($p = 0.03$). *One*, finally, has no real cognates as a HIP in the other two West Germanic languages. It is clear, though, from Figure 1 that, as far as acceptability is concerned, it most closely resembles *men*.

The question whether any variation exists in the scores for HIPs between the three universal-universal contexts can be answered in the negative for both Afrikaans and English. No substantial differences are found for *jy*, *'n mens*, *mens*, *you* or *one* ($p > 0.017$ in every case). The same holds for Dutch *je* and, for the most part, for *men* as well. These results are, in a sense, in line with Gast & van der Auwera (2013: 143-149), who point out that the second person singular, 'one'-pronouns and 'man'-pronouns are general universal-internal HIPs (with the latter often also having non-universal-internal uses, as discussed in Section 4). The Dutch 'man'-pronoun does score significantly lower in the non-veridical non-modal use than in the non-veridical modal one ($p < 0.017$). Neither use differs from the veridical one, though ($p > 0.017$). Unfortunately, we are currently not able to offer an explanation for this finding.

## 3.2 Completion task

In the analysis of this questionnaire's results, we first distinguish strategies that match the description of people in general or people who are not known or whom one would not be able to or want to identify in any specific way from answers that do not, i.e. the "relevant" instances of impersonalization from the "irrelevant" ones. Examples of the latter are the personal pronouns *ze* 'she' in (15a) (cf. 7) and *I* in (15b). Considering the entirely open character of the completion task and the fairly complex functional domain that it deals with, we are not surprised by the occurrence of such cases in the data.

(15) a. D *Ze is een beetje terneergeslagen maar ze blijft ervan overtuigd dat **ze** nooit mag opgeven.*
'She is a little down but she still believes that she should never give up.'

b. E *Your sister gets the carton of milk out of the fridge and takes a big gulp. She then looks at you worryingly and ask you: "What happens if **I** drink sour milk?"*

Table 2 gives, for each language (see the top row) and use (see the leftmost column), the absolute numbers (#) of relevant and irrelevant answers, as well as their proportions (%).

Table 2: Impersonalization strategies versus other answers for the universal-internal uses in the completion task

| Use | | Afrikaans | | Dutch | | English | |
|---|---|---|---|---|---|---|---|
| | | relevant | irrelevant | relevant | irrelevant | relevant | irrelevant |
| UNI-INT-NVER-NMOD | # | 133 | 17 | 78 | 0 | 52 | 4 |
| | *%* | *88.67* | *11.33* | *100.00* | *0.00* | *92.86* | *7.14* |
| UNI-INT-NVER-MOD | # | 111 | 39 | 72 | 6 | 54 | 2 |
| | *%* | *74.00* | *26.00* | *92.31* | *7.69* | *96.43* | *3.57* |
| UNI-INT-VER | # | 136 | 14 | 78 | 0 | 56 | 0 |
| | *%* | *90.76* | *9.33* | *100.00* | *0.00* | *100.00* | *0* |

We acknowledge that there is some variation in the frequency of cases like (15), not only between the languages (e.g. Afrikaans seems to have more irrelevant instances than Dutch and English) but also between the contexts (e.g. the veridical one appears to encourage the use of an impersonalization strategy the most) and that these differences merit a closer look. However, for reasons of space and focus, the present article concentrates on the answers that actually satisfy the conditions of the completion task (see Van Olmen & Breed 2018: 13-15 for a more in-depth discussion of the variation between the uses in Afrikaans).

The next step in the analysis is the classification of the impersonalization strategies. We obviously set the HIPs under investigation apart from the other strategies. *Jy*, *je* and *you* are labeled as 'you', *mens* and *men* as 'man', *'n mens* as 'a man' and *one* as 'one'. For the remaining answers, we make a distinction between miscellaneous and major ones. The former constitute a single category that groups together every form of impersonalization accounting for less than 0.5% of all our responses (including those for the non-universal-internal uses). It comprises, among other things, the three attestations of the Afrikaans indefinite singular noun phrase *'n persoon* 'a person' in non-veridical contexts like (16a) and the use of the HIP 'we' in the veridical context in (16b) (cf. 8b), which occurs two times in Afrikaans, five times in Dutch and three times in English (see Siewierska 2004: 211 on the first person plural as a HIP).

(16) a. A *Jou ma neem nou al vir ses weke lank Franse taallesse. Jy vra haar of sy vir jou 'n stukkie teks na Frans sal vertaal vir werk. Sy is bietjie geïrriteerd en antwoord: "**'n Persoon** kan nie 'n taal in ses weke leer nie."*
'Your mother has been taking French classes for six weeks now. You ask her whether she could translate a text into French for you for work. She is a little bit irritated and replies: "A person cannot learn a language in six weeks."'

b. D ***We** leven maar een keer.*
'We only live once.'

In the universal-internal uses, we come across two types of major impersonalization strategy, which is defined as being more frequent in our entire dataset than a miscellaneous one and is categorized separately: indefinite pronouns ("idf.pro" for short) and the indefinite plural noun phrase 'people', as in (17a) (cf. 15b) and (17b) respectively. Note that the former appear to be characteristic of the non-veridical non-modal use in every language, which confirms Gast & van der Auwera's (2013: 47) argument that, as an existential quantifier, an indefinite pronoun can be employed in this context (and in their number-neutral existential one).

(17) a. E *What happens if **someone** drinks sour milk?*

b. D *Je partner zegt dat hij/zij bezorgd is dat jullie dochter, die altijd leest, misschien een beetje eenzaam is. Jij bent het helemaal niet eens met je partner en zegt: "**Mensen** zijn nooit eenzaam in het gezelschap van een goed boek."*
'Your partner expresses worry that your daughter, who is always reading, might be a little lonely. You completely disagree and retort: "People are never alone with a good book."'

Table 3 presents the absolute numbers and the proportions of the various impersonalization strategies discussed so far (see the top row) in all three languages (see the leftmost column) and all three universal-internal uses (see the second column from the left).

Table 3: Distribution of the impersonalization strategies for the universal-internal uses in the completion task

| Language | Use | | you | man | a man | one | idf.pro | people | misc. |
|---|---|---|---|---|---|---|---|---|---|
| Afrikaans | UNI-INT-NVER-NMOD | # | 30 | 72 | 23 | 0 | 6 | 0 | 2 |
| | | *%* | *22.56* | *54.14* | *17.29* | *0.00* | *4.51* | *0.00* | *1.50* |
| | UNI-INT-NVER-MOD | # | 14 | 73 | 18 | 0 | 2 | 1 | 3 |
| | | *%* | *12.61* | *65.77* | *16.22* | *0.00* | *1.80* | *0.90* | *2.70* |
| | UNI-INT-VER | # | 11 | 91 | 28 | 1[15] | 1 | 2 | 2 |
| | | *%* | *8.09* | *66.91* | *20.59* | *0.74* | *0.74* | *1.47* | *1.47* |
| Dutch | UNI-INT-NVER-NMOD | # | 73 | 1 | 0 | 0 | 3 | 1 | 0 |
| | | *%* | *93.59* | *1.28* | *0.00* | *0.00* | *3.85* | *1.28* | *0.00* |
| | UNI-INT-NVER-MOD | # | 65 | 7 | 0 | 0 | 0 | 0 | 0 |
| | | *%* | *90.28* | *9.72* | *0.00* | *0.00* | *0.00* | *0.00* | *0.00* |
| | UNI-INT-VER | # | 64 | 6 | 0 | 0 | 2 | 1 | 5 |
| | | *%* | *82.05* | *7.69* | *0.00* | *0.00* | *2.56* | *1.28* | *6.41* |
| English | UNI-INT-NVER-NMOD | # | 40 | 0 | 0 | 8 | 3 | 1 | 0 |
| | | *%* | *76.92* | *0.00* | *0.00* | *15.38* | *5.77* | *1.92* | *0.00* |
| | UNI-INT-NVER-MOD | # | 40 | 0 | 0 | 10 | 0 | 4 | 0 |
| | | *%* | *74.07* | *0.00* | *0.00* | *18.52* | *0.00* | *7.41* | *0.00* |
| | UNI-INT-VER | # | 43 | 0 | 0 | 8 | 1 | 1 | 3 |
| | | *%* | *76.79* | *0.00* | *0.00* | *14.29* | *1.79* | *1.79* | *5.36* |

The data is given in a visually more accessible way in Figure 2, in which the UNI-INT component of the uses is left out for practical reasons, the solid areas in the bars correspond to our main HIPs and the striped ones to other impersonalization strategies.

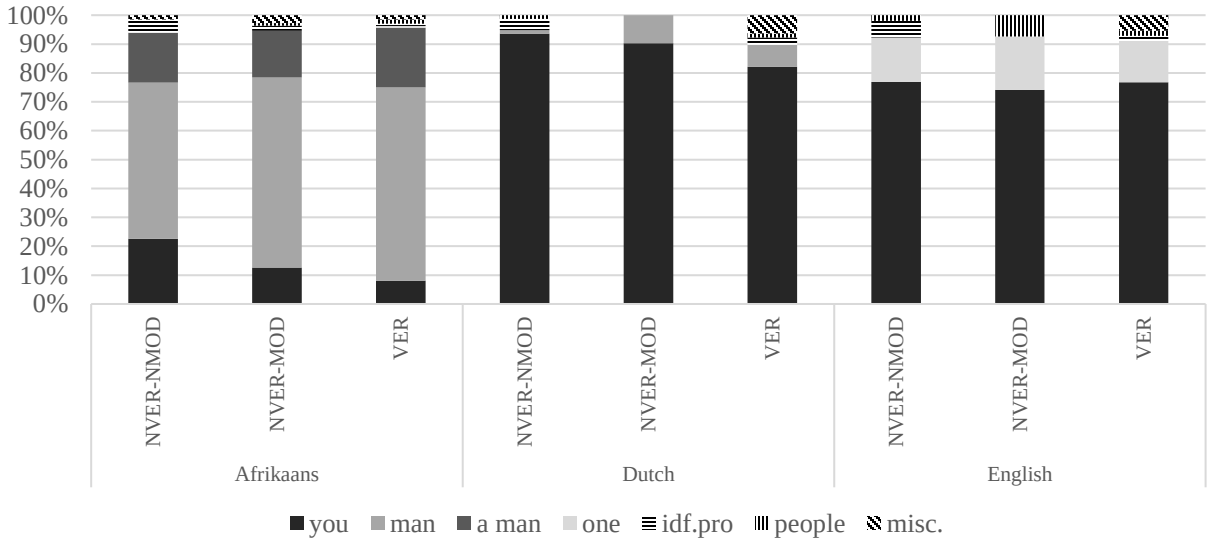


Figure 2: Distribution of the impersonalization strategies for the universal-internal uses in the completion task

Let us now discuss these results in terms of the distribution of pronominal versus non-pronominal impersonalization strategies on the one hand and the distribution of 'you' versus '(a) man' or 'one' on the other. For the sake of completeness, we should first stress, however, that, as expected based on the literature and the acceptability judgment questionnaires (see Section 3.1), 'they' is not used to complete any universal-internal scenario in any language.

[15] See Van Olmen & Breed (2018: 17) on the potential Anglicism *een* 'one' in Afrikaans.

It is immediately clear from Figure 2 that HIPs are very strongly preferred to other forms of impersonalization in universal-internal contexts. In Afrikaans, they make up 93.98% of the non-veridical non-modal answers, 94.59% of the non-veridical modal ones and 97.79% of the veridical ones. In Dutch, the respective proportions are 94.87%, 100% and 96.15% and, in English, 92.31%, 92.59% and 96.42%. Not surprisingly with these percentages, no significant differences exist between the languages or the uses ($p > 0.05$ for all Fisher's exact tests). In Section 4.2, we will compare these findings with those for non-universal-internal contexts to see whether they too exhibit such a marked partiality to HIPs.

As to the relation between 'you' and 'man'/'one', the Dutch and English usage data is in line with the acceptability judgments, in which the second person singular consistently scores significantly higher than its competitor. *Je* accounts for 93.59% of the non-veridical non-modal contexts, 90.28% of the non-veridical modal ones and 82.05% of the veridical ones whereas *men* makes up 1.28%, 9.72% and 7.69% respectively. Similarly, *you* and *one* are responsible for 76.92% versus 15.38% of the non-veridical non-modal answers, 74.07% versus 18.52% of the non-veridical modal ones and 76.79% versus 14.29% of the veridical ones. The completion task results for Afrikaans, by contrast, paint a drastically different picture from the acceptability judgments. According to the latter, *mens* is substantially less acceptable than *jy* and *'n mens*, which are roughly equal. In the completion task data, however, *mens* is more common than *jy* and *'n mens* combined, making up 54.14% of the non-veridical non-modal cases, 65.77% of the non-veridical modal ones and 66.91% of the veridical ones. *Jy* accounts for 22.56%, 12.61% and 8.09% respectively and *'n mens* for 17.29%, 16.22% and 20.59% respectively. In other words, as far as the 'man'-pronoun is concerned, the less grammaticalized form is considered more acceptable but, as suggested by the completion task, it is the more grammaticalized form that people really employ in everyday language. The explanation for this difference, which is explored in more detail in Kirsten (2016: 192-193) and Van Olmen & Breed (2018: 22), probably lies in the fact that the typically more conservative character of written language influences participants seeing *'n mens* next to *mens* in writing in the acceptability judgment questionnaire. In the completion task, neither variant is given and participants, who are asked to fill in the blanks as instinctively as possible, are thus free of this bias and choose the form that they normally use themselves.

A comparison of the three languages demonstrates that the distribution of the main HIPs in Afrikaans differs significantly from that in both Dutch and English for every universal-internal use ($p < 0.017$ across the board). The completion task thus suggests that Afrikaans is a 'man'-prominent language in terms of usage[16] and that it is, in that sense, closer to German (see van der Auwera et al. 2012: 57; Gast 2015: 8 on *man* as the epitomizing impersonalization in this language) than to Dutch and English, which count as 'you'-prominent. For these two languages. van der Auwera et al.'s (2012: 57) one-text parallel corpus study indicates that *je* is even more widespread in Dutch than *you* is in English. There is, however, little support for this claim in our completion questionnaire results (or, for that matter, in our acceptability judgments, where no substantial variation is found between the second person singular in Dutch and its counterpart in English). Only in the non-veridical non-modal use is the distribution of *je* versus *men* significantly different from that of *you* versus *one* ($p = 0.0024$) and is 'you' employed more frequently in Dutch than in English.

As shown in Section 3.1, the answer to the question of whether the type of universal-internal use affects the main HIPs' acceptability is essentially negative for all three languages under investigation. The results of the completion tasks suggest that the type of context does not influence the

[16] Contra Kirsten's (2016: 199) corpus-based claim that the second person singular occurs much more frequently as a HIP than the 'man'-pronoun in Afrikaans. Her analysis is discussed in more detail in Van Olmen & Breed (2018: 22).

HIPs' usage either in Dutch and English ($p > 0.017$ in all cases). *Men*'s significantly lower acceptability in the non-veridical non-modal use compared to the non-veridical modal one – i.e. the only exception – is, in other words, not mirrored by a similar substantial difference in frequency of use. This fact could be taken as an indication that the former finding is coincidental. The same might be true of the only significant difference in the two questionnaires for Afrikaans: *jy* decreases from from 22.56% in non-veridical non-modal contexts to 8.09% in veridical ones while *mens* goes from 54.14% to 66.91% and *'n mens* stays more or less the same at 17.29% and 20.59% respectively ($p = 0.001$). Still, in future research, it would perhaps be useful to take a closer look at the pragmatics of the second person singular versus the 'man'-pronoun in Afrikaans and in Dutch and see whether it could explain some of these results (cf. Deringer et al. 2015, Gast et al. 2015).

### 3.3 Interim conclusion

It is evident from both questionnaires that the third person plural cannot serve as a universal-internal HIP in any of the three languages. The second person singular and the 'man'- and 'one'-pronouns in Afrikaans, Dutch and English, by contrast, are all-purpose universal-internal HIPs according to the acceptability judgment task, with no real differences between contexts. The completion task, which shows that pronominal strategies are strongly preferred usage-wise for universal-internal impersonalization, confirms this overall lack of variation in the HIPs between the different uses. It does suggest that the languages diverge in the frequency with which the HIPs are employed. The lower acceptability of *men* and *one* is reflected in their limited occurrence in the completion questionnaire results. Both findings can be attributed to the fairly formal character of these HIPs and to the possible gradual disappearance of the 'man'-pronoun in Dutch. For 'you', there *is* a discrepancy between the two tasks: *jy*, *je* and *you* all appear to be equally highly acceptable but, unlike Dutch and English, Afrikaans does not seem to use the second person singular very frequently. The former languages can be characterized as 'you'-prominent, with little or no differentiation between them, and the latter language as 'man'-prominent. For (*'n*) *mens* too, the completion questionnaire points in another direction than the acceptability judgment one. *'n Mens* is considered as acceptable as *jy* and as much more acceptable than *mens*. In usage, however, it is the more grammaticalized form that occurs much more often than the less grammaticalized one. This difference probably results from the acceptability judgment task's presentation of the two variations alongside each other in writing, which tends to be more conservative.

## 4 Non-universal-internal uses

### 4.1 Acceptability judgment task

In Table 4, we give the mean scores and standard deviations of every main HIP in Afrikaans, Dutch and English for all non-universal-internal uses.

Table 4: Descriptive statistics of the non-universal-internal uses in the acceptability judgment task

| Use | | Afrikaans | | | | Dutch | | | English | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *jy* | *mens* | *'n mens* | *hulle* | *je* | *men* | *ze* | *you* | *one* | *they* |
| UNI-EXT | mean | 2.25 | 3.10 | 3.38 | 4.90 | 1.49 | 4.41 | 4.74 | 2.30 | 2.13 | 4.93 |
| | *std* | *1.54* | *1.61* | *1.51* | *0.37* | *0.79* | *0.74* | *0.55* | *1.13* | *1.00* | *0.25* |
| EXI-COR | mean | 1.23 | 1.41 | 2.10 | 4.85 | 1.02 | 3.53 | 4.81 | 1.11 | 1.39 | 4.93 |
| | *std* | *0.64* | *0.83* | *1.31* | *0.63* | *0.15* | *0.96* | *0.47* | *0.38* | *0.58* | *0.33* |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| EXI-VAG-PL | mean | 1.42 | 1.79 | 2.34 | 4.75 | 1.03 | 4.05 | 4.69 | 1.07 | 1.20 | 4.37 |
| | *std* | *0.98* | *1.25* | *1.32* | *0.68* | *0.18* | *0.97* | *0.57* | *0.25* | *0.45* | *1.04* |
| EXI-VAG-NN | mean | 1.45 | 1.58 | 3.67 | 4.66 | 1.13 | 3.98 | 4.71 | 1.30 | 1.35 | 4.39 |
| | *std* | *1.07* | *1.02* | *1.31* | *0.82* | *0.47* | *0.94* | *0.58* | *0.84* | *0.67* | *1.08* |
| EXI-INF-PL | mean | 1.41 | 1.48 | 1.84 | 4.74 | 1.03 | 3.13 | 4.65 | 1.04 | 1.33 | 4.02 |
| | *std* | *0.97* | *0.94* | *1.15* | *0.64* | *0.23* | *1.11* | *0.68* | *0.21* | *0.73* | *1.06* |
| EXI-INF-NN | mean | 2.04 | 1.49 | 3.17 | 4.45 | 1.24 | 2.81 | 4.71 | 1.48 | 1.30 | 3.87 |
| | *std* | *1.42* | *0.95* | *1.40* | *1.01* | *0.63* | *1.13* | *0.67* | *0.94* | *0.73* | *1.28* |
| EXI-SPE-PL | mean | 1.21 | 1.74 | 2.38 | 4.54 | 1.03 | 3.31 | 4.70 | 1.02 | 1.20 | 4.11 |
| | *std* | *0.68* | *1.21* | *1.36* | *0.94* | *0.18* | *1.13* | *0.58* | *0.15* | *0.54* | *1.12* |
| EXI-SPE-NN | mean | 1.11 | 1.61 | 3.83 | 3.44 | 1.03 | 3.66 | 4.53 | 1.00 | 1.35 | 3.91 |
| | *std* | *0.50* | *0.99* | *1.19* | *1.45* | *0.18* | *0.97* | *0.80* | *0.00* | *0.79* | *1.07* |
| SAV | mean | 1.72 | 1.72 | 2.15 | 4.91 | 1.06 | 4.23 | 4.67 | 1.26 | 1.72 | 4.78 |
| | *std* | *1.24* | *1.11* | *1.31* | *0.46* | *0.32* | *0.85* | *0.65* | *0.71* | *0.89* | *0.47* |

No statistics are required to conclude from Table 4 that, unsurprisingly (see Gast & van der Auwera 2013: 143-149), 'you' and 'one' are considered much less acceptable or, basically, unacceptable in non-universal-internal uses. Ignoring the universal-external use, we find that the mean scores of *jy* range from 1.11 to 2.04, those of *je* from 1.02 to 1.24 and those of *you* and *one*, respectively, from 1.00 to 1.48 and from 1.20 to 1.72. The same can be said of the more grammaticalized form of the 'man'-pronoun in Afrikaans: the acceptability of *mens* varies from to 1.41 to 1.79. Most of these HIPs, particularly the latter one, score (slightly) higher for the universal-external use: the respective levels of acceptability of *jy*, *mens*, *je*, *you* and *one* are 2.25, 3.10, 1.49, 2.30 and 2.13. We believe that these scores, which remain fairly low, arise from the difficulty of excluding an internal reading completely in a scenario like (18). Seeing the clauses with the universal-internal HIPs on paper in the acceptability judgment task could have encouraged some participants to somehow still identify with the set of referents, however strange the resulting sentence may be in context. Support for this interpretation of the facts comes from the completion task, where universal-internal HIPs are nearly absent from the universal-external data (see Section 4.2), and from discussions with native speakers of Afrikaans about its 'man'-pronoun (see Van Olmen & Breed 2018: 12).

(18) E A friend who is going to Greece for business asks you whether they should rent a car to get around. Based on your own negative experience in that country, you reply: "I don't think that is a good idea. In Greece, **you/one** drive/drives quite unpredictably."

The above explanation also applies to *'n mens* in the universal-external use. The less grammaticalized form of the Afrikaans 'man'-pronoun exhibits some other particular behavior that should be accounted for, though. In the existential-corporate, existential-plural and speech act verb uses, its acceptability ranges from 1.84 to 2.38. In the existential-number-neutral ones, by contrast, it varies from 3.17 to 3.83, which is significantly higher than its scores for the existential-plural counterparts ($p < 0.017$ across the board). Importantly, these findings should not be taken as an indication that *'n mens* can serve as an existential HIP. It is more plausible that the variation is due to the potential interpretation of the less grammaticalized 'man'-pronoun as an indefinite singular noun phrase with the meaning 'a human' in number-neutral uses like (19a) (cf. 11b), even though the resulting sentences are again semantically quite marked.[17] In existential-plural contexts such as (19b) (cf. 11a),

[17] Section 4.2 confirms that 'a human' is indeed a peculiar interpretation in any case: in the completion questionnaire data, it is attested only once in non-universal-internal contexts and, more specifically, in an existential-vague-number-neutral one.

as well as in the necessarily plural existential-corporate and speech act verb uses (see Siewierska & Papastathi 2011: 583, 585), such a reading is impossible. Our understanding of the results is substantiated by the consistently low scores of *mens*, which is less acceptable than *'n mens* in all eight uses at issue ($p < 0.006$ for all t-tests) and on which the dimension of number has no impact at all ($p > 0.05$ for each t-test). If *'n mens* truly had existential uses, which have been shown to develop out of universal-internal ones in 'man'-pronouns (see Giacalone Ramat & Sansò's 2007: 106), its more grammaticalized variant would be expected to have them as well. Not uninterestingly, *'n mens*'s behavior does suggest that the scenarios designed to test the dimension of number are effective.

(19) a. A ***'n Mens*** *het springmielies hierbinne gemaak.*
'A human has made popcorn in here.'
b. A ?****'n Mens*** *het hier bymekaargekom vir 'n partytjie.*
'A human has gathered here for a party.'

For the aforementioned reasons, the second person singular, (*'n*) *mens* and *one* are not considered any further in the remainder of the present section. The only HIP with universal-internal uses that *is* included here is *men*. The acceptability of the Dutch 'man'-pronoun in non-universal-internal contexts goes from 2.81 to 4.41. In other words, notwithstanding the variation, it is clear that *men* is more evolved in its (semantic) grammaticalization than (*'n*) *mens*. In this (and other) respect(s), Dutch can be considered a typical Standard Average European language, like German and French (see van der Auwera 2011): the 'man'-pronouns in these languages have the functional potential to convey every universal or existential impersonal meaning, as shown by Giacalone Ramat & Sansò (2007: 123-127) and Siewierska (2011: 71). Afrikaans, by contrast, may resemble German in its preference for 'man' (to 'you') in universal-internal uses but is closer to, for instance, Frisian (see Hoekstra 2010) in that (*'n*) *mens* has not (yet?) developed beyond the first, universal-internal stage of the normal evolution of 'man'-pronouns (see Giacalone Ramat & Sansò 2007: 106). For an in-depth discussion of the formal and functional (dis)similarities between the West Germanic 'man'-pronouns from the point of view of grammaticalization, we refer to Van Olmen & Breed (2018: 20-21) and Van Olmen et al. (2018: §4).

The results for *men* and for the third person plural in all three languages are given in a visually more manageable manner in Figure 3. Most of the form-related remarks about Figure 1 apply here as well (see Section 3.1). Note, in addition, that 'they' is indicated with solid lines and *men* with a dashed one and that the uses are ordered roughly in line with Siewierska & Papastathi's (2011) and Gast & van der Auwera's (2013) semantic maps.[18]

[18] The universal-external use is linked to the universal-internal-veridical one in Gast & van der Auwera (2013) and to the personal use of 'they', which is not taken into account here, in Siewierska & Papastathi (2011). The existential-corporate use is between the universal-external one and the truly impersonal ones in both maps. As discussed in Section 1, the latter contexts are divided differently by the two pairs of scholars. The arrangement in Figure 3, with the vague contexts first and the inferred and specific ones next (in a random order), does justice to the fact that the vague use is connected to the corporate one and has the inferred and specific ones as its separate off-shoots in Siewierska & Papastathi (2011). The plural contexts preceding the number-neutral ones within these uses captures the fact that, in Gast & van der Auwera's (2013) map, the plural "node" links up with the corporate one and the number-neutral one with the plural one. Finally, in Siewierska & Papastathi's (2011) view, the speech act verb use is a direct and otherwise unconnected off-shoot of the third person plural's personal use. Because personal 'they' is not included in the present study, the speech act verb use simply appears at the end. For the cross-linguistic differences motivating all these distinctions and their relations, we refer to Siewierska & Papastathi (2011: 603-605) and Gast & van der Auwera (2013: 143-151).

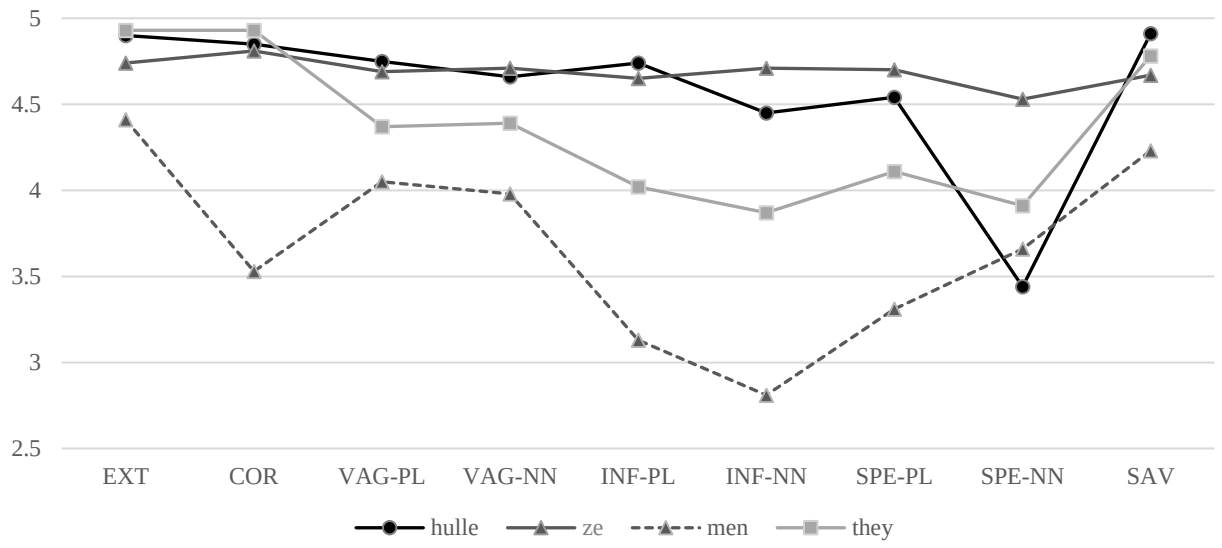


Figure 3: Distribution of the mean scores of 'they' and men for the non-universal-internal uses

Let us know look at results and, more specifically, first those of the third person plural in the three languages and then those of the 'man'-pronoun in Dutch.

According to a one-way ANOVA ($p > 0.05$), there is no substantial variation in the acceptability of *ze* across the non-universal-internal uses. Post hoc t-tests do reveal that Dutch 'they' scores lower in existential-specific-number-neutral contexts than in universal-external ones and existential-corporate ones ($p = 0.034$ and 0.004 respectively) but these differences do not meet a Bonferroni-corrected level of significance. Unsurprisingly, our two-way ANOVA testing the dimensions of (un)knownness and number in the purely impersonal existential uses does not show any effect or interaction either. These findings are quite interesting in light of the literature. They suggest, for instance, that Gast & van der Auwera's (2013: 149) claim about the necessarily plural meaning of *ze* is not true (see Section 1). Moreover, Siewierska & Papastathi (2011: 596-598) observe that, in their questionnaire, the acceptability of the Dutch third person plural follows the hierarchy in (20) (see Section 2.1 about the absence of the speech act verb use here).

(20) UNI-EXT ≈ EXI-COR > EXI-INF > EXI-VAG > EXI-SPE

The differences in our questionnaire between the universal-external and existential-corporate uses on the one hand and the existential-specific-number-neutral one on the other could, in a sense, be seen as being in line with their extreme positions in (20) but our overall lack of variation between, for example, the existential-corporate, -inferred and -vague contexts is definitely at odds with the hierarchy. These discrepancies may be due to the fact that our data comes from three times as many participants as Siewierska & Papastathi's (2011) (see Section 2.4) and, perhaps more likely, to the fact that our questionnaire tests the sheer functional potential of *ze* while the earlier one implicitly compares the acceptability of the HIP with that of non-pronominal forms of impersonalization (see Section 2.3). To see how Dutch 'they' fares compared to other strategies, we will look at the results of the completion task in Section 4.2.

Unlike in Dutch, the acceptability of the third person plural varies across the non-universal-internal uses in both English and Afrikaans, as the one-way ANOVAs make clear ($p < 0.05$ in each case). For English, the post hoc Bonferroni-corrected t-tests shows that 'they' scores significantly

higher in universal-external and existential-corporate contexts than in all others, with the exception of speech act verb contexts. This use in turn has substantially higher scores for *they* than all existential-inferred and -specific ones ($p < 0.001$ across the board). Furthermore, a closer look at the truly impersonal existential uses, by means of a two-way ANOVA, reveals that, separately, number and (un)knownness have no effect on the acceptability of the English third person plural ($p > 0.05$) but that their interaction is significant ($p < 0.05$). The reason for this result is that only in number-neutral contexts does *they* score higher in the vague use than in the inferred ($p = 0.037$) and specific ($p = 0.036$) ones. Together, our findings are more or less in keeping with Siewierska & Papastathi's (2011: 596-598) results in (21) for English.

(21) UNI-EXT ≈ EXI-COR > EXI-VAG > EXI-INF ≈ EXI-SPE

Like in (21), the universal-external and existential-corporate contexts outrank the other ones in our data, which indicates that there is indeed "an overall speaker preference for semi-impersonal uses" (Siewierska & Papastathi 2011: 599) of the third person plural and *they* in particular. In addition, the relation between the existential-vague, -inferred and -specific uses with a number-neutral reading in our questionnaire is identical to their relative positions in the hierarchy. However, the actual scores suggest that it is an exaggeration to say that the existential-specific use "is at best marginal in … English" and that the existential-inferred one "appears to be absent in English" (Siewierska & Papastathi 2011: 598): their respective plural and number-neutral ratings are 4.11 and 3.91 and 4.02 and 3.87. These claims may, of course, be more true of usage, which is discussed in Section 4.2. Another important difference with (21) is that, according to our data, the dimensions of number and (un)knownness, from the two existing semantic maps, interact. This fact hints at the possibility of some type of combined map (see Section 5).

For Afrikaans, the post hoc t-tests show that *hulle* is significantly less acceptable in existential-specific-number-neutral contexts than in all others and that it scores substantially higher in the universal-external, existential-corporate and speech act verb uses than in the existential-inferred-number-neutral one. The universal-external and speech act verb uses are also found to have higher ratings for the third person plural than the existential-specific-plural one ($p < 0.001$ in all cases). In brief, *hulle*'s acceptability, not unlike *they*'s, appears to be the most solid in the semi-impersonal and speech act verb contexts and somewhat weaker in the existential-inferred-number-neutral and existential-specific ones (note, however, that its scores in these uses, ranging from 3.44 to 4.54, are still fairly high). Furthermore, a two-way ANOVA testing (un)knownness and number in the truly impersonal existential uses confirms the need for a combined semantic map: both dimensions have a significant effect on *hulle*'s acceptability and so does their interaction ($p < 0.05$ across the board). Our t-test-based comparison of the various contexts reveals the following (potentially) meaningful differences: Afrikaans 'they' scores higher with a plural reading than a number-neutral one in the existential-inferred ($p = 0.009$) and -specific ($p < 0.001$) uses but not in the existential-vague use ($p = 0.34$) and it is more acceptable in existential-vague contexts than in existential-specific ones, irrespective of its number-neutral ($p < 0.001$) or plural ($p = 0.04$) interpretation. In number-neutral contexts, the existential-inferred use too gets a higher rating for *hulle* than the existential-specific one ($p < 0.001$).

What is important to point out here is that, although *they* and *hulle* do not behave in exactly the same way in the truly impersonal existential domain, the directionality of the differences in the third person plural's acceptability abides by the "logic" of the existing semantic maps in both languages. In no case does 'they' score higher in the existential-inferred and -specific uses than in the

existential-vague one, of which the former are separate off-shoots according to Siewierska & Papastathi (2011). In the same vein, there are no instances in which *they* or *hulle* is significantly more acceptable in an existential-number-neutral context than in an existential-plural one, of which, for third person plural HIPs, the former is an extension in Gast & van der Auwera's (2013) map.

In the preceding paragraphs, the focus is on the language-internal variation in acceptability of *hulle*, *ze* and *they*. The (dis)similarities between the third person plural HIPs in Afrikaans, Dutch and English deserve our attention too. For universal-external, existential-corporate and speech act verb contexts, as well as for both existential-vague ones, no significant differences exist between *hulle*, *ze* and *they* (the only peculiar exception is the higher acceptability of *they* compared to *ze* in universal-external contexts ($p < 0.017$)). The fact that these uses also tend to be those with the highest levels of acceptability suggests that they are the most well-established uses of 'they' in the three languages and, perhaps, in general. Siewierska & Papastathi (2011: 592), for instance, do not find any attestations of third person plural HIPs in existential-inferred or -specific contexts in their parallel corpus study of nine languages. We do not mean to claim, however, that 'they' has all these uses in every language. German *sie* 'they', for example, is known to dislike the speech act verb context (see Van Olmen & Breed 2018: 4). As for differences between Afrikaans, Dutch and English, *ze* and *hulle* are found to be significantly more acceptable than *they* in all existential-inferred and -specific uses – apart from the existential-specific-number-neutral one, where *they* has a higher acceptability rating than *hulle*. In addition, the Dutch third person plural scores higher than its Afrikaans counterpart but only in existential-inferred and -specific contexts with a number-neutral reading ($p < 0.017$ across the board).

These differences between *hulle*, *ze* and *they*, as well as their language-internal variation, can be explained in terms of semantic bleaching. To be more precise, the divergence between *hulle* and *ze* on the one hand and *they* on the other can be attributed to attrition in what Siewierksa & Papastathi (2011: 605) describe as "referent identification". As a personal pronoun, the third person plural explicitly points to a definite set of referents. In the universal-external, existential-corporate and existential-vague uses, where no variation between the languages is found, there is a decline in the degree of "overt referent identification as expressed within the construction" (Siewierska & Papastathi 2011: 584): from the locative phrase in universal-external contexts through the predicate itself in existential-corporate ones to the near lack of any cue for identification in existential-vague ones (see Section 2.1). What distinguishes these uses from the existential-inferred and -specific ones, in which *hulle* and *ze* do outrank *they*, is the latter's "purely situational as opposed to verbal nature of the referent identification" (Siewierska & Papastathi 2011: 584): a piece of circumstantial evidence from which a set of referents responsible for it is deduced in existential-inferred contexts and "some form of contact of the speaker with the referent of the subject, for example, visual (at the door, on TV), audio (on the phone, radio), written (via an e-mail or letter)" (Siewierska & Papastathi 2011: 584) at the moment of utterance in existential-specific contexts. In both uses, the third person plural is even more disconnected from its original definite meaning than in the other impersonal uses, as argued by Siewierska & Papastathi (2011: 600). In short, the variation in our data for these contexts thus appears to indicate that third person plural's referent identification features as a personal pronoun are somewhat less bleached in *they* than in *hulle* and *ze*.

An additional factor accounting for the divergence between the Dutch third person plural and its Afrikaans equivalent is attrition in number. The personal pronoun 'they' obviously refers to two or more referents. In the universal-external, existential-corporate and speech act verb uses, where *ze* and *hulle* have similar levels of acceptability, this plural meaning is still present (see Siewierska & Papastathi 2011: 583, 585). In truly impersonal existential contexts, the interpretation of 'they' may be plural or number-neutral, in Gast & van der Auwera's (2013: 141) view. In the first case,

the original sense of plurality of the pronoun can be said to persist. In the second case, it has clearly eroded to give way to a possible reading of 'they' as 'someone' instead of the compulsory reading as 'some people'. In other words, the (dis)similarities in our data between *ze* and *hulle* in the truly impersonal existential uses seems to suggest that, especially in contexts with more situation-based referent identification, the third person plural's original number feature is slightly less bleached in *hulle* than in *ze*.

The final HIP that needs to be discussed here is, of course, *men*. A first thing to point out is that this 'man'-pronoun's relation to the third person plural in the non-universal-internal domain resembles its relation to the second person singular in the universal-internal domain. More specifically, *men* is considered less acceptable than its competitor in all uses ($p < 0.001$ across the board), which is probably again due to its formality and possible ongoing disappearance (see Section 3.1). The 'man'-pronoun does exhibit some intriguing variation in acceptability between contexts, as a one-way ANOVA ($p < 0.05$) and our post hoc Bonferroni-corrected t-tests suggest. It scores lower in all its universal-internal uses than in the universal-external, existential-vague and speech act verb ones. Perhaps, the latter contexts should be regarded as *men*'s last strongholds. Its relatively high acceptability, i.e. 4.23, in cases like (22) (cf. 10) is of particular interest. As mentioned in Section 2.1, the speech act verb use only appears in Siewierska & Papastathi's (2011) map for 'they', where it is a direct and otherwise unconnected off-shoot of the third person plural's personal use (because, in a language like French, 'they' can be employed impersonally in a variety of ways but not with a speech act verb while, in a language like Finnish, it does not really serve as a HIP but can be used in speech act verb contexts). The occurrence of *men* in (22) can obviously not be linked to the third person plural's personal use. This fact suggests that the speech act verb use *is*, in some way, related to the other existential uses. Gast & van der Auwera (2013) do not immediately see how, though, and exclude the use from their semantic map for all HIPs. Let it suffice to say here that we fully endorse their recommendation for "more (esp. diachronic) investigation" (Gast & van der Auwera 2013: 142) into speech act verb contexts.

(22) D ***Men*** *zegt dat dat huis wemelt van de spoken.*
'They say that that house is crawling with ghosts.'

For the genuinely impersonal existential uses, a two-way ANOVA shows that the acceptability of *men* is also affected by the interaction of the dimensions of (un)knownness and number ($p < 0.05$). A more detailed comparison reveals meaningful differences between the plural and number-neutral existential-vague contexts on the one hand and their inferred ($p < 0.001$ in both cases) and specific ($p < 0.001$ and $p = 0.023$ respectively) equivalents on the other. The fact that *men* scores higher in the former is in line with what we know of 'they'. It is not entirely clear to us, however, why the weakening referent identification relevant for the third person plural would have an impact on 'man'-pronouns. Perhaps, *men*'s lower ratings in existential-inferred and -specific uses are just due to a mixture of the participants' general uncertainty about this rare HIP and the overall infrequency of these particular contexts (as noted by Siewierska & Papastathi 2011: 592 for 'they'). At any rate, for the dimension of number, Gast & van der Auwera (2013: 149) seem reasonable in not assuming any differences in *men* and 'man'-pronouns at large. In our data, there is indeed no number-based variation in existential-vague and -specific contexts. The existential-specific use defies any expectation, though: *men* is actually more acceptable in its number-neutral version than in its plural version ($p = 0.023$). Clearly, further research, possibly involving other fully grammaticalized 'man'-pronouns like German *man* or French *on*, is needed in the future.

## 4.2 Completion task

Table 5 presents, for each language and each non-universal-internal use, the absolute numbers and the proportions of relevant answers, i.e. impersonalization strategies, and irrelevant ones.

Table 5: Impersonalization strategies versus other answers for the non-universal-internal uses in the completion task

| Use | | Afrikaans | | Dutch | | English | |
|---|---|---|---|---|---|---|---|
| | | Relevant | irrelevant | relevant | irrelevant | relevant | irrelevant |
| UNI-EXT | # | 144 | 6 | 78 | 0 | 55 | 1 |
| | *%* | *96.00* | *4.00* | *100.00* | *0.00* | *98.21* | *1.79* |
| EXI-COR | # | 122 | 28 | 76 | 2 | 54 | 2 |
| | *%* | *81.33* | *18.67* | *97.44* | *2.56* | *96.43* | *3.57* |
| EXI-VAG-PL | # | 135 | 15 | 77 | 1 | 54 | 2 |
| | *%* | *90.00* | *10.00* | *98.72* | *1.28* | *96.43* | *3.57* |
| EXI-VAG-NN | # | 115 | 35 | 78 | 0 | 54 | 2 |
| | *%* | *76.67* | *23.33* | *100.00* | *0.00* | *96.43* | *3.57* |
| EXI-INF-PL | # | 127 | 23 | 72 | 6 | 54 | 2 |
| | *%* | *84.67* | *15.33* | *92.31* | *7.69* | *96.43* | *3.57* |
| EXI-INF-NN | # | 132 | 18 | 74 | 4 | 56 | 0 |
| | *%* | *88.00* | *12.00* | *94.87* | *5.13* | *100.00* | *0.00* |
| EXI-SPE-PL | # | 139 | 11 | 78 | 0 | 56 | 0 |
| | *%* | *92.67* | *7.33* | *100.00* | *0.00* | *100.00* | *0.00* |
| EXI-SPE-NN | # | 139 | 11 | 78 | 0 | 56 | 0 |
| | *%* | *92.67* | *7.33* | *100.00* | *0.00* | *100.00* | *0.00* |
| SAV | # | 129 | 21 | 78 | 0 | 53 | 3 |
| | *%* | *86.00* | *14.00* | *100.00* | *0.00* | *94.64* | *5.36* |

An in-depth discussion of the variation between languages and/or between uses, though potentially not uninteresting, is beyond the scope of this article (but see Van Olmen & Breed 2018: 13-15 on Afrikaans). Let the examples of irrelevant answers in (23a) to (23c) (cf. 10, 11a and 9b respectively) and one comment suffice here.

(23) a. A ***Kenners*** *sê dat daardie huis wemel van die spoke.*
'Experts say that that house is crawling with ghosts.'
b. D ***De jongeren*** *zijn hier bij elkaar gekomen voor een feestje.*
'The young people have gathered here for a party.'
c. E ***The police*** *have installed speed cameras here.*

In contexts like (23c), overt references to the police, the government or other such institutions are found in all three languages. This fact is not so surprising, of course, given that the answers testify to the corporate nature of the particular existential use.

Within the relevant answers, we again distinguish a set of miscellaneous impersonalization strategies from a number of major ones. The only miscellaneous form that occurs more than once is the passive in both Afrikaans and Dutch, as in (24). What is intriguing about this strategy, which does not fit the HIP-favoring blank (see Section 2.2), is that participants must have felt so strongly about its use that they changed the entire sentence. It definitely deserves to be studied in more detail as a form of impersonalization in follow-up research.

(24) a. A ***Daar word gesê*** *dat daardie huis wemel van die spoke.*
'It is said that that house is crawling with ghosts.'

b. D *Jij en je vrienden trekken rond middernacht naar het strand om te gaan zwemmen. Wanneer jullie er aankomen, zien jullie in het zand een tafeltje met twee stoelen, een bijna opgebrande kaars, een boeket rozen en wat etensrestjes. Je zegt tegen je vrienden: "Hier* ***is*** *net een romantisch etentje* ***gehouden****.'*
'You and your friends are going to the beach for a midnight swim. When you get there, you see a table with two chairs, a dying candle, a bouquet of roses and some leftover food. You tell your friends: 'A romantic dinner has just been had here.'

The major impersonalization strategies are the definite plural noun phrase 'the people', the indefinite plural noun phrase 'people' and indefinite pronouns. Although the first one *is* possible in Dutch and English, it is found here only in Afrikaans, where it is typical of the universal-external use, as in (25a) (cf. 9a). The other two strategies appear in all three languages and, to a large extent, they are complementary: with just a few exceptions, 'people' is restricted to universal-external, speech act verb and existential-plural contexts while the indefinite pronouns, most of which are cases of 'someone', occur most frequently in existential-number-neutral ones, as in (25b) and (25c) respectively (cf. 12c and 12d). Their distribution is, in other words, determined by number.

(25) a A *In Griekeland ry* ***die mense*** *nogal onvoorspelbaar.*
'In Greece, the people drive quite unpredictably.'
b. D ***Mensen*** *bellen je op je twee telefoons.*
'People are calling you on both your phones.'
c. E ***Someone*** *is knocking on the door.*

The only intrinsically plural context that hardly ever features an indefinite plural noun phrase is the existential-corporate one. This "anomaly" is due to the fact that the predicate in such a use limits the reference to a more or less definite institutional subset of human participants, with which 'people' is incompatible. Another deviation from the number-based distribution of indefinite pronouns and indefinite plural noun phrases is the appearance of 'someone' in truly impersonal existential-plural contexts. It is always considerably more frequent in the number-neutral equivalents but answers like (26a) and (26c) (cf. 11a and 12a) are not uncommon.

(26) a. A ***Iemand*** *het hier bymekaargekom vir 'n partytjie*.
'Someone has gathered here for a party.'
b. E *Oh, after he was released from prison,* ***someone*** *lynched him.*

We have no real explanation for these results at this point. Maybe, some participants did not feel entirely comfortable with 'they' or *men* in contexts like (26) and struggled to fill out the blank and comply with the description in the instructions at the same time. 'Someone', arguably one of the most straightforward forms of existential quantification in all three languages, may have been their last resort, forcing them to construe cases such as (26) as number-neutral in some way.

Table 6 provides the absolute numbers and the proportions of the various forms of impersonalization in the three languages under examination and in the nine non-universal-internal uses. Note that the second person singular is included here because its use in universal-external contexts, interpreted as universal-internal ones (see Section 4.1), counts as a HIP. The same is true of Afrikaans '(a) man' in this use. The data is presented in a visually more attractive way in Figure 4.

Table 6: Distribution of the impersonalization strategies for the non-universal-internal uses in the completion task

| Language | Use | | you | (a) man | they | idf.pro | people | the people | misc. |
|---|---|---|---|---|---|---|---|---|---|
| Afrikaans | UNI-EXT | # | 3 | 13 | 53 | 3 | 56 | 16 | 0 |
| | | *%* | *2.08* | *9.03* | *36.81* | *2.08* | *38.89* | *11.11* | *0.00* |
| | EXI-COR | # | 0 | 0 | 104 | 14 | 3 | 1 | 0 |
| | | *%* | *0.00* | *0.00* | *85.25* | *11.48* | *2.46* | *0.82* | *0.00* |
| | EXI-VAG-PL | # | 0 | 0 | 62 | 10 | 59 | 3 | 1 |
| | | *%* | *0.00* | *0.00* | *45.93* | *7.41* | *43.70* | *2.22* | *0.74* |
| | EXI-VAG-NN | # | 0 | 1 | 28 | 82 | 1 | 0 | 3 |
| | | *%* | *0.00* | *0.87* | *24.35* | *71.30* | *0.87* | *0.00* | *2.61* |
| | EXI-INF-PL | # | 0 | 0 | 13 | 87 | 25 | 1 | 1 |
| | | *%* | *0.00* | *0.00* | *10.24* | *68.50* | *19.69* | *0.79* | *0.79* |
| | EXI-INF-NN | # | 0 | 0 | 8 | 122 | 1 | 0 | 1 |
| | | *%* | *0.00* | *0.00* | *6.06* | *92.42* | *0.76* | *0.00* | *0.76* |
| | EXI-SPE-PL | # | 0 | 0 | 30 | 45 | 64 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *21.58* | *32.37* | *46.04* | *0.00* | *0.00* |
| | EXI-SPE-NN | # | 0 | 0 | 1 | 134 | 2 | 0 | 2 |
| | | *%* | *0.00* | *0.00* | *0.72* | *96.40* | *1.44* | *0.00* | *1.44* |
| | SAV | # | 0 | 0 | 74 | 5 | 47 | 2 | 1 |
| | | *%* | *0.00* | *0.00* | *57.63* | *3.88* | *36.43* | *1.55* | *0.78* |
| Dutch | UNI-EXT | # | 0 | 8 | 68 | 1 | 1 | 0 | 0 |
| | | *%* | *0.00* | *10.26* | *87.18* | *1.28* | *1.28* | *0.00* | *0.00* |
| | EXI-COR | # | 0 | 1 | 75 | 0 | 0 | 0 | 0 |
| | | *%* | *0.00* | *1.32* | *98.68* | *0.00* | *0.00* | *0.00* | *0.00* |
| | EXI-VAG-PL | # | 0 | 6 | 65 | 0 | 6 | 0 | 0 |
| | | *%* | *0.00* | *7.79* | *84.42* | *0.00* | *7.79* | *0.00* | *0.00* |
| | EXI-VAG-NN | # | 0 | 5 | 47 | 26 | 0 | 0 | 0 |
| | | *%* | *0.00* | *6.41* | *60.26* | *33.33* | *0.00* | *0.00* | *0.00* |
| | EXI-INF-PL | # | 0 | 7 | 35 | 25 | 3 | 0 | 2 |
| | | *%* | *0.00* | *9.72* | *48.61* | *34.72* | *4.17* | *0.00* | *2.78* |
| | EXI-INF-NN | # | 0 | 3 | 28 | 43 | 0 | 0 | 0 |
| | | *%* | *0.00* | *4.05* | *37.84* | *58.11* | *0.00* | *0.00* | *0.00* |
| | EXI-SPE-PL | # | 0 | 4 | 61 | 10 | 3 | 0 | 0 |
| | | *%* | *0.00* | *5.13* | *78.21* | *12.82* | *3.85* | *0.00* | *0.00* |
| | EXI-SPE-NN | # | 0 | 1 | 30 | 47 | 0 | 0 | 0 |
| | | *%* | *0.00* | *1.28* | *38.46* | *60.26* | *0.00* | *0.00* | *0.00* |
| | SAV | # | 0 | 18 | 59 | 0 | 0 | 0 | 1 |
| | | *%* | *0.00* | *23.08* | *75.64* | *0.00* | *0.00* | *0.00* | *1.28* |
| English | UNI-EXT | # | 0 | 0 | 37 | 0 | 18 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *67.27* | *0.00* | *32.73* | *0.00* | *0.00* |
| | EXI-COR | # | 0 | 0 | 53 | 1 | 0 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *98.15* | *1.85* | *0.00* | *0.00* | *0.00* |
| | EXI-VAG-PL | # | 0 | 0 | 28 | 5 | 21 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *51.85* | *9.26* | *38.89* | *0.00* | *0.00* |
| | EXI-VAG-NN | # | 0 | 0 | 8 | 46 | 0 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *14.81* | *85.19* | *0.00* | *0.00* | *0.00* |
| | EXI-INF-PL | # | 0 | 0 | 5 | 40 | 9 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *9.26* | *74.07* | *16.67* | *0.00* | *0.00* |
| | EXI-INF-NN | # | 0 | 0 | 1 | 52 | 2 | 0 | 1 |
| | | *%* | *0.00* | *0.00* | *1.79* | *92.86* | *3.57* | *0.00* | *1.79* |
| | EXI-SPE-PL | # | 0 | 0 | 12 | 25 | 19 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *21.43* | *44.64* | *33.93* | *0.00* | *0.00* |
| | EXI-SPE-NN | # | 0 | 0 | 1 | 55 | 0 | 0 | 0 |
| | | *%* | *0.00* | *0.00* | *1.79* | *98.21* | *0.00* | *0.00* | *0.00* |
| | SAV | # | 0 | 0 | 32 | 3 | 18 | 0 | 0 |
| | | % | *0.00* | *0.00* | 60.38 | 5.66 | 33.96 | *0.00* | *0.00* |

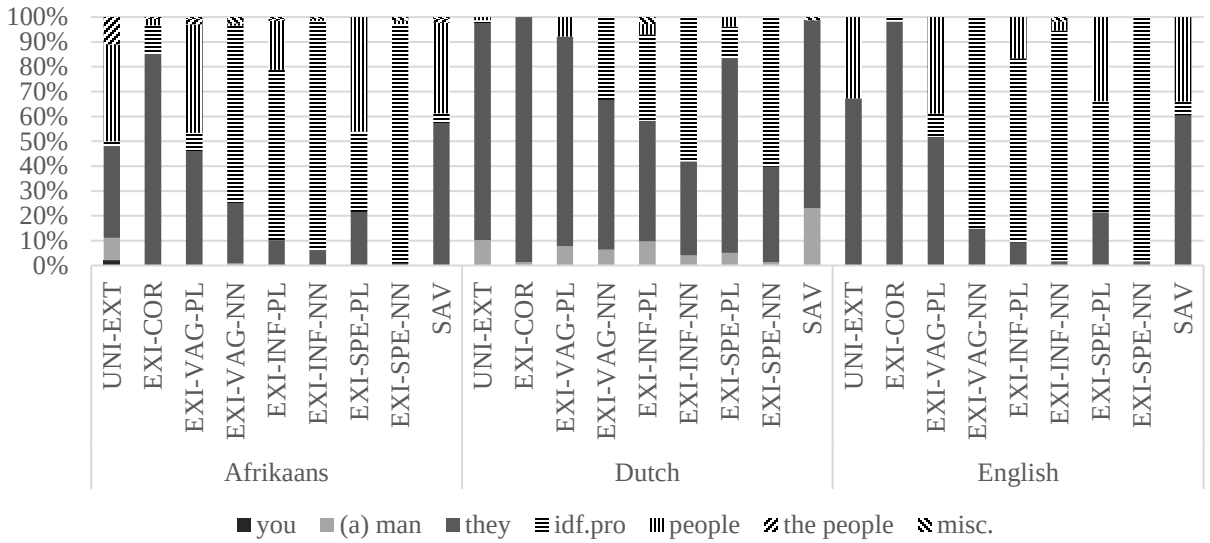


Figure 4: Distribution of the impersonalization strategies for the non-universal-internal uses in the completion task

In the rest of the present section, we first discuss the preference for HIPs in non-universal-internal contexts, then turn to the behavior of 'they' (and *men*) in each language and, finally, compare the use of the third person plural across the three languages.

It is fairly clear from Figures 2 and 4 that non-universal-internal contexts differ from universal-internal ones in the use of HIPs. In Afrikaans, *jy* and (*'n*) *mens* account for roughly 95% of the universal-internal answers. The two non-universal-internal uses that come closest are the existential-corporate and speech act verb ones, with *hulle* making up 85.25% and just 57.63% respectively. In no other context is this HIP responsible for more than half of the cases. The Bonferroni-corrected Fischer's exact tests confirm that each universal-internal use has a significantly higher proportion of HIPs than each non-universal-internal one, with the exception of existential-corporate contexts ($p < 0.001$ across the board). In Dutch too, HIPs are employed in, on average, 95% of the universal-internal cases. The differences with the universal-external, existential-corporate, existential-vague-plural, existential-specific-plural and speech act verb uses are not so pronounced: *ze* and *men* make up 97.44%, 100%, 92.21%, 83.33% and 98.72 respectively. The universal-internal contexts only have a significantly larger number of HIPs than the existential-inferred-plural one and all existential-number-neutral ones ($p < 0.001$ in each case). English, lastly, is more like Afrikaans. All universal-internal uses, with an average HIP percentage of 92%, have a significantly higher proportion of pronominal impersonalization strategies than all non-universal-internal ones ($p < 0.001$ across the board), except for the universal-external and existential-corporate ones. In short, generalizing over the three languages, we can say that there is indeed a stronger preference of HIPs in universal-internal contexts than in non-universal-internal ones, though especially the existential-corporate, universal-internal and speech act verb uses appear to be quite keen on pronominal forms of impersonalization too. These results suggest that at least part of the explanation for the infrequency with which HIPs are attested with an existential meaning in corpora (see Section 1) lies in the fact that they are simply not the preferred strategy for existential uses and in particular those of the genuinely impersonal type.

A more in-depth look at *hulle* shows that its use, like its acceptability (see Section 4.1), is the most solid in the semi-impersonal and speech act verb contexts: the HIP is significantly more common in the existential-corporate use than in the speech act verb and universal-external ones (as well as the existential-vague-plural ones), where it is substantially more frequent than in the remaining uses ($p < 0.001$ in all cases). In fact, in these truly impersonal existential contexts, the use of *hulle* is fairly restricted, ranging from 0.72% to 25.22%, which is slightly at odds with its relatively high levels of acceptability here. This limited variation is still interesting, though. A comparison of the contexts reveals that the third person plural is employed more in existential-vague-plural (45.93%) and existential-specific-plural (21.58%) uses than in their number-neutral equivalents (24.35% and 0.72% respectively) ($p < 0.001$ in both cases). No difference exists between the two existential-inferred contexts (10.24% and 6.06%). *Hulle* also occurs significantly more often in both existential-vague uses than in their existential-inferred and -specific counterparts, which do not vary from one another in a substantial way. These differences in usage may not be identical to those in acceptability (see Section 4.1) but, like the latter, they do follow the logic of Siewierksa & Papastathi's (2011) and Gast & van der Auwera's (2013) semantic maps: if there is number-based variation, it will be the plural contexts that exhibit larger numbers of third person plural HIPs and, in the same vein, if (un)knownness-based variation exists, it will be the vague uses that have a higher frequency of third person plural HIPs.

As discussed in the previous section, *ze* exhibits no real variation in acceptability. As far as usage is concerned, the differences between the existential-corporate (98.68%), universal-external (87.18%), existential-vague-plural (84.42%), existential-specific-plural (78.21%) and speech act verb uses are also quite limited, with only the last two having substantially fewer cases of *ze* than the first one ($p < 0.001$ in both cases). Yet, the third person plural does tend to occur significantly more often in these contexts, the semi-impersonal ones in particular, than in the remaining ones, in which its usage can drop to less than two fifths of the answers. We simply do not see such dissimilarities in the acceptability judgment task. Moreover, in the existential-vague and -specific uses, *ze*'s use drops from 84.42% and 78.21% with a plural reading to 60.26% and 38.46% with a number-neutral one ($p = 0.002$ and $p < 0.001$ respectively) and, in existential-plural uses, it decreases from 84.42% with a vague interpretation to 48.61 with an inferred one ($p < 0.001$) and, in existential-number-neutral contexts, from 60.26% with a vague reading to 37.84% with an inferred one and 38.46% with a specific one ($p = 0.006$ in each case). In other words, the frequency with which the third person plural is employed, unlike its acceptability, does seem to be affected by the dimensions of (un)knownness and number in expected ways.

For the other non-universal-internal HIP in Dutch, i.e. *men*, the number of occurrences in the completion questionnaire is too low to see any of the effects of (un)knownness and number found in the acceptability judgment data. In fact, more generally, the only use that occasionally displays a significantly larger proportion of instances of the 'man'-pronoun than another one (e.g. the universal-internal-non-veridical-non-modal and existential-corporate contexts with 1.28% and 1.32% respectively) is the speech act verb use (23.08%) ($p < 0.001$ in each case). Unlike in *men*'s acceptability, there is little evidence in its usage for some type of special status of the universal-external and existential-vague contexts. Perhaps, the speech act verb use is really the only stronghold of the Dutch 'man'-pronoun. Here, it is just three times less frequent than *ze* while, in the other contexts, it occurs at least five times less often. These results are, of course, in line with *men*'s consistently lower acceptability compared to the third person plural.

The picture for the English third person plural looks similar to that for its Afrikaans counterpart: the HIP is substantially more frequent in existential-corporate contexts (98.15%) than in universal-external (67.27%), speech act verb (60.38%) and existential-vague-plural (51.85%) ones, in

which it occurs significantly more often than in all of the remaining contexts ($p < 0.001$ across the board). These findings nicely match the results of the acceptability judgment task, where the universal-external, existential-corporate and speech act verb uses stand out (cf. 21 as well). In the truly impersonal existential uses, we can observe a decrease from 51.85% in vague-plural ones to 9.26% in inferred-plural ones and 21.43% in specific-plural ones ($p < 0.001$ in both cases) and a drop from 14.81% in vague-number-neutral ones to 1.79% in inferred- and specific-number neutral ones ($p = 0.0153$ in each case). In existential-vague and -specific contexts, the use of *they* is also significantly higher with a plural interpretation than with a number-neutral one ($p < 0.001$ in both cases). These facts suggest that, in English too, the dimensions of (un)knownness and number have an impact on the usage of the third person plural. What is more, unlike our acceptability judgment scores, the very low frequencies of *they* in the completion questionnaire for the genuinely impersonal existential contexts of the non-vague type support Siewierska & Papastathi's (2011: 598) claim that the existential-specific use "is at best marginal in … English" and the existential-inferred one "appears to be absent".

Lastly, let us compare the usage of 'they' in the three languages. Note that there are no differences in acceptability between Afrikaans, English and Dutch in the universal-external, existential-corporate, speech act verb and existential-vague uses and that *hulle* and *ze* are, on the whole, more acceptable than *they* in the other contexts and *ze* is more acceptable than *hulle* with a number-neutral interpretation in these contexts (see Section 4.1). The completion task suggests that, despite being the only third person plural with a competitor, *ze* is the most established HIP of the three in the non-universal-internal domain more generally: it occurs significantly more often than *hulle* and *they* in every context ($p < 0.017$), apart from the existential-corporate and speech act verb ones, where *they* is used to more or less the same extent. There is, however, no evidence that *hulle* is, in some way, a more established HIP in Afrikaans than *they* is in English: overall, no differences are found in the frequency with which they are employed in non-universal-internal contexts. In fact, in the uses where they do differ substantially, i.e. the universal-external and existential-corporate ones, it is *they* that occurs more often. In short, usage-wise, Dutch could be characterized as a more 'they'-prominent language than Afrikaans and English and English in turn as more 'they'-prominent than Afrikaans at least in the semi-impersonal domain.

### 4.3 Interim conclusion

It is clear from both questionnaires that the second person singular, (*'n*) *mens* and *one* cannot function as non-universal-internal HIPs. *Men*, by contrast, can be used in this domain, which indicates that the Dutch 'man'-pronoun is more grammaticalized than its Afrikaans equivalent not only formally but also functionally. The two questionnaires show, though, that, like in universal-internal contexts, it is the less acceptable and less frequent HIP. The explanation for these findings probably lies in its relatively formal and potentially disappearing character. Perhaps the only use where *men* remains somewhat strong according to the acceptability judgment and completion data together is the speech act verb one. This fact highlights the so far unsolved issue of how the context in question is connected to the other non-universal-internal contexts. For third person plural HIPs, it has been linked directly to their personal use but this account obviously does not work for 'man'-pronouns. More research into the matter is required.

Beside *men*, the only other HIPs able to serve non-universal-internal purposes are instantiations of 'they'. One of the things that Afrikaans, English and Dutch share is that – admittedly, to varying degrees – the third person plural is most acceptable and most frequent in universal-external, existential-corporate and speech act verb contexts (and, possibly, also existential-vague ones). This

result is in line with its previously observed preference for semi-impersonal uses, which can be attributed to the fact that they are closest to the personal use of 'they' with regard to overt referent identification and plurality. Another point that *hulle*, *ze* and *they* have in common, though again with considerable variation between them, is their decreasing acceptability and especially their decreasing usage in the truly impersonal existential domain along the dimensions of (un)known-ness and number: if a difference exists between existential-vague and existential-inferred and -specific uses or between existential-plural and existential-number-neutral ones, it is always the existential-inferred, -specific or -number-neutral one where 'they' has a lower level of acceptability or is employed less often. These drops can be explained in terms of referent identification and number as well: the third person plural's original sense of definite reference and plurality needs to be bleached substantially for it to be able to occur in contexts where reference is more situation-based or potentially singular. Moreover, the interaction between (un)knownness and number in the decreases in all three languages suggests that a combined semantic map, as first proposed in our own earlier study of Afrikaans on its own (see Van Olmen & Breed 2018: 25) is really needed.

The comparison of *hulle*, *ze* and *they* reveals few differences in acceptability. Only in existential-inferred and -specific contexts are *hulle* and *ze* considered more acceptable than *they* and only in the number-neutral subtypes of these contexts is *ze* judged as more acceptable than *hulle*. This variation could be interpreted as pointing to diverging degrees of semantic attrition of the third person plural's original personal meaning. As far as usage is concerned, our completion task shows that, given the almost consistently higher frequency with which *ze* is employed in all uses compared to *hulle* and *ze*, Dutch is more 'they'-prominent than Afrikaans and English. The latter language is found to be more 'they'-prominent than Afrikaans in the two semi-impersonal contexts. To conclude, the completion questionnaire also points out that the universal-internal domain has a much stronger preference for pronominal forms of impersonalization than the non-universal-internal one, with the exception of the universal-external and existential-corporate uses. This result can be taken as an indication that the low frequencies with which HIPs are attested with an existential function in corpus studies are at least partially due to the fact that they are simply not the preferred strategy for existential contexts and especially the truly impersonal ones.

## 5 Conclusion

For a summary of our results, we refer to the interim conclusions in Sections 3.3 and 4.3. Our focus here is on the combination of Siewierska & Papastathi's (2011) semantic map and Gast & van der Auwera's (2013), first proposed in a slightly different form by Van Olmen & Breed (2018: 25)[19] to describe the Afrikaans data and further supported by the present findings for Dutch and English. The map is presented in Figure 5.

[19] More specifically, the universal-internal uses *are* included in Figure 5 and the speech act verb context is no longer linked to the universal-external one through the third person plural's personal use.

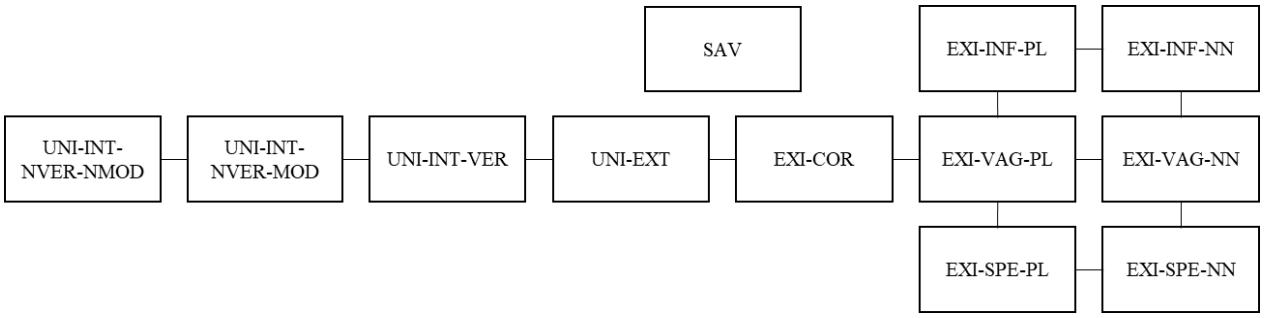


Figure 5: Combined semantic maps for all HIPs

The arrangement of the universal-internal, universal-external and existential-corporate contexts is based on Gast & van der Auwera (2013) (see also footnote 11). The order of the universal-external and existential-corporate uses is the same in Siewierska & Papastathi (2011). Their map includes a speech act verb use, unlike Gast & van der Auwera's (2013), and posits the third person plural's personal use, referred to as "known (anaphoric)", as its connecting context with the universal-external one. If Figure 5 *is* to capture the functional behavior of HIPs, the presence of the speech act verb use is essential. However, given the problems that its supposed link to the universal-external context via 'they' as a personal pronoun poses for *men* (see Section 4.1), we refrain from relating it to another context at this point. Our six truly impersonal existential uses combine the dimensions of (un)knownness and number from the two existing semantic maps. The link between the existential-corporate and existential-vague-plural contexts originates from the former's connection to the existential-vague use in Siewierska & Papastathi (2011) and to the existential-plural one in Gast & van der Auwera (2013). The arrangement of the remaining uses reflects the earlier maps' analyses of the existential-inferred and -specific ones as separate off-shoots of the existential-vague one and of the existential-number-neutral one as an extension of the existential-plural one (see also footnote 18).

The semantic map in Figure 5 works in the following way: (i) if a HIP is known to possess a particular use, it will also have all the uses to the left of it on the same line, up to the point at which the HIP entered the impersonal domain in the first place, (ii) if a HIP possesses a use on the top and/or bottom line, it will also have the use on the center line next to it (and therefore the uses to the left of the latter use too), (iii) these principles do not apply to the unconnected speech act verb use. Let us now illustrate the map with English 'you', 'one' and 'they' and with the Dutch 'man'-pronoun (it works for *jy*, *je*, *hulle*, *ze* and (*'n*) *mens* too, though).

In Figures 6 and 7, we present, respectively, *you*'s ratings from the acceptability judgment task and its frequencies of use from the completion task. Different shades of gray are used to represent (dis)similarities in acceptability and usage. The actual value that each tint of gray stands for is given inside the rectangles.

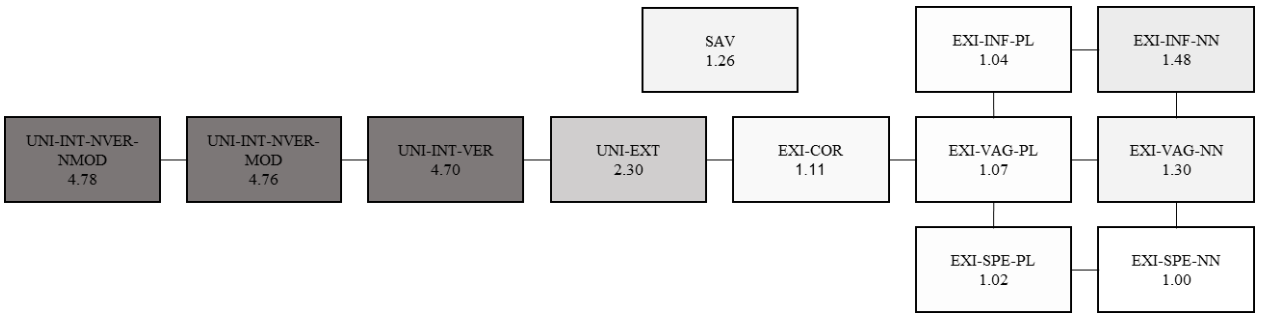

Figure 6: *You* in the acceptability judgment task on the combined semantic map

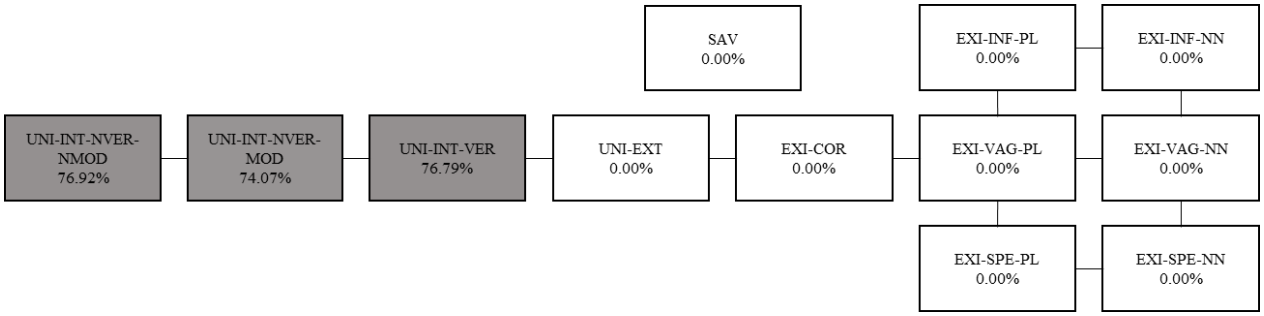


Figure 7: *You* in the completion task on the combined semantic map

These figures nicely show that *you* is a universal-internal HIP. The only slightly darker use in the non-universal-internal domain is the universal-external one in Figure 6, which has been argued to be due to participants imposing an internal perspective on a universal-external scenario when faced with the second person singular (see Section 4.1). Figures 6 and 7 can also be said to exemplify the organization of the map: *you* can be employed in the universal-internal-veridical use and therefore also in the two universal-non-veridical ones to the left of it. All of the above is, of course, true of *one* in Figures 8 and 9 as well.

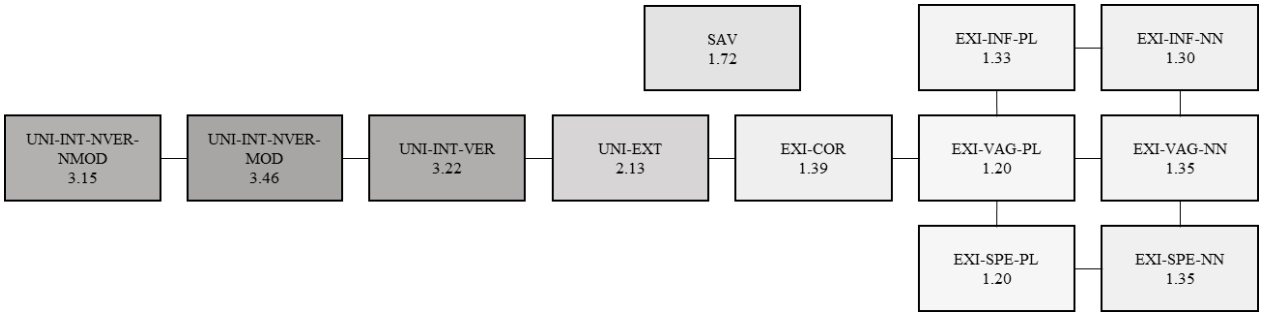


Figure 8: *One* in the acceptability judgment task on the combined semantic map

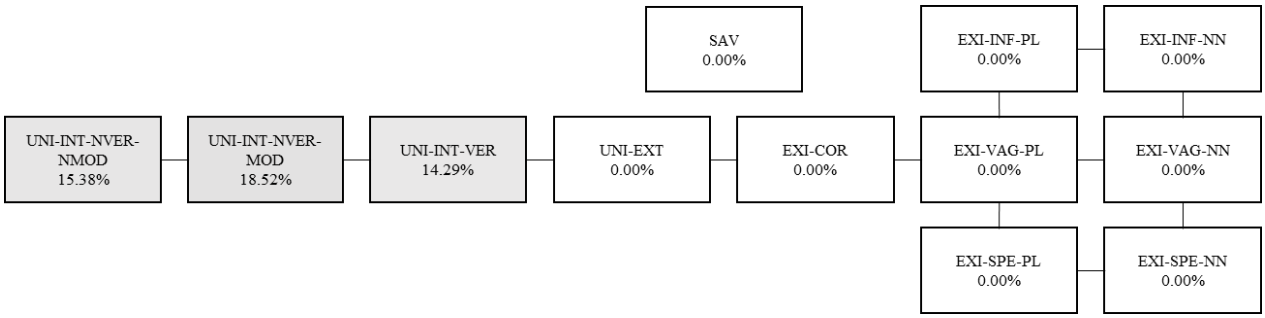


Figure 9: *One* in the completion task on the combined semantic map

The generally lighter shades of gray reflect the fact that this HIP is considered less acceptable and is employed less often than the second person singular.

Figures 10 and 11 capture *they*'s scores in the acceptability judgment questionnaire and its proportions of use in the completion questionnaire respectively.

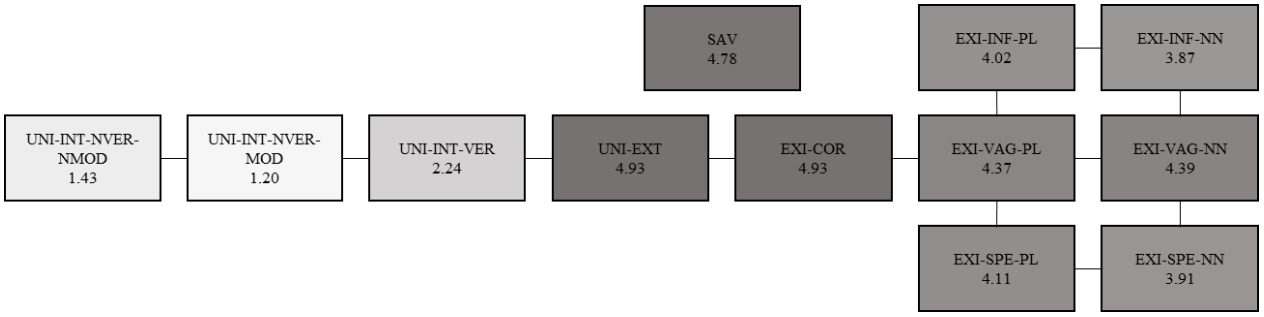


Figure 10: *They* in the acceptability judgment task on the combined semantic map

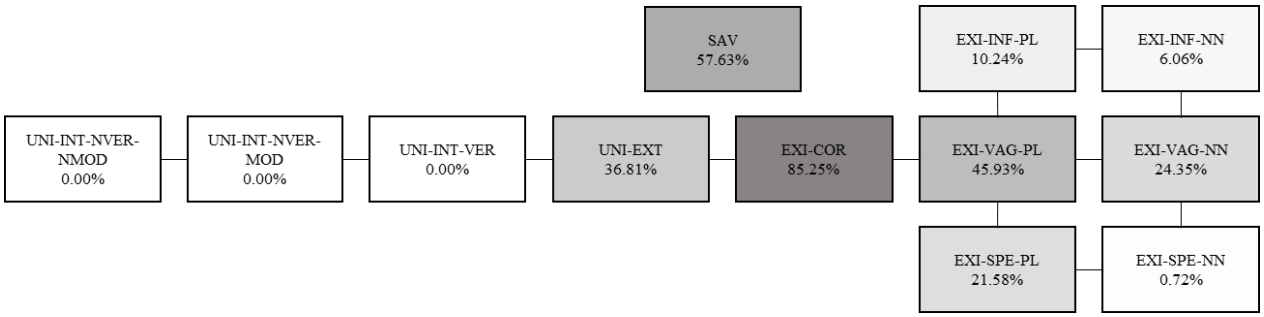


Figure 11: *They* in the completion task on the combined semantic map

The almost complete lack of color at the left side of the maps clearly demonstrates that *they* serves as a non-universal-internal HIP only. It is important to note that this absence of universal-internal uses, which are to the left of the universal-external one, does not actually violate the map's "logic". The reason, mentioned in (i) above, is that 'they' comes into the impersonal domain at the universal-external point. It is also evident from Figures 10 and 11 that the third person plural is most well-established in semi-impersonal, speech act verb and existential-vague-plural contexts in English. Furthermore, the right side of the maps, in particular that of *they*'s frequency of use, illustrates their organization in the genuinely impersonal existential domain – and, thus, the interaction of the dimensions of (un)knownness and number – very well. No use on the top or bottom line is significantly darker than its adjacent use on the center line and, similarly, no use at the extreme right is substantially darker than the use immediately to the left of it. In other words, the possible and only the possible distinctions predicted by the semantic map in Figure 5 are present in *they*, acceptability- and especially usage-wise. They are, however, more a matter of degree than of discrete cut-off points for the third person plural in English. This also holds for the other HIPs examined in this article, in spite of, for instance, Gast & van der Auwera's (2013) categorical claim about *ze*'s restriction to existential-plural uses. It remains to be seen in future research whether any HIP in any other language makes more discrete distinctions in the existential domain.

Mapping *men*, finally, points to what may seem like a potential problem for the arrangement of uses. The entirely gray nature of Figure 12 shows that the Dutch 'man'-pronoun can, in theory, serve any impersonal purpose but, according to Figure 13, there is, in reality, just one context that it is still more or less regularly employed for, i.e. the speech act verb one. This fact could make one think of circumstances in which *men* has specialized completely and only has the speech act verb use left. From a synchronic point of view, such a situation would go against the logic of the map. Under the assumption that the speech act verb context is somehow connected to the existential ones (see Section 4.1) and in accordance with (i) above, *men* would be expected to have at least all universal uses as well because 'man' always enters the impersonal domain as a universal-internal

HIP.

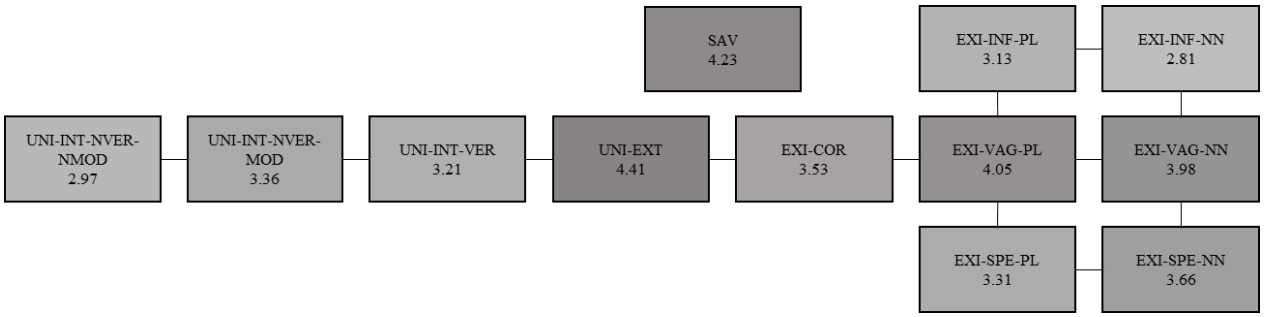


Figure 12: *Men* in the acceptability judgment task on the combined semantic map

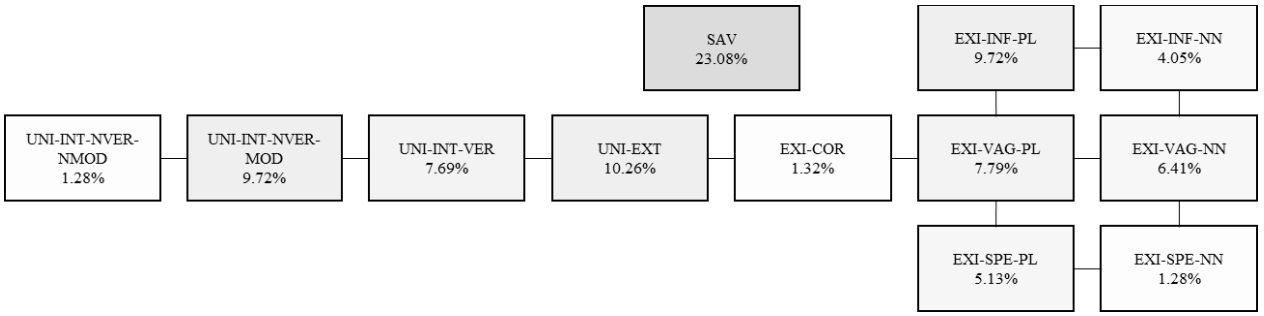


Figure 13: *Men* in the completion task on the combined semantic map

However, the semantic map would be saved by the diachronic perspective. The unfulfilled expectation about speech-act-verb-only *men* would immediately compel us to hypothesize that the HIP must have had the universal uses at some point and this hypothesis would be confirmed by the historical facts (cf. van der Auwera & Plungian 1998: 113 on deontic and concessive *mogen* 'may' in Present-day Dutch, the missing link between the two of epistemic possibility on modality's semantic map and the latter meaning's presence in an older stage of the language). Let us conclude, though, with the observation that, to our knowledge, the similar situation of a 'man'-pronoun which can be used existentially but not universally anymore simply does not exist (see Giacalone Ramat & Sansò 2007, Siewierska 2011).

References

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

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

Cabredo Hofherr, P. 2010. Binding properties of impersonal human pronouns in generic and episodic contexts. http://archive.sfl.cnrs.fr/sites/sfl/IMG/pdf/impersMay2010CabredoHofherrManOnLa.pdf (accessed on 6 April 2017)

Cheshire, J. Grammaticalisation in social context: The emergence of a new English pronoun. *Journal of Sociolinguistics* 17, 608-633.

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

De Hoop, H. & S. Tarenskeen. 2015. It's all about *you* in Dutch. *Journal of Pragmatics* 88, 163-175.
Deringer, L., V. Gast, F. Haas & O. Rudolf. 2015. Impersonal uses of the second person singular and generalized empathy: An exploratory corpus study of English, German and Russian. In L. Gardelle & S. Sorlin (eds.), *The Pragmatics of Personal Pronouns*. Amsterdam: John Benjamins, 311-334.
Donaldson, B.C. 1993. *A Grammar of Afrikaans*. Berlin: Mouton de Gruyter.
Duinhoven, A.M. 1990. Verdwijnt *men*? In J.B. den Besten, A.M. Duinhoven & J.P.A. Stroop (eds.), *Vragende Wijs: Vragen over Tekst, Taal en Taalgeschiedenis*. Amsterdam: Rodopi, 70-80.
Fonesca-Greber, B. & L.R. Waugh. 2003. On the radical difference between the subject personal pronouns in written and spoken European French. In P. Leistyna & 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, 4-33.
Gast, V. 2017. *Online Database for a Typology of Human Impersonal Pronouns*. http://www.personal.uni-jena.de/~mu65qev/improtype (accessed on 4 July 2017)
Gast, V., L. Deringer, F. Haas & 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. & J. van der Auwera. 2013. Towards a distributional typology of human impersonal pronouns, based on data from European languages. In D. Bakker & M. Haspelmath (eds.*), Languages Across Boundaries: Studies in the Memory of Anna Siewierska*. Berlin: Mouton de Gruyter, 119-158.
Giacalone Ramat, A. & A. Sansò. 2007. The spread and decline of indefinite man-constructions in European languages: An areal perspective. In P. Ramat & E. Roma (eds.), *Europe and the Mediterranean Linguistic Areas: Convergences from a Historical and Typological Perspective*. Amsterdam: John Benjamins, 95-131.
Greenbaum, S. & R. Quirk. 1970. *Elicitation Experiments in English: Linguistic Studies in Use and Attitude*. London: Longman.
Hoekstra, J. 2010. On the impersonal pronoun *men* in Modern West Frisian. *Journal of Comparative Germanic Linguistics* 13, 31-59.
Hofmeister, P., T.F. Jaeger, I. Sag, I. Arnon & N. Snider. 2007. Locality and accessibility in wh-questions. In S. Featherston & W. Sternefeld (eds.), *Roots: Linguistics in Search of its Evidential Base*. Berlin: De Gruyter, 185-206.
IBM Corp. 2013. *IBM SPSS statistics for Windows: Version 22.0*. Armonk: IBM Corp.
Kirsten, J. 2016. *Grammatikale Verandering in Afrikaans van 1911-2010*. PhD dissertation. Vanderbijlpark: North-West University.
Malamud, S.A. 2013. (In)definiteness-driven typology of arbitrary items. *Lingua* 126, 1-31.
Posio, P. & M. Vilkuna. 2013. Referential dimensions of human impersonals in dialectal European Portuguese and Finnish. *Linguistics* 51, 177-229.
Rasinger, S.M. 2013. *Quantitative Research in Linguistics: An Introduction*. London: Bloomsbury.
Siewierska, A. 2004. *Person*. Cambridge: Cambridge University Press.
Siewierska, A. 2011. Overlap and complementarity in reference impersonals: Man-constructions vs. third person plural-impersonal in the languages of Europe. In A. Malchukov & A. Siewierska (eds.), *Impersonal Constructions: A Cross-Linguistic Perspective.* Amsterdam: John

Benjamins, 57-89.
Siewierska, A. & M. Papastathi. 2011. Towards a typology of third personal plural impersonals. *Linguistics* 49, 575-610.
Sprouse, J. & D. Almeida. 2012. Assessing the reliability of textbook data in syntax: Adger's *Core Syntax*. *Journal of Linguistics* 48, 609-652.
van der Auwera, J. 2011. Standard Average European. In B. Kortmann & J. van der Auwera (eds.), *The Languages and Linguistics of Europe: A Comprehensive Guide*. Berlin: De Gruyter, 291-306.
van der Auwera, J., V. Gast & J. Vanderbiesen. 2012. Human impersonal pronoun uses in English, Dutch and German. *Leuvense Bijdragen* 98, 27-64.
van der Auwera, J. & V.A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2, 79-124.
Van Olmen, D. & A. Breed. 2018. Human impersonal pronouns in Afrikaans: A double questionnaire-based study. *Language Sciences* 69, 1-29.
Van Olmen, D., A. Breed & B. Verhoeven. 2018. A corpus-based study of the human impersonal pronoun (*'n*) *mens* in Afrikaans: Compared to *men* and *een mens* in Dutch. *Languages in Contrast* (*Online First*)
http://www.jbe-platform.com/content/journals/10.1075/lic.17004.van (accessed on 17 August 2018)
Weerman, F. 2006. It's the economy, stupid: Een vergelijkende blik op *men* en *man*. In M. Hüning, U. Vogl, T. van der Wouden & A. Verhagen (eds.), *Nederlands tussen Duits en Engels*. Leiden: Stichting Neerlandistiek Leiden, 19-47.
Zifonun, G. 2001. *Man lebt nur einmal*: Morphosyntax und Semantik des Pronomens *man*. *Deutsche* 3, 232-253.
Zobel, S. 2016. A pragmatic analysis of German impersonally used first person singular *ich*. *Pragmatics* 26, 379-416.