Human impersonal pronouns in Swedish and Dutch. A contrastive study of *man* and *men*

Evie Coussé, Johan van der Auwera

University of Gothenburg, University of Antwerp

This paper presents a contrastive study of the human impersonal pronouns *man* in Swedish and *men* in Dutch. Both impersonal pronouns are etymologically derived from *man* ‘human being’ and they more or less have the same meaning. However, there are important differences in the usage of these pronouns. In this study, the similarities and differences between Swedish *man* and Dutch *men* are studied in a Dutch-Swedish parallel corpus. Analyzing a parallel corpus has the advantage of allowing one to both study the distribution of *man* en *men* in original texts and to contrast the use of these pronouns with their translations.

**Keywords:** human impersonal pronouns, contrastive analysis, parallel corpus; Swedish/Dutch

1. **Introduction**

The Germanic languages have or had a human impersonal pronoun that is etymologically derived from *man* ‘human being’, e.g. *man* in German, Swedish and Norwegian, and *men* in Dutch. In this contribution, we will compare the use of two of these pronouns in more detail, viz. Swedish *man* versus Dutch *men*.

A typical use involves human generic reference: both *man* and *men* refer to ‘people in general’ and can accordingly be replaced by the impersonal pronoun *one* or the indefinite pronouns *anyone* or *everyone* in English. (1) illustrates this uses with Dutch *men*.

(1) *Men leeft maar één keer.* (Weerman 2006:26)

“One only lives once.”

(1) holds true for all human beings. Ramat and Sansò (2007:99) point out that these types of sentences are close to generic statements such as *the lion is dangerous* or *lions are dangerous*. It is therefore not surprising that this usage of *man/men* has been called ‘generic reference’ (Teleman et al. 1999:2:394). The generic meaning of *man/men* can be traced back to a generically used noun ‘man’ referring to the entire human race (Ramat and Sansò 2007:99). This origin can also be held responsible for some morphosyntactic properties. Dutch *men,*
for instance, combines with a third person singular verb despite its plural reference.¹ *Man* and *men* can also refer to a delimited group of people or even to individual human beings. In what follows, we describe these uses in terms of the hierarchy of Givón (1984:387), also used by Ramat and Sansò (2007). The use we have just described is the one on the left.²

```
generic > non-referential indefinite > referential indefinite > definite
```

The second use is the ‘non-referential indefinite’ one: human generic reference can be contextually delimited to a subgroup of humanity. Pettersson (1978:21) illustrates this with example (3), in which *man* only refers to people living or staying in France.

(2)  *Man dricker vanligen vin till maten i Frankrike.* (Pettersson 1978:21)
“One usually drinks wine with food in France.”

Non-referential impersonal *man* and *men* are (a) indefinite as they can be argued to refer to a group of individuals that has not been mentioned before and that cannot be identified more precisely in the given context, and (b) non-referential since they are used in a general context.

The pronouns *man* and *men* can also refer to indefinite human referents in episodic contexts, as is illustrated below.

(3)  *Men praat er al over in de stad.* (Haeseryn et al. 1997:258)
“Some people already talk about it in the city.”

Ramat and Sansò (2007:102) call this type of reference ‘referential indefinite’ but it is also known as ‘vague reference’ (Teleman et al. 1999:2:395, Altenberg 2004:94) or ‘arbitrary reference’ (Egerland 2003:73). The pronoun can be used in order to keep the identity of the referent vague. In (3) it is hard to determine who exactly is talking in the city. The function of vague *man/men* is similar to the passive: the speaker wants to remain vague about the exact reference of the agent. This vagueness can also be expressed by subjects such as *someone* or *some people* (Ramat and Sansò 2007).

Finally, *man* (but not, or much less clearly, *men*) can be used to refer to a definite referent in an episodic context.³ In (4) and (5) *man* refers to a known specific referent, more specifically to the speaker, and it can be replaced by the personal pronoun *jag* ‘I’.

(4)  *Uh uh, vad man blir trött och slö av den här hettan.* (Pettersson 1978:22)
“He, he, I get really tired and slow from this heat.”

(5)  *Nåja, man har väl läst ett antal böcker om astronomi.* (Pettersson 1978:22)
“Well, I have read a couple of books about astronomy.”

The definite usage of an impersonal pronoun is a fascinating phenomenon. Why does the speaker use the impersonal pronoun *man* to refer to himself instead of the simple first person singular personal pronoun? Pettersson (1978:22) argues that *man* in (4) expresses the personal experience of the speaker and at the same
time indicates that everyone in the same situation would experience the same thing. By using an impersonal pronoun, the speaker thus generalizes his own personal experience to a universal level (Linell and Norén 2005:122). Another motivation for using man with a definite usage can be observed in (5). Pettersson (1978:22) argues that man expresses modesty and reservation on behalf of the speaker. Using man instead of jag can here be seen as a strategy by which “the ego hides his or her intentions and desires in an anonymous mass” (Mühlhäuser and Harré 1990:199).

This brief survey shows that both pronouns largely refer to the same types of referents. However, there are important differences in the usage of the pronoun in both languages. First, it appears that man is a stylistically versatile pronoun, whereas men has a restricted use. Haeseryn et al. (1997:256) indicate that men is restricted to formal language. Weerman (2006:31) supports this observation with frequency counts indicating that men prevails in formal written language. He assumes – following Duinhoven (1990) and Paardekooper (1991) – that men is on its way out of the language. Typical of such elements is that they find a last stronghold in an educated written register, which is acquired at a later age. Swedish man, however, does not show any such signs of regress. Altenberg (2004:93) reports a high frequency of man in both fiction and non-fiction prose, which points to the pronoun’s versatility.

A second difference between man and men is that man is by far the most frequent pronoun in the language used to express impersonal reference, whereas men competes with what are ‘normally’ the personal pronouns je ‘you’, we ‘we’ and ze ‘they’, as illustrated in (6) and (7), as alternatives to (1), and in (8), a counterpart to Swedish (2).

(6) Je leeft maar één keer.
(7) We leven maar één keer.
(8) In Frankrijk drinken ze normaal wijn bij het eten.

It should be noted that these examples are not entirely synonymous with the sentences containing men. As opposed to men, these personal pronouns are specific about the speaker being included or excluded from the impersonal reference. Haeseryn et al. (1997:258) argue that impersonal je and we always include the speaker in the reference, whereas impersonal ze is speaker-exclusive.

Swedish personal pronouns do not exhibit the same flexibility to express impersonal reference as in Dutch. Telemann et al. (1999:2:395) mention only in a footnote that man can be replaced in spoken language by dom ‘they’, if the speaker or hearer are not included in the reference. Pettersson (1978:25) notes that the personal pronoun du ‘you’ can be used instead of impersonal man in Swedish. However, he disapproves of a routine replacement of man by du as one “runs the risk to irritate the hearer and to simply make a fool of oneself”. This recommendation does not however stop younger generations from increasingly using impersonal du in everyday speech (Törnudd-Jalovaara 1997, Fremer 1998, 2000). Note that the literature does not mention any impersonal usage of the personal pronoun vi ‘we’.

In this contribution, we will study the similarities and differences between man and men in detail in a Dutch-Swedish parallel corpus. The contribution builds on some existing contrastive studies of impersonal pronouns in Germanic, such as Norell (1996), Johansson (2003) and esp. Altenberg (2004). The latter is
a study of the correspondences between Swedish man and its English cognate one in an English-Swedish parallel corpus. It shows that man is used more frequently than one and that it is translated by a wide array of alternative forms in English.

The rest of this article is structured as follows. In section 2, we introduce the Dutch-Swedish parallel corpus that forms the empirical basis of this study. In section 3, the distribution of man and men in the different parts of the parallel corpus is discussed. In sections 4 and 5, the correspondences of man and men are analyzed in detail. The results of the study are summarized in section 6.

2. Corpus and method

The empirical basis of our contrastive analysis is the Dutch-Swedish parallel corpus. This parallel corpus was compiled within the SALT project (Språkbankens Arkiv för Länkade Texter ‘The Swedish Language Bank Archive of Aligned Texts’) at the University of Gothenburg. The Dutch-Swedish parallel corpus consists of Swedish originals that are sentence-aligned with their translations into Dutch on the one hand and Dutch originals that are sentence-aligned with their Swedish translations on the other. The corpus contains both contemporary novels and non-fiction texts in Swedish and Dutch. The investigated corpus comprises seven novels in Swedish together with their translations and five Dutch novels plus translations. The non-fiction part of the corpus contains four Swedish texts and one Dutch text together with their translations. Table 1 shows the number of words in the parallel corpus.

Table 1. Size of the Dutch-Swedish parallel corpus (in this study).

<table>
<thead>
<tr>
<th></th>
<th>Non-fiction</th>
<th>Fiction</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch original</td>
<td>44 761</td>
<td>286 624</td>
<td>331 385</td>
</tr>
<tr>
<td>Swedish original</td>
<td>205 472</td>
<td>674 272</td>
<td>879 744</td>
</tr>
<tr>
<td>Dutch translation</td>
<td>224 462</td>
<td>735 553</td>
<td>960 015</td>
</tr>
<tr>
<td>Swedish translation</td>
<td>42 835</td>
<td>275 985</td>
<td>318 820</td>
</tr>
<tr>
<td>Total</td>
<td>517 530</td>
<td>1 972 434</td>
<td>2 489 964</td>
</tr>
</tbody>
</table>

Table 1 shows that the fiction texts outnumber the non-fiction texts (1 972 434 words vs. 517 530 words). This imbalance has to do with a general shortage of non-fiction texts in Dutch and Swedish that are translated into the other language. The relative paucity of non-fiction texts does not allow a comparison of man and men in the two text genres.

Table 1 also reveals that the total size of the available Swedish original texts exceeds the size of the Dutch originals (879 744 words vs. 331 385 words). This unbalanced distribution is the result of practical problems during the compilation of the Dutch-Swedish parallel corpus. Rawoens (2008:121) indicates that a number of Dutch originals could not be integrated in the corpus because of unresolved copy right issues. The uneven distribution in the originals, however, is balanced by the inverse distribution of Dutch and Swedish translations (960 015 words vs. 318 820 words). In order to ensure a correct representation of man and men in this study, we will present their distribution in both original and translated texts.
Via the online search interface GLOSSA, all instances of *man* and *men* were extracted from the corpus both in the original texts and translations. The pronouns were retrieved together with the sentence they occurred in and the corresponding sentence in the other language. All hits were manually checked. In total, we investigated 3991 instances of the impersonal pronouns *man* and *men*.

### 3. Frequencies of *man* and *men*

Table 2 shows the relative frequency of *man* and *men* in the original texts and translations. In addition to the absolute numbers, the frequencies are given in number of tokens per ten thousand words in order to normalize the different sizes of the investigated subcorpora.

<table>
<thead>
<tr>
<th>Text type</th>
<th>Swedish original</th>
<th>Swed. translation</th>
<th>Dutch original</th>
<th>Dutch translation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>n</em></td>
<td><em>r</em></td>
<td><em>n</em></td>
<td><em>r</em></td>
</tr>
<tr>
<td>Fiction</td>
<td>1787</td>
<td>26.5</td>
<td>911</td>
<td>33.0</td>
</tr>
<tr>
<td>Non-fiction</td>
<td>530</td>
<td>25.8</td>
<td>128</td>
<td>29.9</td>
</tr>
<tr>
<td>Total</td>
<td>2317</td>
<td>26.3</td>
<td>1039</td>
<td>32.6</td>
</tr>
</tbody>
</table>

Let us first have a look at the relative frequency of *man* and *men* in the original texts. Table 2 shows that *man* is used more than five times as frequently as *men* (26.3 vs. 5.1). This finding confirms the expectation that *man* is more common in Swedish compared to *men* in Dutch. Moreover, it appears that *man* occurs about as often in fiction as in non-fiction texts (26.5 vs. 25.8). Statistical testing shows that text genre does not have a significant influence on the distribution of *man* (Pearson $\chi^2 = 0.3$, df = 1, $p = 0.58$). This suggests that *man* is a versatile pronoun, which can readily be used in diverse types of (written) language. Dutch *men*, on the contrary, is used relatively often only in non-fiction texts (17.9 vs. 3.1 in novels). The effect of text genre is highly significant ($\chi^2 = 163.89$, df = 1, $p < 0.05$). Similar results have been reported by Weerman (2006:31). He found that *men* is far more common in academic texts (31 tokens per 10 000 words) and newspapers (11 per 10 000) compared to novels (5 per 10 000).

The table allows us to complement the results for the original texts with the relative frequencies of *man* and *men* in the translated texts. The same tendencies can be observed. First, *man* is almost seven times as frequent as *men* in translations (32.6 vs. 4.7). Moreover, *man* is about as frequent in fiction and non-fiction translated texts (33.0 vs. 29.9) and there is no statistical effect of text genre ($\chi^2 = 1.12$, df = 1, $p = 0.29$). This is quite surprising as the Swedish translations are based on Dutch originals which have been shown to have a different preference for *men* depending on text genre. The reverse tendency can be observed in the Dutch translations. Despite the fact that text genre does not significantly affect the distribution of *man* in the Swedish originals, the translated non-fiction texts do show a significant higher preference for *men* than fiction texts in Dutch ($\chi^2 = 159.45$, df = 1, $p < 0.05$). The findings in the translated texts support our earlier conclusion that the distribution of *man* is
unaffected by text genre as opposed to *men*, which is restricted to non-fiction text genres.

Finally, the table gives us the opportunity to compare the translations with the source texts. It shows that the relative frequencies of *man*/*men* are surprisingly similar in the translations compared to the original texts. Especially the distribution of *men* in both originals and translations (5.1 vs. 4.7) does not show a significant difference ($\chi^2 = 0.82$, df = 1, $p = 0.37$). This finding indicates that the Dutch translators have succeeded in reducing the many instances of *man* in the Swedish originals to a level appropriate for Dutch. This of course implies that the translators must have resorted to other translation strategies in order to find alternative equivalents for *man*. We will give more details on the exact nature of these translation pairs in Sections 4 and 5. The relative frequency of *man* is somewhat higher in Swedish translations compared to originals in the same language (32.6 vs. 26.3). The difference is significant ($\chi^2 = 32.76$, df = 1, $p < 0.05$). However, we should probably consider the relative frequency of *man* in our original texts to be on the low side. Altenberg (2004:99) reports an average of 34.9 times *man* per ten thousand words in his original texts, which is closer to the ratio of 32.6 found in our translations.

## 4. Dutch correspondences of *men*

We now go on to investigate the Swedish correspondences of *men* in the parallel corpus. Table 3 provides the frequency of different Swedish translations of *men* (columns 2-3) and the frequency of different source constructions in Swedish that were translated into *men* (columns 4-5). The data are presented in the same way as Table 3 in Altenberg (2004:104) in order to ensure comparability – we have only chosen other labels.

### Table 3. Swedish correspondences of *men*

<table>
<thead>
<tr>
<th>Swedish correspondence</th>
<th>Swedish translation</th>
<th>Swedish source</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Congruent subject</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>men</em> <em>‘one’</em></td>
<td>155</td>
<td>304</td>
</tr>
<tr>
<td>personal pronoun</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>general noun</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>indefinite pronoun</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>(b) Other subject</td>
<td>8</td>
<td>116</td>
</tr>
<tr>
<td>passive</td>
<td>6</td>
<td>102</td>
</tr>
<tr>
<td>other subject replacement</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>(c) No subject</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>non-finite clause</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>nominalisation</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>ellipsis</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>(d) No correspondence</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>170</td>
<td>454</td>
</tr>
</tbody>
</table>

Table 3 shows that *man* is the preferred correspondence for *men*, both as a translation and as a source. The correspondence is strongest when *men* is translated into Swedish. This suggests that *man* can be considered to be the neutral translation equivalent of *men*. We can illustrate these correspondences with some examples taken from the parallel corpus.6
In een gammele hut kan men even gelukkig zijn als in een burcht. [Pae]
Man kan bli lika lycklig i en rankig hydda som i en borg.
“In a ruinous shed, one can be as happy as in a castle.”

På sensommaren 1135 kunde man gå nästan torrskodd över Donau. [Nil]
In de nazomer van 1135 kon men de Donau bijna met droge voeten oversteken.
“In the late summer of 1135, one could almost cross the Danube with dry feet.”

Men mag dus aannemen dat de meeste vrouwen die verliefd werden op een andere vrouw, hun gevoelens niet konden plaatsen of benoemen. [Dek]
Man kan alltså anta att de flesta kvinnor som blev förälskade i en annan kvinna inte kunde placera eller benämna sina känslor.
“One can thus presume that most women who fell in love with another woman could not place or name their feelings.”

These examples demonstrate that *men* and *man* can express the various types of impersonal reference presented in the introduction. In (9) both pronouns are used for generic reference, in (10) for indefinite reference to a contextually delimited subgroup of humanity, and in (11) for vague reference. There are no corresponding definite uses of *man* and *men*, however: for *men* this is impossible or difficult.

Other – less frequent – correspondences of *men* can be divided into some broad categories. First, *men* corresponds to other subject pronouns and nouns which express some kind of impersonal reference. Altenberg (2004:104) refers to this group as ‘Congruent subject’. Examples include *de* ‘they’ (12) and the general noun *folk* ‘people’ (13).

Men was gelaten geweest, nieuwsgierig, en vooral helder. [Kra]
De hade känt resignation, nyfikenhet och framför allt klarhet.
“They had been calm, curious, and above all clear.”

‘René wil niet dat men weet dat hij hier is’, zegt Alma later in bed. [Cla]
René vill inte folk ska veta att han är här’, sa Alma sen i sängen.
“The is not want people to know that he is here”, Alma later said in bed.”

Second, the correspondences of *men* may be reorganized so that the subject position is filled by another referent; this is the category ‘Other subject’ in Table 3. Most of these cases are agentless passives (22%).

Han påstods vara utomlands. [Sjö]
Men zei dat hij naar het buitenland was.
“He is said to be abroad.”

Most of the passives are synthetic with the passive suffix -s attached to the finite verb. These *s*-passives predominantly use mental verbs or verbs of saying that are combined with an infinitive with the same subject as the finite verb (see Teleman et al. 1999:2:397 and 1999:3:575). This type of raising construction can also be found in English, especially with verbs of saying, e.g. *he is said to
Evie Coussé, Johan van der Auwera

be abroad, but is hardly possible in Dutch, e.g. *Hij werd gezegd in het buitenland (te) zijn (see van der Auwera and Noël 2011:25-26). Translators thus have to resort to alternative constructions in order to translate this relatively frequent Swedish passive construction. One possibility is to reorganize the passive clause as an active one with men as a vague placeholder for the agent. This translation strategy can be held responsible for the relatively great number of passive correspondences for men in the corpus.

Apart from passives, there are other, less common restructuring strategies. In example (15), a subjectless infinitival clause is translated into a finite clause with men.

(15) Det är en storsak att mottaga sitt första beundrarbrev. [Tun]

*Het is heel wat wanneer men voor het eerst een brief van een bewonderaar krijgt.*

“It is big thing to receive your first letters of admirers.”

In sum, men appears to express all types of impersonal reference distinguished in the introduction. Moreover, men shares this versatile impersonal reference with man, which is shown to be the preferred correspondence of men. It is important to note that men is not found with definite reference, neither in the source texts nor in the translations, and when man translates men, man understandably also has no definite reference.

5. Dutch correspondences of man

Table 4 presents the correspondences for man in Dutch translations and sources texts.

<table>
<thead>
<tr>
<th>Dutch correspondence</th>
<th>Dutch translation</th>
<th>Dutch original</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Congruent subject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>je/ge/u ‘you’</td>
<td>1888</td>
<td>1047</td>
</tr>
<tr>
<td>men ‘one’</td>
<td>290</td>
<td>150</td>
</tr>
<tr>
<td>ze ‘they’</td>
<td>295</td>
<td>56</td>
</tr>
<tr>
<td>we ‘we’</td>
<td>49</td>
<td>4</td>
</tr>
<tr>
<td>ik ‘I’</td>
<td>76</td>
<td>6</td>
</tr>
<tr>
<td>hij ‘he’, zij ‘she’, het ‘it’</td>
<td>66</td>
<td>15</td>
</tr>
<tr>
<td>noun phrase</td>
<td>38</td>
<td>15</td>
</tr>
<tr>
<td>indefinite pronoun</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td>(b) Other subject</td>
<td>320</td>
<td>251</td>
</tr>
<tr>
<td>passive</td>
<td>212</td>
<td>147</td>
</tr>
<tr>
<td>zijn/vallen te infinitive</td>
<td>30</td>
<td>27</td>
</tr>
<tr>
<td>other subject replacement</td>
<td>78</td>
<td>77</td>
</tr>
<tr>
<td>(c) No subject</td>
<td>61</td>
<td>92</td>
</tr>
<tr>
<td>non-finite clause</td>
<td>27</td>
<td>35</td>
</tr>
<tr>
<td>phrase</td>
<td>21</td>
<td>44</td>
</tr>
<tr>
<td>ellipsis</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>(d) No correspondence</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>Total</td>
<td>2317</td>
<td>1039</td>
</tr>
</tbody>
</table>

Table 4 shows that the most common correspondence of man is not the cognate pronoun men but rather the personal pronoun je, both in the translations and in
Human impersonal pronouns in Swedish and Dutch

the sources (45% and 40%). As exemplified in the introduction, the personal pronoun je can be used in Dutch to express impersonal reference. Sometimes, however, it is not clear whether je is used with impersonal or deictic reference, as illustrated in example (16). This sentence can be interpreted as a suggestion directed to the hearer to put cloves in his tooth as an old remedy to ease tooth ache. In this case, je has definite reference. The same utterance can also be read as a general wisdom that cloves help to ease tooth ache. In that case, je is used generically to refer to a human being in general.

(16) Die moet je in je kies stoppen. [Mul]
Man trycker in den i tanden.
“You need to put that one in your molar.”

Note that the table conflates the results for je, ge and u, which all function as second person singular pronouns in Dutch. In our corpus ge occurs in southern texts and it has the same referential flexibility as je. U is the polite form. In the corpus it always has definite reference:

(17) Känns det aldrig svårt och blir man inte rädd?[Gui]
Is dat niet verschrikkelijk moeilijk en bent u nooit bang?
“Does it never feel difficult and are you never afraid?”

There some rare correspondences of other personal pronouns, viz. ze ‘they’ and we ‘we’:

(18) Gammalt släppte man ålar i brunnar, det visste han. [Ekm]
Vroeger stopten ze palingen in putten, dat wist hij.
“Earlier, they put eels in wells, that he knew.”

(19) Ser man till kvinnans civilstånd var den generella dödligheten högre bland gifta än bland ogifta fram till 1880. [Joh]
Kijken we naar de burgerlijke staat van de vrouw, dan is tot 1880 de algemene sterfte onder getrouwde vrouwen hoger dan onder ongetrouwde.
“If we look at the marital status of women, then until 1880 the general mortality was higher among married women compared to unmarried ones.”

In (18) the third person plural pronoun ze refers to people that lived in the past. It is appropriate to use speaker-exclusive ze for this type of non-referential impersonal reference as the context excludes the speaker and hearer from the potential referents. In (19) the first person plural pronoun we occurs as a translation equivalent of man used with a vague reference. The impersonal usage of we is more specific than vague man since it explicitly includes both speaker and hearer in the reference. The pronoun is often used in formal written texts in Dutch as a pluralis modestiae in order to avoid the use of the first person singular pronoun ik, which is felt to be too direct and immodest (Haeseryn et al. 1997:238). Using the pronoun we instead of ik has the rhetorical advantage that the reader gets involved in the argumentation and the observation of the writer is generalized to everyone who reads the text.

The personal pronouns ze, we and also ik can also refer to definite referents in an episodic context. (20) is an illustration for ze, and (21) for ik.
Evie Coussé, Johan van der Auwera

(20) -- Hur fan kan man ställa barackerna så tätt? frågade Kurt Wallander. [Man]
   ‘Hoe kunnen ze die barakken nou zo dicht elkaar neerzetten!’ vroeg Wallander.
   ‘‘How the hell can they put the barracks so close to each other?’ Kurt Wallander asked.’

(21) -- Det kan man knappast säga. [Man]
   ‘Zo zou ik het nou ook weer niet willen formuleren.
   ‘I would not put it like that.’

The translation of man with the personal pronouns ze, we and especially ik (which cannot be used with impersonal reference) gives us a good insight into the motivations for the definite usage of man. There seems to be quite a discrepancy in the pragmatics of the originals and the translations. In example (20), for instance, the speaker asks a rhetorical question about everyone except himself. However, if we read the translation, we may assume that the character knows who is responsible for the position of the barracks. It is clear that the character expresses himself a bit more indirectly with respect to the identity of the referent in the original than in the translation. We could cast this difference in the words of Mühlhäusler and Harré (1990:199): “the ego hides his or her intentions and desires in an anonymous mass” in the Swedish original, whereas the Dutch alter ego dares to speak out in the translation. Interestingly, Altenberg (2004:105) makes the same observation for English personal pronouns corresponding to man in Swedish. He notices that man has a tactful effect in Swedish whereas the corresponding personal pronouns they, we and I are used with a more direct reference in English.

Table 4 also shows that man sometimes corresponds to the third person pronouns hij, zij and het, mainly in the translations (3%). (22) illustrates that these personal pronouns are not used with impersonal reference but that they refer anaphorically to a referent that has been introduced before. In (22) this referent is familie ‘family’.

(22) Toen de familie weigerde het losgeld te betalen kreeg zij foto’s thuis waarop de gegijzelde poedelnaakt vastgebonden op een brits lag. [Cla]
   När familjen vägrade att betala lösensumman fick man hem foton där gisslan låg spritt naken fastsurrad på en brits.
   “When the family refused to pay the ransom, they got the pictures sent home where the hostage lied tied to a bed stark naked.”

Swedish man also corresponds to Dutch indefinite pronouns, again mainly in the translations (1%). Thus iemand ‘someone’ vaguely refers to one or more individuals in (23). Interestingly, indefinite pronouns are also used to refer to a definite referent, and in (24) it seems to refer to the speaker himself; it has a similar pragmatic effect as the definite use of man in the original, i.e., the speaker hides his identity behind the vague reference of the indefinite pronoun.

(23) Man skulle behöva göra undersökningar för att få fram det. [Ekm]
   Om dat boven tafel te krijgen zou iemand onderzoek moeten doen.
   “Someone would need to do research to uncover that.”
(24) **Varför får man inte ha sina hemligheter?** [Nil]

> Varför får man inte ha sina hemligheter?
>  
> “Why can’t one have secrets?”

A last type of congruent subjects are noun phrases with nouns such as `mens` ‘human being’ in either the singular or plural form accompanied by the definite or indefinite article (`de mens, de mensen, een mens` and `mensen`). (25) and (26) offer illustrations with `de mens` and `een mens`; the latter is typical for southern Dutch.

(25) **Hade Gud gjort det osynligt, eftersom man ändå inte såg någonting?** [Nil]

> Hade Gud gjort det osynligt, eftersom man ändå inte såg någonting?
>  
> “Did God make it invisible, because man did not see anything anyway?”

(26) **Man vill ju gärna dela med sig av det lilla man vet om livet.** [Tun]

> Man vill ju gärna dela med sig av det lilla man vet om livet.
>  
> “One just wants to share the little one knows about life with others.”

Apart from noun phrases with `mens` various other noun phrases occur as correspondences of `man`. They are therefore more restricted in their referentiality. Some examples are:

(27) **Så kan man slippa samliv och nya graviditeter.** [Joh]

> Op die manier kan de vrouw gemeenschap en nieuwe zwangerschappen vermijden.
>  
> “This way, one avoids intercourse and new pregnancies.”

(28) **Pojken luktade liten pojke, och man fick nog tänka sig att han inte hade bytt tröja på rätt så många dar.** [Gus]

> De jongen rook zoals kleine jongens ruiken, en Stig vermoedde dat hij al een paar dagen in dezelfde trui rondliep.
>  
> “The boy smelled like small boys smell, and one could suspect that he had not changed his sweater for a number of days.”

In (27), the noun phrase `de vrouw` ‘the woman’ is more specific than the original generic pronoun `man` in that it only refers to the female sex instead of to humanity as a whole. However, the reference of `man` in the original version is similarly restricted thanks to general knowledge of the world. Example (28) illustrates how `man` can refer to a known individual.

Let us turn to the category ‘Other subject’ in Table 4. This category includes correspondences of `man` where the subject position is filled by another referent. Agentless passives are the most frequent strategy for replacing `man` by another subject referent, both in translations (9%) and sources (14%).

(29) **Man har funnit spår av liv som fanns här före istiden.** [Nil]

> Er zijn sporen gevonden van leven vóór de ijstijd.
>  
> “One has found traces of life that existed here before the ice age.”
A second type of correspondence of *man* that resembles the passives above is illustrated in (30) and (31). In these examples, the verb phrase in Swedish corresponds to a modal infinitive in Dutch, i.e., a construction with a modal expressing possibility, which consists of a finite verb *zijn* ‘to be’ or *vallen* ‘to fall’ and a *to* infinitive of the main verb (see Haeseryn et al. 1997:1035-1037, 1046-1047). In Dutch, the subject referent of *man* is backgrounded and replaced by the direct object referent of the Swedish clause, if available. Often, the Swedish main verb is combined with a modal auxiliary.

“The French on the contrary could be found on the small field every day.”

“That was probably also what one could expect.”

Apart from these two rather frequent types of correspondences, we found some minor strategies to replace *man* with another subject referent, collapsed in Table 4 under ‘other subject replacement’. In example (32), the Swedish verb phrase corresponds to the finite verb *zijn* ‘to be’ and an adjective expressing the meaning of the main verb in Dutch. This adjective is often derived from the main verb with the suffix *-baar* ‘-able’. The Swedish verb phrase usually contains a modal, as is also noticed by Altenberg (2004:109) for similar correspondences of *man* in English.

(32) Kunde man tänka sig, frågade han Vilhelm från Rubruck, att påven skulle låna ut hästar så att han lättare kunde ta sig till Compostela? [Nil] Was het denkbaar, vroeg hij aan Willem van Ruysbroeck, dat de paus hem paarden wilde lenen zodat hij gemakkelijker in Compostela kon komen?
“Could one imagine, he asked William of Rubruck, that the Pope would lend out horses so that he could get to Compostela more easily?”

A third category of correspondences (type c in Table 4) contains non-finite clauses without a subject referent corresponding to *man*. These clauses mostly take the form of agentless *to* infinitive clauses, as illustrated in (33).

(33) Vänskap kan beseglas genom att man stänker sitt blod i varandras fotspår. [Nil] Vriendschap kan bezegeld worden door bloed in elkanders voetsporen te sprenkelen.
“Friendship can be sealed through sprinkling blood in each other’s footsteps.”

A finite clause with *man* can also be reduced to a single phrase or word. In example (34), the subordinate clause with *man* corresponds to a Dutch noun phrase. The finite modal clause containing *man* in (35) corresponds to the modal adverbial *vermoedelijk* ‘probably’ in Dutch.

(34) Vänskap är som om detta konstiga hus hade något slags inverkan på hur man tänker. [Gus]
“Friendship is as if this strange house had some kind of influence on how one thinks.”

(35) Man kunde ana att de starkröda läpparna hade färgats med rödpenna vars spets först blivit uppböpt. [Ekm]

Vermoedelijk waren de felrode lippen geverfd met een kleurpotlood waarvan de punt eerst was bevochtigd.

“One could guess that the intense red lips were painted with a red pencil of which the tip was first made wet.”

Finally, the subject corresponding to man can be omitted in various ways, for example by coordination, as in (36). And for (37), one could say that the translator has chosen a totally different construction (the ‘No correspondence’ label of Table 4).

(36) Man hittade (enligt det ockulta ryktet) ett bergigt landskap som var aldeles fritt från is och snö, och där började man bygga en militärbas. [Nil]

Men trof daar (volgens het occulte gerucht) een bergachtig landschap aan, waar geen ijs en sneeuw lag, en begon met de bouw van een militaire basis.

“One found there (according to the occult rumor) a mountainous landscape that was entirely free from ice and snow, and one started to build a military base there.”

(37) -- Det kan man lugnt säga om Malm också, sa Martin Beck. [Sjö]

‘Dat was bij Malm ook het geval,’ zei Martin Beck.

“One could easily say that about Malm too, Martin Beck said.”

6. Conclusion

We can now summarize the main findings of this study. First, the data show that man and men are used with overlapping meanings in the Dutch-Swedish parallel corpus. Both pronouns are used in original texts and translations in order to express similar generic human reference, non-referential indefinite reference and referential indefinite (or vague) reference. Within these domains of reference, they are perfectly interchangeable and thus can serve as translation equivalents. With respect to definite reference, the study confirms that man and men are different. In our corpus men was never used to refer to definite referents as opposed to man, which could refer to the speaker himself or to another known referent.

Despite their overlapping meaning, man and men appear to be used rather differently in the Dutch-Swedish parallel corpus. First, man occurs relatively frequently in both fiction and non-fiction texts, whereas men appears to be mainly restricted to non-fiction prose. These findings indicate that man is stylistically more flexible than men, which is limited to more formal registers. Moreover, man seems to dominate the domain of impersonal reference in Swedish, whereas men competes with an array of alternative strategies to
express impersonal reference. The analysis of the Swedish correspondences of *men* showed that that *man* is by far the preferred form in both translations and source texts (88% and 65% respectively). Alternative strategies to express the impersonal reference of *men* do occur in Swedish but they are low in frequency. These findings contrast strongly with the Dutch correspondences of *man*, which are more diverse in nature. Of these, the personal pronoun *je* is most frequent in the corpus (up till 45% in translations and 40% in source texts), a point also made in van der Auwera et al. (2011). Another common Dutch alternative in the corpus is the third person plural pronoun *ze* (13% in translations and 5% in sources).

The use of the personal pronouns *je* and *ze* as frequent correspondences of *man* shows similarities with the findings of Altenberg (2004:104) in his comparison of Swedish and English. He showed that the most common English correspondence of Swedish *man* was not the impersonal pronoun *one* but rather the personal pronoun *you* (in 25% of all cases). Moreover, *they* was also relatively common as a correspondence of *man* in his corpus (in 5% of all cases). It is striking that Dutch and English seem to show similar correspondences of Swedish *man* despite the fact that they have impersonal pronouns with a different etymological background.8

Apart from these corresponding pronouns, the study has uncovered a wide array of alternative strategies to express impersonal reference. A frequent type of correspondence for both *man* and *men* is, for instance, the agentless passive and, to a lesser extent, non-finite clauses with *to* infinitives. These types of correspondences have also been reported by Altenberg (2004) for the English-Swedish parallel corpus. Other correspondences in this study are more dependent on specific structural differences between Swedish and Dutch, such as the tendency to translate *s*-passives of mental verbs and verbs of saying by active clauses with *men*.

References


van der Auwera, J., Gast, V. and Vanderbiesen, J. Forthc. “Human impersonal pronouns in English, Dutch and German”. Leuvense Bijdragen

Notes

1 Finite verbs in Swedish are not coded for number and person distinctions.
2 This classification is relatively abstract. For further semantic distinctions, see Cabredo Hofherr (2003, 2006), van der Auwera (2012) and van der Auwera et al. (Forthc). For this paper ‘use’ and ‘meaning’ can be taken to refer to context or meaning-in-context and we remain agnostic on the question whether the pronouns maintain one single meaning throughout these contexts.
3 Egerland (2003:74) notes that this usage is traditionally considered to be substandard and that not all native speakers of Swedish accept it. Other descriptions of definite man (Pettersson 1978, Teleman et al. 1999, Linell and Norén 2005) do not comment on the sociolinguistic status of this usage in Swedish.
4 The project homepage is http://spraakbanken.gu.se/salt/. Full bibliographic details can be found in Rawoens (2008:347-349). Note that we did not investigate Dynamisk psykiatri i teori och praxis – Moderne psychiatrie by Johan Cullberg and Het meesterstuk – Mästerprovet by Anna Enqvist. These texts were not available via the online search interface GLOSSA at the website of the Swedish Language Bank. This implies that our corpus is somewhat smaller in size than the corpus described and investigated by Rawoens (2 489 964 words instead of 2 990 599 words).
5 We would like to thank Gösta Ericsson for his help with part of the corpus analysis.
6 Source sentences are systematically placed before their translations and are identified by a code that identifies the text in the corpus (cf. Rawoens 2008:347-349).
7 The table abstracts away from the fact that most personal pronouns in Dutch have a so-called strong form with a full vowel (i.e. jij, gij, zij, wij) and a weak unstressed form with a schwa (e.g. je, ge, ze, we). It is suggested in the literature (e.g. Haeseryn et al. 1997:258) that only the weak forms can be used with indefinite reference. In our corpus, however, we do find a few occurrences of a full form expressing indefinite reference:

    [...] och man måste alltid utgå från att motståndaren var mer intelligent och bättre informerad än man själv. [Gui]
    [...] en je moest er altijd van uitgaan dat de tegenstander intelligenter en beter geïnformeerd was dan jij zelf.
    “[…] and you always had to assume that the enemy was more intelligent and better informed than yourself.”

8 However, Dutch shows a stronger preference for personal pronouns than English. This difference is significant for both pronouns. Dutch je is used in 1465 of 3356 cases vs. you in 962 of 3918 cases, yielding Pearson $\chi^2 = 291.85$ (df = 1, p < 0.05). Dutch ze is used in 351 of 3356 cases vs. they in 204 of 3918 cases, yielding Pearson $\chi^2 = 70.75$ (df = 1, p < 0.05).

Authors’ addresses

Evie Coussé
Department of Languages and Literatures
University of Gothenburg
Box 200
405 30 Gothenburg
Sweden

evie.cousse@gu.se

Johan van der Auwera
Center for Grammar, Cognition and Typology
University of Antwerp