

From you to me (and back)

The flexible meaning of the second
person pronoun in Dutch

Sammie Tarenskeen

Optimal Communication
Department of Linguistics
Radboud University Nijmegen
P.O. Box 9103
6500 HD Nijmegen
The Netherlands
www.ru.nl/pionier
sammietarenskeen@gmail.com

From you to me (and back)

The flexible meaning of the second person pronoun in Dutch

Master's thesis
General Linguistics
Faculty of Arts
Radboud University Nijmegen

July 2010

Sammie Tarenskeen
0448176

First supervisor	Helen de Hoop
Second supervisor	Kees de Schepper

Acknowledgements

The first I want to thank is, of course, Helen de Hoop. Thank you for being the most wonderful supervisor I can imagine, for your suggestions, support and comments, for encouraging me every time when I felt hopeless again, but most of all for your involvement. I also want to thank Kees de Schepper, my second supervisor, for your helpful comments. Thank you and the other organisers of the workshop Local Pronouns for giving me the opportunity to give a talk on this occasion. Now, I can see that this was the starting point for a new stage in my thesis: the stage of writing quickly, and with pleasure. I also want to mention Mirjam Broersma here, the nicest person I have ever worked for. Thank you for the conversations about the talk and all other things; it always helps me.

Special thanks go to Peia Prawiro-Atmodjo, for the nice moments, for forcing me sternly to formulate my first ideas, and for giving me the ‘kick under the ass’ that I needed to start writing. I also thank Geertje van Bergen, for your help with the corpus of course, but especially for persuading me to give a practising talk before the workshop mentioned above, which was, apparently, a magic trick to diminish my nervousity, and finally, for being a nice roommate on Thursdays. This brings me to Sander Lestrade and Margot van den Berg; thanks a lot for lending me your rooms once a week, which was not only good for my thesis, but also for my idea of being part of the Linguistics department a bit.

Of course, I also thank my family and friends. Many of you asked me about my thesis, not only how it went, but also in more detail about my topic, my method, my findings. I am grateful for your questions, for your suggestions, for your interest. In particular, I thank Verena and Elisabeth for being my ‘scriptiemaatjes’ in the final writing stage; I enjoyed working, lunching, and talking together so much. Thank you papa and mama, for always being there, thank you Meenah, for always cheering me up when I am losing confidence again, thank you Frank and Angela, for always being interested, and finally, thank you Merlijn, for always everything.

Table of contents

Acknowledgements	v
Table of contents	vii
Conventions	x
Abbreviations	x
1 Introduction	1
2 <i>Je</i> in the literature	3
2.1 The original meanings: deictic and generic <i>je</i>	3
2.1.1 <i>Je</i> as a second person pronoun	3
2.1.2 <i>Je</i> as an impersonal pronoun	5
2.1.3 Summary	7
2.2 Previous accounts of self-referring <i>je</i>	7
2.2.1 Link with deictic <i>je</i> only: speech accommodation	8
2.2.1.1 Appel's first account	8
2.2.1.2 The problems	9
2.2.2 Link with generic <i>je</i> only: modesty	10
2.2.2.1 Bennis et al. and Appel	10
2.2.2.2 The problems	11
2.2.3 Link with both deictic and generic <i>je</i> : empathy	14
2.2.3.1 Van Wassenaar	15
2.2.3.2 Onrust	16
2.2.3.3 Klaasse	17
2.2.4 Conclusions on the literature	18
2.3 General conclusions	19
3 <i>Je</i> in spontaneous speech: a corpus study	21
3.1 Data collection	21
3.1.1 Corpus I. Teachers of Dutch	22
3.1.2 Corpus II. Spontaneous speech	23
3.1.3 Summary	24
3.2 The correct interpretation of <i>je</i>	24
3.2.1 A grammatical cue: episodicity	25

3.2.2	The role of context	27
3.2.3	Self-referring <i>je</i> and context	29
3.2.4	Conclusions	31
3.3	Self-referring <i>je</i>	31
3.3.1	Explicit cues: questions and pronoun switches	32
3.3.2	Implicit cues: probability and recognition	34
3.3.3	Subjectivity and the role of the hearer	37
3.3.4	Conclusions	39
3.4	Generic <i>je</i> and self-referring <i>je</i>	39
3.4.1	Explicit restrictions	40
3.4.2	Implicit restrictions	42
3.4.3	Self-referring <i>je</i> as an instance of generic <i>je</i>	43
3.4.4	A first motivation for the use of self-referring <i>je</i>	45
3.4.5	Conclusions	47
3.5	Deictic <i>je</i> , generic <i>je</i> and self-referring <i>je</i>	48
3.5.1	The ambiguous cases	48
3.5.2	From personal to impersonal pronoun	49
3.5.3	A final motivation for the use of self-referring <i>je</i>	51
3.5.4	Conclusions	53
3.6	General conclusions	54
4	Impersonals and seconds: a crosslinguistic comparison	57
4.1	Generic constructions and self-reference	57
4.1.1	Generic and arbitrary	58
4.1.2	Specific reference in episodic contexts	60
4.1.2	The suggestion of a generalisation	63
4.1.2.1	Swedish <i>man</i> and French <i>on</i> : pragmatics	63
4.1.2.2	More indirect speaker-reference	65
4.1.2.3	Non-first person specific reference	68
4.1.3	Conclusions	69
4.2	Second person pronouns, genericity, and the first person	70
4.2.1	Second person pronouns and generic reference	70
4.2.2	Generic use of emphatic forms	72
4.2.3	Second person pronouns and the first person	76
4.2.4	Conclusions	78
4.3	Impersonals versus seconds	79
4.3.1	Impersonal constructions: effects of genericity	79
4.3.2	Seconds: generalisation and involvement	81
4.3.2.1	The origin of generic seconds	81

4.3.2.2 The effects of self-referring seconds	82
4.3.2.3 From second to first person pronoun?	85
4.3.3 Conclusions	86
4.4 General conclusions	86
5. Conclusion	89
References	93

Conventions

Throughout this thesis, many examples have been taken from the CGN (Spoken Dutch Corpus). For the sake of clarity, I left out those hesitations and particles (the latter being used a lot in Dutch) that might blur the intention of the speaker. In all cases, the CGN code is given, which makes it possible for the reader to recover the original utterance in the CGN.

Dutch is closely related to English and therefore in many ways very similar. As I do not focus on morphosyntax but on semantics and pragmatics in this thesis, I found that interlinear glosses would be confusing while they are in fact unnecessary. I therefore decided not to present glosses, but only an English paraphrase which is as literal as possible. When a correct English paraphrase deviated too much from the Dutch original utterance, I added a more literal translation. Languages other than Dutch and English are glossed and translated in the conventional way. All bold markings in those examples were added by me.

Abbreviations

COMP	complementiser	PL	plural
DET	determiner	POL	polite
IMP	imperfective	REF	reflexive
NEG	negation	SG	singular
PART	particle	STR	strong
PF	perfective		

Introduction

It has widely been observed that the Dutch pronoun *je* ‘you’ sometimes refers to the speaker. At first sight, this phenomenon is quite surprising, because *je* is a second person pronoun. As *je* can also refer generically, this form already corresponded to more than one meaning; with its apparent first person reference, it is even more ambiguous. This ambiguity can lead to misinterpretations, as shown in the following example from Klaasse (2005).

- (1) SP1 Hoe is het?
 ‘How are you?’ (lit: ‘How is it?’)
 SP2 Hm, **je** begint behoefte te krijgen aan vakantie...
 ‘Well, **you** begin to feel the need for a holiday...’
 SP1 Ik? Ik ben net terug!
 ‘Me? I’ve just come back!’
 SP1 Nee, ik.
 ‘No, me.’
 [Klaasse 2005:4 (01)]

The first speaker in (1) interprets *je* as a second person pronoun, and assumes that it refers to herself, the addressee. However, the second speaker did not intend to refer to the addressee: he is referring to himself instead.

This kind of miscommunication does probably not occur very often. Usually, it is clear who is meant by *je* for both the speaker and the addressee, in spite of its apparently flexible meaning. The question arises why speakers actually use *je* to refer to themselves, instead of the first person pronoun *ik*. This question is twofold: firstly, we can wonder what makes it possible to use *je* with first person reference, and secondly, we can ask ourselves what makes this possibility attractive for speakers. These questions form the central point in this thesis.

2 From you to me (and back)

In Chapter 2, I will first shortly discuss the original meanings of *je* as a second person and an impersonal pronoun, and next, an overview of the Dutch literature on self-referring *je* will be presented. In Chapter 3, I will examine *je*, and self-referring *je* in particular, in more detail on the basis of corpus data. With my analysis of self-referring *je*, I will review the crosslinguistic data of impersonal and second person pronouns and their possibilities of reference to the speaker in Chapter 4. In Chapter 5, I will round off this thesis by presenting my final conclusions.

Je in the literature

In this chapter, an overview will be presented of Dutch self-referring *je* in the literature. I will discuss several accounts, while focusing, like the authors do themselves, on the speaker's motivation for choosing a second person pronoun to refer to herself. Before I start this overview, however, I will first outline very shortly the original meanings of *je* in section 2.1. In section 2.2, the overview of the Dutch literature will follow, and I will end this chapter in 2.3 with some concluding remarks.

2.1 The original meanings: deictic and generic *je*

In Modern Dutch, *je* is one out of several pronominal forms that are used for reference to the addressee. Besides its second person meaning, it can also be used as an impersonal pronoun. And, as we have seen, it sometimes seems to refer to the speaker. Throughout this thesis, I will call *je* 'deictic' when it has second person reference, 'generic' when it is used as an impersonal pronoun, and 'self-referring' when it refers to the speaker. In the following subsections, I will shortly discuss *je* as a deictic and a generic pronoun.

2.1.1 Je as a second person pronoun

In Modern Dutch as spoken in the Netherlands¹, *je* is the most frequently used pronoun for second person reference. The Dutch second person pronouns can be divided along three dimensions: number, politeness, and emphasis. Following these dimensions, *je*

¹ In Flemish Dutch (the Dutch spoken in Flanders, the non-francophone part of Belgium), the use of second person pronouns is different from the situation in the Netherlands. In this thesis, I will exclusively discuss Netherlandic Dutch.

can be characterised as singular, informal, and weak. Its plural counterpart is *jullie*. The formal second person pronoun is *u*, which can be singular or plural. Finally, its strong counterparts are the subject form *jij* and the oblique form *jou*; *je* itself is not specified for case. Thus, we can conclude that *je* is, in all respects, an unmarked form.

In Dutch, many pronouns have a strong and a weak counterpart². Weak pronouns differ from strong pronouns in their syntactic behaviour, and whereas the former have a reduced vowel (the phonological form of *je* is /jə/), and cannot have stress, the latter usually have a full vowel (the phonological forms of *jij* and *jou* are /jɛi/ and /jʌu/, respectively), can be stressed, and are associated with emphasis. The second person pronouns *u* and *jullie* do not have a weak counterpart.

Finally, *je* can also be used as a second person possessive and reflexive pronoun. As a possessive pronoun, it is again the weak counterpart of a strong pronoun, *jouw*. I will not pay any attention to these usages in this thesis, but *je* will occur as a possessive or reflexive pronoun in several examples in Chapter 2 and 3.

The use of *je* and its strong counterparts is fairly new. According to Vermaas (2004), the forms first occurred in the 17th century as regional, vernacular variants of *ghy*. At that time, *ghy* and its oblique form *u* were the plural counterparts of *du/di*, which could be used as polite forms as well. Initially, *ghy/u* and *jy/jou/je* were semantically equal. That means that *je* could refer to both singular and plural, and that it was a polite counterpart of *du/di*. However, *jij/jou/je* tended to replace *du/di* more and more during the 18th century.

In the 19th century, *du* and *di* had disappeared, and a semantic difference had emerged between *gij/u* on the one hand, and *jij/jou/je* on the other hand. Whereas the former kept their original polite connotation and their singular and plural reference, the latter three

² The status of weak pronouns, which are also called ‘reduced’ or ‘deficient’, is not entirely clear; they may either be analysed as clitics, or as an apart class which is different from both strong pronouns and clitics. This question falls outside the scope of this thesis, but see Cardinaletti & Starke (1999) for details on weak pronouns in Germanic and Romance languages.

were now associated with singular³, and were used for intimate relations. At the same time, the use of *u* in subject position increased, and *gij* disappeared, initially from the spoken language, and in the beginning of the 20th century from the written language as well. In this period, the plural form *jullie* also emerged. Thus, the paradigm as used in Modern Dutch had arisen. In Table 1, the ‘old’, 17th century paradigm, and the ‘new’, 20th century paradigm are given.

Table 1 Second person pronouns in 17th and 20th century Dutch

	17th century		20th century	
	nominative	oblique	nominative	oblique
polite/plural	ghy, jy	u, jou, je	u, jullie	u, jullie
informal/singular	du	di	jij, je	jou, je

During the 20th century, new developments in the second person pronoun system took place; no more forms appeared or disappeared, but the balance between formal and informal pronouns changed. This shift went hand in hand with social developments, like changes in hierarchical relationships. Whereas the polite form *u* used to be the default second person pronoun, this pronoun was more and more associated with social distance, and its informal counterparts, nowadays associated with solidarity, changed into the default form.

2.1.2 *Je as an impersonal pronoun*

In Modern Dutch, many personal pronouns can be used impersonally. However, only *je* and the third person plural pronoun *ze* are common in this use. Besides the impersonal use of originally personal pronouns, Dutch also has the pronoun *men*, which behaves

³ When having the plural form *jullie* as an antecedent, *je* can be used as a plural form, as in (i).

- (i) Dus **jullie** denken dat **je** samen kunt werken?
 ‘So you-PL think that you-SG can work together?’

However, it is unclear whether *je* can also have plural reference ‘on its own’, that is, without having a plural antecedent. Probably, *je* with plural reference is very infrequent.

syntactically like a third person singular pronoun, and which can only refer impersonally. *Je*, *ze*, and *men* can all have generic human reference, that is, all of them can be used in generic statements about human people. This is shown in the following examples.

- (2) In Nederland mag **je** stemmen op je achttiende
'In the Netherlands, **you** are allowed to vote at eighteen.'
- (3) In Brazilië spreken **ze** Portugees.
'In Brazil, **they** speak Portuguese.'
- (4) In Italië eet **men** vaak pasta.
'In Italy, **people** often eat pasta.'

Although all three pronouns can refer generically, there are both semantic and syntactic differences. One of the semantic differences is that *ze* obligatorily excludes the speaker and the hearer, whereas *je* and *men* do not. Another semantic difference between *je* and *ze* will be discussed in Chapter 3.

The most important syntactic difference is that *je* and *ze* both have a complete paradigm with oblique, reflexive, and possessive forms, which can all be used generically as well, whereas *men* is more limited: it can only occur in subject position and does not have an oblique counterpart.

The generic use of *je* is a more recent phenomenon than the use of *men*, but while generic *je* is extremely common nowadays (especially in the spoken language), *men* is marginal in spoken standard Dutch, and even in written language it is not frequent (Weerman 2006). It is generally perceived as highly formal or even archaic. In that, *men* is different from its German and Swedish counterpart *man*, which are both still frequent in spoken and written language. The Norwegian situation, however, is quite similar to the Dutch, in that Norwegian *man* is more formal and less frequently used than impersonally used personal pronouns (Johansson 2004). The English cognate *men/man* already disappeared in the 15th century (Los 2002). Generally, *men* and its cognates in the Germanic languages and in French are not particularly stable (Weerman 2006).

Dutch *men* used to be highly common in the Middle Ages and only started to decrease, particularly in the spoken language, at the end of the 18th century (Paardekooper 1991). At this point in time,

the generic use of *je* started to increase. Initially, it was mainly used in non-subject position. It has therefore been suggested that the advantage of a complete paradigm is the reason, or one of the reasons, why *je* has replaced *men* (Paardekooper 1991; Weerman 2006).

Apparently, the deictic and generic use of *je* emerged roughly at the same time. It must be noted, however, that it has taken much longer before *men* disappeared from written language. Klaasse (2005) has compared the use of impersonal pronouns in 1885-1887 to their use in 1985-2003, on the basis of two corpora of literary novels. In her 19th century corpus, 88% of the impersonal pronouns are instances of *men*, whereas generic *je* does not occur at all. The 20th century corpus shows a dramatically different picture: generic *je* occurs in 57% of the cases, forming the majority of the impersonal pronoun occurrences, while only 17% of them are instances of *men*. Thus, the use of generic *je* must have grown more common during the 20th century.

2.1.3 Summary

In this section, we have seen firstly, that *je* is the most frequently used second person pronoun and the most common impersonal pronoun in modern Dutch, and secondly, that this situation is fairly new. In both usages, *je* is the more informal, unmarked variant. The literature shows that deictic and generic *je* emerged at the same point in time, but also that generic *je* probably became common a bit later than deictic *je*.

2.2 Previous accounts of self-referring *je*

In this section, I will give an overview of the literature on self-referring *je*. As far as I know, the only accounts of this phenomenon have been written in Dutch. This literature is limited in three respects. First, not much is written about this phenomenon. I know of only six publications in which the observation is mentioned, five of which suggesting some possible speakers' motivations to use self-referring *je*, and only one of them referring back to earlier literature on the topic. Second, in none of the papers the conclusions on self-referring *je* are drawn from scientific research on the basis of

empirical data; in fact, four out of the five publications were not written for a linguistically specialised audience. This perhaps explains why it is a rather persisting impression in the Netherlands that self-referring *je* is mainly, or even only, used by soccer players. And third, in none of the papers self-referring *je* is placed in a crosslinguistic context, that is, none of the authors raises the question whether similar phenomena exist in other European languages.

Despite these limitations, it is certainly interesting to pay attention to the Dutch literature. In the following sections, I will discuss the five publications in which a speaker's motivation of self-referring *je* is given; the sixth publication (Van Hout 2003), in which the phenomenon is only observed, I will leave out of consideration. The accounts can be subdivided into three types: self-referring *je* can be a) linked to deictic *je*, b) linked to generic *je*, and c) linked to both deictic and generic *je*. In the following three sections, I will discuss each of these types and their problems.

2.2.1 *Link with deictic je only: speech accommodation*

2.2.1.1 Appel's first account

Appel (1993), in a paper with the indicative title 'Not only soccer players often use *je* when they mean *I*'⁴, relates self-referring *je* to deictic *je*. He argues that the use of self-referring *je* is an instance of speech accommodation (the phenomenon that speakers adapt their own speech to their conversational partner's). Speech accommodation is associated with bilingual speech, but it also occurs in monolingual speech. An example from Appel is shown in (5).

- (5) SP1 Ben **je** in *je* trots gekrenkt?
 'Are **you** hurt in your pride?'
 SP2 Ja, natuurlijk ben **je** in *je* trots gekrenkt.
 'Yes, of course **you** are (I am) hurt in my pride.'
 [Appel 1993:98 (8)]

⁴ Original title: 'Niet alleen voetballers gebruiken vaak 'je' als ze 'ik' bedoelen'.

Appel argues that *je* in the interviewee's utterance is in fact 'elicited' by the interview situation: the deictic *je* from the interviewer question is 'mirrored' in the interviewee's answer. He motivates this by pointing out that the speaker who uttered (5), a well-known Dutch soccer player, would never start a practice session by giving an instruction as given in (6).

- (6) **Je** wil dat jullie nu eerst tien rondjes om het veld lopen...
 'You (I) want you to run ten times in circles around the field first...'
 [Appel 1993:99]

According to Appel, the use of self-referring *je* has recently been increasing. He explains this as a consequence of the increase in frequency of deictic *je* at the cost of its polite counterpart *u*: as interviewers tend to use the informal second person pronoun *je* in their questions more and more, their interviewees will more often mirror this *je* in their answers.

2.2.1.2 The problems

Appel's account meets several problems. At first, he does not consider self-referring *je* to be a strategy to avoid the first person pronoun *ik*. That *je* can refer to the speaker is thus an accidental consequence of the fact that interviewers often use *je* in their questions. Following this argumentation, it is problematic for Appel's account that *u* has, as far as we know, never been 'mirrored' in this type of situation. Although *u* is less frequent today than it was a hundred years ago, it is still a productive pronoun which does not necessarily sound archaic⁵, and which is not uncommon in interviews.

The assumption that deictic *je* has been gaining ground at the cost of *u* for the last decades is uncontroversial; several authors (e.g.,

⁵ *U* may sound archaic in specific situations; as it is unusual that children use *u* when speaking to their parents nowadays, it sounds archaic to most speakers who were born after about 1970 when children do use *u* in this situation. However, it is very common that students in secondary school address their teachers with *u*. For details or more situations see Vermaas (2004).

Appel et al. 1976; Vermaas 2004) have shown this change in balance, which started in the second half of the 20th century, as was pointed out in section 2.1.1. That the frequency of self-referring *je* has been increasing during the same period, however, has never been demonstrated.

2.2.2 Link with generic *je* only: modesty

2.2.2.1 Bennis et al. and Appel

In the remaining accounts, self-referring *je* is explained as a strategy to avoid the first person pronoun *ik* for some reason. In two papers, (false) modesty is argued to motivate this choice. Self-referring *je* is considered to be a variant of generic *je* in both of these accounts. One of the accounts was formulated by Bennis et al. (2004:29), who claim that self-referring *je* is a phenomenon which is typical for the speech of soccer players (*voetbal-Nederlands* ‘soccer-Dutch’). They point out that a substantial part of the Dutch population shares the opinion that it is inappropriate or even indecent to use *ik* frequently, because it sounds egocentric and immodest to talk about oneself all the time. As soccer players talk about their own deeds in public (for radio and television) more often than other people, they have ‘invented’ a ‘bescheidenheidspronomen’ (‘pronoun of modesty’): a strategy to talk about themselves without seeming immodest.

Bennis et al. leave an important question unanswered: what makes *je* such a good candidate to reach the goal of modesty? Appel (1993), in an alternative explanation added to the main account that was discussed in the previous subsection, is a bit more explicit about this. He suggests that the great interest in private, personal issues, which is assumed to be a characteristic of the 80s and 90s of the 20th century, might play a role. These decades are even called the *ik-tijdperk* (‘I-era’), because people are said to talk a lot about themselves in this period. As a reaction to the *more* personal questions that are asked, speakers may feel the need to talk about themselves in a *less* personal way, out of (false) modesty; therefore, they choose a pronoun that is *less* personal. Thus, the hypothesis is that personal questions in a conversation will lead to an increase of self-referring *je*.

Bennis et al. and Appel present quite similar accounts, but they differ in some respects. While they share the assumption that self-referring *je* is a fairly new phenomenon, which is particularly frequent in soccer players' speech, only Bennis et al. claim that it *originates* from soccer players. Appel, in contrast, places the phenomenon in a broader context, providing more insight in the phenomenon. A second difference is that although they both argue that modesty plays a role, only Appel adds the adjective *false* (with some cautiousness though). This difference may seem trivial, but it is in fact quite important, as Appel signals that a speaker using self-referring *je* presents things somewhat differently from how they really are: the speaker *seems* to be modest, but it remains a question whether she actually *is*. In Chapter 3, I will come back to this observation.

To sum up, the modesty accounts share three basic assumptions: 1) the development of self-referring *je* is a recent one, 2) speakers use it in order to avoid *ik*, and 3) this choice is motivated by modesty, either true or false. The third assumption implies that a sentence containing self-referring *je* has a flavour of modesty which its counterpart with *ik* does not have (or at least to a lesser extent).

2.2.2.2 The problems

The main problem with the two accounts above is the assumption that self-referring *je* is a recent phenomenon. It is not easy to find the first instances of self-referring *je*, but I have at least found some instances from the start of the 20th century, which implies that it is by far not as recent as assumed.

At first, many examples were found in letters of, and interviews with, the famous Dutch author Annie M.G. Schmidt, as cited in Van der Zijl's (2004) biography. (7) is a quote from an oral interview in 1955.

- (7) Eerlijk gezegd benauwt 't **me** wel een heel klein beetje. 't Maakt namelijk zo weinig actief, dat buitenleven. **Je** zit maar een beetje te staren voor het raam, en een boekje te lezen, en alles wat in de stad zo verschrikkelijk belangrijk lijkt, is hier helemaal niets. Ik bedoel dit: in de stad zit **je** als maar te denken: o, wat afschuwelijk! Nou heb ik dit nog niet gedaan,

en o, wat vreselijk, nou is dat ook nog niet klaar, maar hierbuiten heb **je** alleen maar zo'n idee van: *waarvoor moet dat nou?*'

'To be honest, it oppresses **me** a bit. Life in the country is quite inactive. **You** are just sitting, staring out of the window, reading a book, and everything that seems terribly important in the city, is nothing here. I mean this: in the city, **you** keep thinking all the time: oh, how horrible! Now I didn't do this, and oh, how dreadful, that is also not finished yet, but here, on the country, **you** only have the idea of *why should I?*' (lit: '*why must that now?*')

[Asser 1955:13, cited in Van der Zijl 2004:225]

An even older example occurs in a letter from Schmidt to her mother, written in 1933 or 1934.

- (8) **We** vloeken nu samen het publiek stijf, 't is zo grappig: op de cursus krijgen **we** prachtige theoretische lessen hoe we met 't publiek om moeten gaan, hoe we ze moeten opvoeden en hun smaak ontwikkelen hoe we de kinderen teeder moeten leiden tot hogere ambities enz, enz. Maar als **je** in de Gorze tegenover die horde vuile communisten staat dan kun **je** al die idealen wel aan de kapstok hangen en alleen maar denken aan zelfverdediging.

'**We** are swearing at the visitors together now, it's so funny: at the course, **we**'re taught all those beautiful things: how we should get on with the visitors, how we should educate them and develop their taste, how we should lead the children tenderly to higher ambitions, etc, etc. But when **you** are in the Gorze, standing face to face with that horde of filthy communists, then **you**'d better throw away all those ideals and only think about self-defence.'

[cited in Van der Zijl 2004:74]

Self-referring *je* was also used in *De HBS-tijd van Joop ter Heul* ('Joop ter Heul's HBS⁶ time'), written in 1919 by Cissy van

⁶ The HBS (Hogere Burgerschool 'Higher Citizen School') was a type of secondary education that existed in the Netherlands until 1968.

Marxveldt. Joop ter Heul was one of the first characters in a book for young girls who used the vernacular of the Dutch upper class schoolgirls (Daniëls 2007). The book consists of (diary) letters that are written in a quite informal style. (9) is an example of self-referring *je* from a letter that was written by the main character about her older sister.

- (9) Pa is twee weken naar Engeland voor zaken, en die Julie doet nu zoo raar. [...] Och, 't kan me eigenlijk niets schelen, maar bij alles zegt ze tegen **mij**: “Kind, waar bemoei je je mee?” En dat klinkt zoo verachtelijk en daar word **je** dan wel eens kriegel van.

‘Dad is making a business trip to England for two weeks, and that Julie is behaving so strange now. [...] Well, I don’t care actually, but she keeps saying to **me**: “Kid, mind your own business.” And that sounds so disdainful, and it makes **you** feel touchy sometimes.’

[Van Marxveldt 1919:3]

In (8), we already saw that self-referring *je* sometimes refers to ‘we’ rather than to ‘I’. We see a clearer example of this in (10), which was taken from a diary letter.

- (10) Georgien heeft een leuke wijs gemaakt op mijn clublied; we repeteeren het nu voor schooltijd in het park, maar dan gaan **we** al om kwart over acht van huis, anders zijn er zooveel kantoorlui, die **je** uitlachen.

‘Georgien has composed a nice melody for my club song; we rehearse it in the park before school now, but then **we** leave home at a quarter past eight already, for otherwise, there are a lot of office workers who make fun of **you**.’

[Van Marxveldt 1919:221]

Obviously, *je* in (10) refers to ‘we’, the girls who are rehearsing their club song together. This is an indication that *je* is not simply a way to avoid the use of *ik*.

The examples above show that self-referring *je* was already used in the first part of the 20th century. This is problematic for both accounts that were described above. In 1919, soccer players were

not interviewed for radio and television yet, which makes it unlikely that they invented this use of the pronoun. Self-referring *je* also occurred far before the so-called ‘I-era’.

Apart from the problem of time, the explanations of (false) modesty are problematic too. In none of the examples above, *je* gives the fragment a flavour of modesty, true or false. Thus, we can conclude that neither of the accounts is satisfying.

2.2.3 Link with both deictic and generic *je*: empathy

The three remaining authors are Van Wassenaar (1994), Onrust (1995), and Klaasse (2004). They share the intuition that the use of self-referring *je* can be motivated by the speaker’s need for empathy. The reason why *je* is suitable for this need is that it can appeal to the addressee’s willingness to imagine herself into the situation of the speaker. However, the accounts also show some differences.

At first, Van Wassenaar and Klaasse, like Appel, observe that self-referring *je* is not only used by soccer players or sportsmen. Onrust, however, considers self-referring *je* to be exclusive for the speech of soccer players, in particular for Johan Crujff, who is assumed to be a pioneer in this respect. In the light of the previous subsection, we know that Crujff cannot possibly be the inventor of self-referring *je*, since he was born in 1947. Onrust also discusses the idea of *je* as a strategy for modesty. However, she rejects this as an explanation of ‘voetbal-*je*’ (‘soccer player *je*’), suggesting that soccer players, and especially Crujff, are not exactly known for their modesty⁷.

A second difference is that each of the three authors explains the link with generic and deictic *je* in a different way. I will discuss the way they work out this link and the problems that their accounts face separately in the following subsections.

⁷ Moreover, Onrust (1995: 14) refers to a dissertation written by J.E. Bergsma, who argues that the increasing use of *je* is caused by fear of personal responsibility. Unfortunately, this reference is incomplete and I did not succeed in recovering this publication.

2.2.3.1 Van Wassenaar

Van Wassenaar's account is rather untransparent, as he does not use linguistic terms⁸. He does mention the possibility of *je* having generic reference, but the link with the motivation for self-referring *je* he proposes is unclear. He focuses on the need for empathy only, arguing that the speaker actually wants to say 'please imagine that you were in my situation'. The speaker is thus assumed to use *je* in order to project herself on the addressee and *je* is more deictic than generic then. This is typically done when the speaker feels the need to create some distance between herself and the utterance, particularly in the case of personal failure. Note that in such a situation, the choice for *je* instead of *ik* is motivated by the opposite of modesty: whereas a modest speaker is supposed to tone down something positive, a speaker creating distance because of failure will rather try to conceal something negative.

The idea of self-referring *je* as a means of inviting the addressee to feel empathy for the speaker is interesting. Yet, it is not completely clear how this works: why can a generic pronoun, which often refers to all humanity, appeal to the addressee in this way? Van Wassenaar does not specify this mechanism. A second problem is the situation of personal failure. Indeed, self-referring *je* can occur (and maybe even often occurs) in this kind of situation, but this is not necessarily the case. An example is (11), which is Cruijff's answer to the question if he found himself a genius.

- (11) Daar sta **je** nooit bij stil. **Je** deed wat **je** moest doen.
 'That never occurs to **you**. **You** did what **you** had to do.'
 [Appel 1993:98 (9)]

(11) is obviously not an example of personal failure; on the contrary, the speaker was apparently found to be brilliant. Thus, although personal failure might be explanative sometimes, it cannot be the whole explanation.

⁸ Van Wassenaar is, as far as I can see, not a linguist. Yet, his paper was published in a (non-scientific) periodical and it is referred to by Klaasse (2005). Besides, his intuitions are interesting enough to discuss here.

2.2.3.2 Onrust

Onrust's account is unique in that she makes a subdivision in self-referring *je*. She argues that it can have two different sources: it can either be an extension of deictic *je* or an extension of generic *je*. In case of deictic *je*, the speaker appeals to the imagination of the addressee to evoke empathic feelings. This is illustrated in (12), which is a soccer player's answer to the question why he chose the Turkish club Istanbulspor instead of the Dutch club Ajax. In the case of the generic pronoun, *je* is used to generalise from a specific situation, changing it into a general pattern. This is shown in (13), in which a soccer player is telling his interviewer about his tasks in his new club.

(12) **Je** bent 26 en **je** wilt graag spelen, en **je** hoopt dan een basisplaats bij Ajax maar als **je** daar het afgelopen jaar niet in slaagt dan heb **je** toch gekozen voor Istanbulspor...

'**You** are 26, and **you** really want to play, and **you** hope for a permanent position at Ajax, but if/when you did not succeed last year, then **you** have chosen Istanbulspor instead...'

(13) Het is de bedoeling dat **je** wat meer balans in je elftal brengt, en dat **je** ervoor zorgt dat het elftal als team geheel beter gaat voetballen.

'The purpose is that **you** bring some more balance in your team, and that **you** make it play soccer better as a whole.'

[both examples Onrust 1995:13]

Onrust judges both forms of self-referring *je* to be inappropriate: in the more deictic case, as illustrated in (12), the speaker adds so many concrete details that the addressee cannot imagine to be in a similar situation anymore, and in the more generic case, as shown in (13), the speaker's situation is unique and therefore not generalisable. Of course, we could also summarise that in both cases, the situation of the speaker is too specific to use *je*. However, it is a remarkable conclusion that *je* 'cannot be used this way', as it is clear that it is actually used like this quite frequently. Moreover, I do not share the intuition that the *je* tokens in (12) are fundamentally different from those in (13), and none of the other authors has made such a distinction either. To me, the instances of *je* in both (12) and

(13) have the flavour of a generalisation, which makes it unlikely that the tokens in (12) are instances of deictic *je*. Again, the phenomenon is not explained in a satisfying way.

2.2.3.3 Klaasse

Finally, I discuss Klaasse's account, which is more formal than the other ones. As said in section 2.1, Klaasse compares the use of impersonal pronouns in the end of the 19th century to their use in the end of the 20th century. We saw that generic *je* did not occur at all in her first corpus. Moreover, self-referring *je* was also absent. The fact that self-referring *je* was not found in the corpus in which generic *je* does not occur either indicates that it may well be an instance of generic *je*. However, her corpora do not represent spoken language, so we cannot draw strong conclusions from her findings.

Klaasse is the only Dutch author who refers to earlier publications on self-referring *je*, and she agrees with Van Wassenaar and Onrust that speakers use self-referring *je* in order to create involvement, out of a need for empathy. Therefore, she calls self-referring *je* 'inlevings-*je*' (*inleving* meaning 'imagining yourself to be in the situation of someone else', which is the essence of empathy). However, Klaasse does not link this to deictic *je*, but to generic *je* instead:

Door *je* te gebruiken in zinnen die op de spreker zelf van toepassing zijn, wordt de schijn van algemeenheid gewekt. De spreker suggereert dat wat voor hemzelf geldt, voor iedereen in die situatie gegolden zou hebben. Degene die wordt aangesproken wordt hierdoor meer betrokken in het verhaal. Dit effect kan door de spreker gebruikt worden om zijn verhaal interessanter, indrukwekkender en levendiger te maken.

'By using *je* in utterances that apply to the speaker, an appearance of generality is created. The speaker suggests that what goes for him, would go for anyone else who was in the same situation. The effect is that the addressee feels more involved in the story. The speaker can use this strategy to make his story more interesting, more impressive, and more lively.' [Klaasse 2005:6]

Klaasse argues that both *men* and *je* can be placed on a continuum of generality: the two extremes are super-generic and super-specific. The idea of a continuum of genericity is not unique: it has often been observed that generic pronouns vary in their degree of generality (e.g., Ramat & Sansò 2007). However, as far as I know, Klaasse is unique in placing deictic *je* on this continuum as well. In its super-generic use, *je* refers to human people in general, as in (14).

- (14) **Je** gaat een keer dood.
'Some day **you** will die.' (lit: '**You** go dead one time')
[Klaasse 2005:19 (38)]

The more limited the subgroup of human beings, the less generic *je* is. In its super-specific use, Klaasse assumes *je* to refer to the addressee, as in (15).

- (15) **Je** staat op mijn teen.
'**You** are standing on my toe.'
[Klaasse 2005:19 (42)]

Klaasse places self-referring *je* one step before deictic *je*, arguing that it is rather specific, but less specific than addressee-referring *je*.

The idea of a continuum is attractive, because generic pronouns like *je* indeed have many degrees of generality. However, there is no reason to assume that reference to the speaker is a more general kind of reference than reference to the addressee, which is implied by Klaasse's account. Placing deictic *je* on the continuum of generality is thus problematic again.

2.2.4 Conclusions on the literature

We have seen that in the Dutch literature, three types of explanations of self-referring *je* have been suggested. According to Appel, the use of self-referring *je* was caused by the increasing use of *je* in its original, deictic meaning, as an effect of speech accommodation. However, we have no indications that *u* can be self-referring as well, and there is no satisfying answer to the question why exactly *je* would be mirrored.

The second type of explanation is the idea that generic *je* can be used as a pronoun of (false) modesty, either as a consequence of talking about oneself in public often (Bennis et al.), or as a reaction on the unbridled interest for the individual (Appel). Both accounts assume that self-referring *je* is a recent development. This assumption turned out to be false. Instances of self-referring *je* from the early 20th century also rule out the possibility that soccer players initiated this usage. Moreover, modesty was found to be unsatisfying as an explaining factor for the use of self-referring *je*.

The third intuition, suggested by three different authors, is that self-referring *je* is motivated by the speaker's need for empathy. Van Wassenaar points out that it often occurs in situations of personal failure, in which the speaker wishes to create distance between herself and the message. Although this idea is interesting, we have seen that self-referring *je* does not only occur in situations of personal failure. Onrust, who ascribed the use of self-referring *je* to soccer players only, suggests a two-way distinction between more deictic and more generic self-referring *je*. However, no evidence was found for a fundamental distinction between the two. Eventually, the continuum of generality suggested by Klaasse was discussed. This representation turned out to be problematic because it implies that reference to the speaker is less specific than reference to the addressee.

2.3 General conclusions

I started this chapter with a short outline of *je* as a deictic and a generic pronoun. In both usages, the form *je* is relatively recent, but yet, it is far more frequent than its respective counterparts *jij*, *u*, and *men*. Next, an overview of the Dutch literature on self-referring *je* followed. It turned out that self-referring *je* has a mysterious status in the Dutch literature in several respects: its origin and usage (soccer-players or more general), its relation to deictic and generic *je*, and why speakers use it (modesty, empathy, or generalisation). It was argued in the start of this chapter that the lack of empirical data and a crosslinguistic comparison in the Dutch literature are serious limitations of the research on this topic. In the next chapter, I will discuss my own study, which was based on empirical data.

Je in spontaneous speech: a corpus study

We have seen in the previous chapter that the data collection was one of the main problems in the Dutch literature on self-referring *je*: all conclusions were based on a combination of constructed sentences and utterances from a personal collection of relevant utterances. In order to find out what is really happening in Dutch, a more systematic way of collecting data is of major importance. I therefore examined data from the Spoken Dutch Corpus (Corpus Gesproken Nederlands, henceforth CGN). In this chapter, I will discuss this corpus study in detail.

The chapter is structured as follows. In section 3.1, I will present how I collected my data. The cues that determine our interpretation of *je* as addressee-referring, generic or self-referring will be discussed in section 3.2. In section 3.3, self-referring *je* will be examined in more detail. I will show that self-referring *je* is in fact generic *je*. This relation will be studied in depth in section 3.4. In this section, I will also suggest a possible speaker's motivation for the use of self-referring *je*. Finally, I will show in section 3.5 that generic and self-referring *je* still have traces of their deictic origin. This will be the last step to my final account of self-referring *je*. Concluding remarks will be made in section 3.6.

3.1 Data collection

The CGN is a database of Dutch speech, collected between 1998 and 2004, and it consists of about nine million words. Both Netherlandic and Flemish Dutch are represented, but as pointed out before, I excluded Flemish Dutch from my study.

The CGN has fifteen components that represent different speech situations, varying in their degree of spontaneity from natural

conversations to read speech. As I aimed to examine natural speech, it was necessary to select only those components that provide the lowest chance of utterances that were prepared before the start of recording. These components are: spontaneous face-to-face conversations, interviews with teachers of Dutch, and telephone dialogues⁹. From these components, I extracted two subcorpora which I will call ‘Corpus I’ and ‘Corpus II’ throughout this thesis. All examples from the CGN that I give are recoverable as follows. At first, the subcorpus is given, with ‘C.I.’ referring to Corpus I, and ‘C.II’ to Corpus II. This is followed by the component: ‘CON’ for conversations, ‘INT’ for interviews with teachers, and ‘TEL’ for telephone dialogues. Finally, the CGN code is given. Thus, for example, [C.II, CON, fn000541.298] refers to an utterance from a conversation in Corpus II, which is represented in the CGN as ‘fn000541.298’.

In the following two sections, I will describe my method of setting up both corpora.

3.1.1 Corpus I. Teachers of Dutch

Initially, I intended to focus on non-deictic *je* only and I therefore decided to focus on the component in which the amount of deictic *je* was least likely to be an overwhelming majority: the interviews with teachers of Dutch. The CGN contains unstructured interviews with 80 teachers of Dutch from the Netherlands, each having a duration of about 20 minutes. I divided the interviews in eight blocks of ten speakers, following the order as represented in the CGN. In five blocks, men and women were equally divided, and I randomly took one of these blocks¹⁰.

From this selection of speakers, all nominative forms of the personal pronoun *je* were taken, which were 502 tokens. As the interviewer utterances were expected to contain mainly deictic *je*, I eliminated all interviewer tokens from the data. Next, the remaining

⁹ The CGN contains telephone dialogues recorded via switchboard and on MD with a local interface. I included both types of recording in my data.

¹⁰ This block consists of the following speakers: 139, 140, 186, 187, 188, 189, 190, 206, 207, and 209. One of the speakers who were labeled as ‘male’ in the CGN turned out to be a woman, and therefore, speaker 188 was replaced by a male speaker of the same age, speaker 086.

unambiguously deictic tokens were eliminated as well. As I assumed the context to be highly informative, two extra criteria were formulated in order to avoid misinterpretations: a) *je* is part of an utterance that is complete insofar that it has a main verb, and if this main verb is transitive, the complement is present, and b) the utterance is interpretable. In my final selection, I had 263 tokens for further examination.

In the course of the study, I realised that I had focused on a very restricted subset of speakers (namely, only teachers of Dutch), which might influence the representativity. Therefore, I constructed a new, more representative corpus: Corpus II.

3.1.2 Corpus II. Spontaneous speech

For Corpus II, I took all tokens of the personal pronoun *je* from the components that represent spontaneous speech, as listed in section 3.1. The total number of *je* tokens was 56796: 54422 tokens having nominative case, and 2374 tokens having oblique case. I randomly drew 1% from these tokens, evenly distributed over the four components, which resulted in a subcorpus of 569 tokens. I maintained the two criteria formulated in the previous subsection. Moreover, I eliminated five tokens that were incorrectly annotated as a personal pronoun (being a possessive pronoun or a diminutive suffix, which both have the same form as the personal pronoun) and one token that was part of read (that is, non-spontaneous) speech. The elimination round resulted in a final subcorpus of 467 tokens.

At first, all tokens were again annotated for their basic reference. I divided reference of *je* into three basic categories: *deictic* (referring to addressee), *non-deictic* (not referring to the addressee), and *ambiguous* (both interpretations possible). The non-deictic category thus contains both generic and self-referring *je*. This may seem strange, as first-person pronouns are deictic too, which suggests that self-referring *je* is deictic. However, in the literature on impersonally used personal pronouns, the term ‘deictic’ is often used to distinguish addressee-referring *je* from generic *je*. Moreover, I will argue later on that self-referring *je* is, strictly speaking, an instantiation of generic *je*, and therefore not deictic.

The results of this annotation are represented in Table 2.

Table 2 Basic reference of *je* in Corpus II

reference	absolute	relative
deictic	247	53%
non-deictic	183	39%
ambiguous	37	8%
total	467	100%

It is noteworthy that the originally second person pronoun *je* refers unambiguously to the addressee in only half of the cases. Moreover, among the *je* tokens are some more or less fixed expressions. I did not take them as a distinct category, because it is often difficult to say whether an expression is fixed or not. Included in the deictic *je* category were clearly fixed expressions like *dank je (wel)* ('thank you'), which occurred 2 times, *weet je (wel)* ('you know'), which occurred 28 times, and *zie je* ('you see'), which occurred 4 times. In the category of non-deictic tokens, the majority of the expressions that could be considered as fixed occurred only once, e.g., *en voor je 't weet* ('and before you know'), *als je niet beter wist* ('if you didn't know better'). In section 3.2.2, special attention will go to the fixed expression *je hebt...* ('you have'), which occurred 17 times and seems to be used as an existential construction in which the direct object is the NP in focus.

3.1.3 Summary

From the CGN, two subcorpora were extracted: Corpus I, which consists of only interviewee utterances of teachers of Dutch with *je* tokens that do not refer to the addressee, and Corpus II, including three situations of spontaneous speech, in which no tokens were eliminated on the basis of their reference beforehand, thus including deictic tokens as well. In the remainder of this chapter, I will cite from both corpora and I will not go into any possible differences between them.

3.2 The correct interpretation of *je*

The first question that needs to be answered is how we actually know whether *je* should be interpreted deictically or generically. Of course, the pronoun itself does not give us a cue for the correct

interpretation, as it does not have something like a marker for genericity or deixis. The finite verb does not have such a marker either, nor does the sentence. In this section, I will discuss the cues that lead listeners to the correct basic interpretation of *je*.

3.2.1 A grammatical cue: episodicity

One cue for a deictic interpretation of *je* may be found in the structure of the sentence. Several semanticists, e.g. Malamud (2006), have pointed out that second person pronouns cannot refer generically in episodic sentences. Thus, it is predicted that *you* can only refer to the addressee in (16a), and the same goes for its Dutch counterpart in (16b).

(16a) **You** burned a house just now. [Malamud 2006:10 (17b)]

(16b) **Je** hebt net een huis in brand gestoken.

‘**You** set fire to a house just now.’

However, a generic interpretation of *je* is perfectly possible when (16b) is part of a conditional, whether in the condition, as in (17a), or in the consequent, as in (17b).

(17a) Als **je** net een huis in brand hebt gestoken, dan voel **je** je voldaan.

‘If **you** have just set fire to a house, **you** feel satisfied.’

(17b) Altijd als **je** je voldaan voelt, dan heb **je** net een huis in brand gestoken.

‘Always when **you** feel satisfied, **you** have just set fire to a house.’

It may be hard to imagine a world in which people usually feel satisfied if (as in (17a)), and even only if (as in (17b)), they have just burned down a house, but this makes the sentence only pragmatically odd. Of course, (17a) and (17b) are not episodic sentences any longer, and we are not forced to reject the prediction above. However, it emphasises the role of pragmatics, because an utterance can be implicitly or partly a conditional, that is, without the characteristic *if* or *when* and *then*. This is shown in (18).

- (18) Oh dus **je** hebt nog helemaal niet gegeten?
'Oh, so **you** haven't eaten anything yet?'
[C.II, TEL, fn008265.274]

If the context is absent, we cannot see if this sentence is episodic, even though it has perfective aspect, which often indicates episodicity. Implemented in two different imaginary contexts, we can interpret *je* either deictically or generically.

- (19a) SP1 Ik ben vanaf vijf uur alleen maar aan het rennen
geweest dus nu heb ik
verschrikkelijk veel honger.
'I have been busy all the time since five, so I am
terribly hungry now'.
SP2 Oh dus **je** hebt nog helemaal niet gegeten?
(19b)SP1 (instructing speaker 2, who will be a participant in an
experiment)
Na deze taak hebben de meeste mensen altijd erge
honger, want je moet er meteen aan beginnen als je
wakker wordt.
'People are always very hungry after this task, as you
have to start it immediately after waking up.'
SP2 Oh dus **je** hebt nog helemaal niet gegeten?

The final sentence in (19b) has exactly the same form as the final sentence in (19a), which is episodic. In (19b), however, it is implicitly part of the conditional 'when you start this task, you haven't eaten anything yet', which is not episodic of course. It is left to the hearer to infer from the pragmatic context that this is the case.

Note that Malamud's prediction does not work the other way around; that is, *je* does not automatically have a generic interpretation in non-episodic sentences, as shown in (20).

- (20) Maar als **je** 't niet meer wil dan kan **je** 't ook zeggen.
'But if **you** don't want it anymore, **you** can say it.'
[C.II, TEL, fn008102.423]

After reading (20) in this isolated position, it is impossible to say whether the two instances of *je* in (20) are to be interpreted

deictically or generically. However, the context in the CGN makes clear that it is intended as deictic: the speaker has offered to copy a CD for the hearer and in (20), the speaker makes sure that she is really doing the hearer a favour, by giving her the chance to politely decline this offer. Thus, both *je* tokens in (20) refer to the addressee.

3.2.2 *The role of context*

We have seen that episodicity can be a cue for the correct interpretation of *je*. However, it turned out that it is not the only cue, as the absence of episodicity does not tell us anything, and as it is often unclear whether a sentence is episodic or not. What, then, are the cues which tell us more about the right interpretation? I argue that the context is decisive.

At first, it is important to realise that we do not tell our hearers much about themselves, as they often know themselves better than we do – at least, that is what we assume. This is illustrated in (21).

- (21) ja 't is een studentenstad ja dat merkte **je** goed.
'it is a college town, that was obvious'
or: 'it is a college town, that was easy to see'
(lit: '**you** noticed that well')
[C.II, TEL, fn006776.148]

Normally, people will tell their conversational partner that they have noticed something themselves; it is odd to tell her that she noticed something, for she is supposed to know that better than they do. Indeed, the *je* in (21) does not refer to the addressee. Instead, the speaker is telling about his own experiences in an excursion to Heidelberg, Germany, and his nights out there. (21) is the answer to his conversational partner's question if the night life in Heidelberg is fun. Thus, *je* refers to a generic person or perhaps even merely to the speaker himself.

As we usually do not tell people things about themselves, it is probably not a coincidence that about half of the deictic tokens are part of a question, whereas only 5% of the generic tokens are, as shown in Table 3.

Table 3 Deictic and generic tokens and sentence type

sentence type	deictic		non-deictic	
	absolute	relative	absolute	relative
questions	110	45%	10	5%
non-questions	137	55%	173	95%
total	247	100%	183	100%

Of course, this is nothing more than an (admittedly strong) tendency: as Table 2 shows, it is perfectly possible to use deictic *je* in utterances which are not a question, and the same goes for the use of generic *je* in questions (although this seems to be quite rare).

The determining factors in the interpretation of *je* are knowledge of the world and pragmatic principles in combination with the context. I give some examples of a construction that seems to have grown into a fixed expression: *je hebt* (or, through inversion, *heb je*), ‘you have’. As said before, this construction occurs 17 times in my subcorpus and in fact, it functions as an existential construction, as shown below. I will discuss how we come to this interpretation for each of the following sentences.

- (22) ja nou **je** hebt meerdere Eries hè.
 ‘Well, **you** have more than one Erie, you know.’
 [C.II, TEL, fn006821.106]
- (23) ah nu heb **je** echt van die achtmetercampers.
 ‘Ah, now **you** have those eight meters long campers.’
 [C.II, TEL, fn006752.110]
- (24) maar **je** hebt toch van die van die van die van die hardhouten drempels hè.
 ‘But **you** have those hard wooden thresholds, haven’t you?’
 [C.II, CON, fn000303.219]

Even without a context, it is clear that the speaker of (22) is not securing that the addressee has more than one Erie, as Erie is the name of several cities in the United States. He could as well have been uttered (22’).

- (22’) Ja nou er bestaan meerdere Eries hè.
 ‘Well, more than one Erie exists’.

I will not go into the reason why the speaker chose *je hebt* instead of the original existential construction *er zijn*; my point is that *je hebt* can be used in a context in which literally having something is impossible, and as such, it can lose its original meaning completely.

As for (23), it is not impossible that someone has several campers with a length of eight meters, but it is at least unusual. Moreover, even if someone had those campers, it would be unnecessary to tell her so, as it is very unlikely that someone has several large campers without knowing. The only situation that I can imagine which renders a deictic reading of *je* in (23) is that it is a reaction of astonishment to the conversational partner's utterance in which he tells that he was in the surprising position of getting some campers. However, the actual context from which (23) was taken is simply a conversation about the difference between the first campers and the much longer campers that are made nowadays. Earlier in the conversation, the addressee told the speaker of (23) that she would like to have a small camper, which implies that she does not have any, large or small. This rules out a deictic reading of *je*, and the same goes for a purely generic reading in which it is typical for humans to have an eight meters long camper. Again, *je hebt* (or the inverted order *heb je*) is an existential construction.

In contrast to (22) and (23), a deictic reading of *je hebt* in (24) is not unlikely without context: the speaker might be checking the truth of her belief that the addressee has hard wooden thresholds. A purely generic reading is implausible because we know that hard wooden thresholds are not one of those objects that the typical person owns. However, the context shows that *je hebt* again functions as an existential construction, as the two conversational partners live together and thus have their thresholds together as well, and as they discuss what kind of thresholds they would like to have, not having any hard wooden thresholds at the time of conversation.

3.2.3 *Self-referring je and context*

As we have seen in the previous two subsections, the interpretation of *je* is highly context-dependent. The question arises whether context is also (mainly) decisive in the case of self-referring *je*. I argue that this is indeed the case. An example from Appel (1993), which I cited in section 2.2.1.1, illustrates this.

- (25) Natuurlijk ben **je** in je trots gekrenkt.
'Of course **you** are hurt in your pride.'

Without any context, the reference of *je* in (25) is completely unclear. I give two imaginary contexts in (26a) and (26b).

- (26a) SP1 **Ik** voelde me zo...
'**I** felt so...'
SP2 In je trots gekrenkt?
'... being hurt in your pride?'
SP1 Nee, helemaal niet!
'No, not at all!'
SP2 Natuurlijk ben **je** in je trots gekrenkt. Dat zie ik duidelijk.
'Of course **you** are hurt in your pride. That is obvious to me.'
- (26b) SP1 ... en toen lachten ze hem allemaal uit, dus **hij** voelde zich natuurlijk verschrikkelijk gekleineerd.
'... and then they all laughed at him, so **he** felt terribly disparaged of course.'
SP2 Het lijkt me echt vreselijk om dat mee te maken: **je** denkt dat **je** iedereen zal overtuigen en in plaats daarvan krijg **je** zo'n reactie.
'It must be horrible to go through: **you** believe that **you** will persuade everyone and then **you** meet such a reaction.'
SP1 Precies, het maakt **je** onzeker, en natuurlijk ben **je** in je trots gekrenkt, als **je** zo wordt uitgelachen.
'Exactly, it makes **you** feel insecure, and of course **you** are hurt in your pride, when people make fun of **you** that way.'

In (26a), the speakers are talking about the feelings the first speaker has. Therefore, it is clear that *je* in the second speaker's utterance refers deictically. In (26b), the second speaker introduces generic *je* to generalise a specific experience of an unspecified male person. The first speaker extends this generalisation and also uses generic *je*.

However, in the ‘real’ context as cited by Appel, the speaker of (25) is asked a question about his particular feelings, as was already shown in (5), repeated as (26c).

- (26c) SP1 Ben **je** in je trots gekrenkt?
‘Are you hurt in your pride?’
SP2 Ja, natuurlijk ben **je** in je trots gekrenkt.
‘Yes, of course are you hurt in your pride.’

The main reason why the reader believes that the speaker says that he himself is hurt in his pride, is the fact that this is the only possible interpretation of an answer to the preceding question, as the first speaker explicitly asks something about the second speaker’s feelings, using a personal pronoun with second person reference. Again, the context determines how *je* should be interpreted, whereas the utterance itself does not give us a cue.

3.2.4 Conclusions

As we have seen, the interpretation of *je* is highly flexible and hearers use several kinds of cues to find out whether *je* refers to the second, generic, or first person, sometimes from the grammatical structure of the sentence, but most of all, from the semantics and pragmatics in the context.

In the following section, I will discuss self-referring *je* in more detail, elaborating on the idea of the context-dependent nature of *je* in general.

3.3 Self-referring *je*

After annotating the tokens from Corpus II for basic reference, I intended to make a more precise annotation, which was meant to distinguish the self-referring *je* tokens from those tokens that are not self-referring. This would make a comparison between the two types of *je* possible. I defined self-referring *je* as follows: the speaker is the referent in the first place. However, this annotation turned out to be highly problematic: I was not able to formulate strict conditions that a token must meet to be annotated as self-referring. And although some tokens seemed clearly self-referring to me, it was

impossible to make a clear distinction between the clear cases and the less clear cases.

The main problem with self-referring *je* is, that *je* is not simply a new word meaning ‘I’. It is a word which has had its own meaning since long (having deictic reference to the addressee), and which has, at a certain point in history, started to function as a generic pronoun as well. Its ‘new’ function as a first-person-referring pronoun must be related in some way to one or both of these original meanings. This indeed turned out to be the case, but unlike the case of deictic and generic reference, the distinction between self-referring and non-self-referring generic *je* is highly subjective and not clear-cut at all. This problematic status will be illustrated here.

3.3.1 *Explicit cues: questions and pronoun switches*

We have seen in section 3.2.3 that a possible cue for the listener to interpret *je* as self-referring is an interviewer question for specific information about the interviewee. This mechanism also occurs in my data, as shown in (27). The interviewee in (27) is a teacher of Dutch as a second language for immigrants, and the speakers are discussing the prejudices of Dutch people towards immigrants and vice versa.

- (27) SP1 is **je** eigen beeld van Nederland veranderd door wat **je** van buitenlanders over Nederland hoort?
 ‘Has **your** own image of the Netherlands changed by what **you** hear foreigners say about it?’
- SP2 **mijn** beeld van Nederland veranderd?
 ‘**My** image of the Netherlands changed?’
- SP1 Ja
 ‘Yes.’
- SP2 **mijn** beeld van Nederland veranderd oei dat is moeilijk nou aan de ene kant leer **je** dan door je cursisten dat er veel Nederlanders zijn die haat hebben voor deze mensen aan de andere kant ben **ik** in de afgelopen periode eigenlijk het laatste jaar toch ook verschillende keren meningen tegengekomen en dat is dan van de

mensen die geen persoonlijk contact hebben waarvan **ik** schrik.

‘**My** image of the Netherlands changed? Gosh, that is a hard question. Well, on the one hand students tell **you** that many Dutch people have feelings of hatred against these people [immigrants]. On the other hand, last period or actually last year, **I** met some opinions of people who are not in personal contact [with immigrants] that really shock **me**.’

[C.I, INT, fn000190.79 - fn000190.89]

In (27), the interviewer asks a question about the interviewee’s particular experiences. It is obvious that this question is interpreted correctly by the interviewee: she repeats it twice and the second time she even stresses the possessive pronoun *mijn*. Yet, she initially uses *je* in her answer. Although we know that she is probably talking about herself, the utterance is presented as a generic statement, which is not only indicated by the use of a generic pronoun, but also by the use of the present tense, the particle *dan* ‘then’, and the vague verb *leren* ‘learn’. A second indication that the speaker is talking about her own experiences is the structure of the sentence (*aan de ene kant... aan de andere kant* ‘on the one hand... on the other hand’) in combination with a pronoun switch to *ik* ‘I’ in the second part of the answer. This part tells us about the speaker’s experiences more specifically than the first part. In sum, contextual cues tell us that the *content* of (27) is specific, whereas the first part is *presented* as a generalisation.

In (27), the interviewer question is the most obvious cue for the hearer to interpret *je* as self-referring. However, a pronoun switch only can be an indication of *je* being clearly self-referring as well. This is shown in the following sentences; in (28), we see a switch from *je* to a first person pronoun, and a switch vice versa is shown in (29).

- (28) want **je** zit toch wel 'ns zo rond te kijken als ze aan 't werk zijn dan denk **ik** goh 't is eigenlijk wel leuk hè die kinderen zo
'For sometimes **you** are looking around when they are working, and then **I** think: gosh, it is quite fun actually, those children here.'
[C.I, INT, fn000086.53]
- (29) maar dat zijn wel schrijvers die **me** interesseren en die **je** blijft volgen in hun ontwikkeling.
'But those are authors who interest **me** and whose development **you** keep following.'
[C.I, INT, fn000139.274]

(28) is uttered by a teacher of Dutch on a secondary school, who is talking about some negative experiences in his job, which he opposes to the situation sketched in (28). The structure of this sentence indicates that *je* refers to the speaker: *dan* 'then' refers to the situation of looking around when the pupils are working, and as the speaker is the one who has these specific thoughts in that situation, it must also be the speaker himself who is looking around.

In (29), we see a switch in the reverse direction: from a first person pronoun to *je*. This speaker is talking about two books, written by different authors, that he recently read. The first part of the sentence is presented as a personal experience: the first person pronoun *me* 'me' is used. This presentation suggests that the second part of the sentence is not a genuine generalisation, as it is unlikely that people in general, or even teachers in general, would keep following the authors whom this specific speaker is interested in.

3.3.2 *Implicit cues: probability and recognition*

The fact that it is sometimes unlikely that *je* refers to people in general turns out to be the central point in the identification of self-referring *je*. We have seen that there are some explicit cues that make us be quite certain of the assumption that *je* refers to the speaker: an interviewer question with specific reference to the addressee, and a pronoun switch from first to second person or vice versa. However, what settles the matter for us as a hearer or reader is that we suppose the situation, feeling or thought expressed to be

too specific for being a real generalisation. This is illustrated in the following example, in which no explicit cue is present. The utterance was taken from a conversation about the film version of Tolkien's *The Lord of the Rings*, which was shot in New Zealand. The speaker tells that she has since long wished to travel to New Zealand one day and then adds (30).

- (30) en als **je** dan weet dat 't daar opgenomen is dan kijk **je** toch even met extra ogen van ja 't ziet er toch wel oogverblindend uit.
'And if **you** know that it was shot there, **you** will look from a different view, like, well, it does look dazzling, really.'¹¹
[C.II, CON, fn007997.160]

The occurrences of *je* are self-referring insofar that the speaker seems to talk about her own experience: since *she* wants to see New Zealand, and as *she* knows that the movie in question was shot there, *she* was all eyes while watching it. However, the sentence is unquestionably presented as a generic statement: the literal claim is that anyone having this particular information would be all eyes while watching this movie. Generic *je* perfectly fits in such a statement.

In (30), the speaker switches from describing a specific event in the past tense to a conditional in the present tense. The combination of conditional and present tense makes the generalising force of the utterance explicit. The same mechanism can be less explicit, as is the case in the next example. The speaker of (31) is a teacher of Dutch as a second language for immigrants, and she is talking about the influence of cultural differences on the situation in class. She illustrates this with something that had happened earlier that day. She was teaching her students how to ask their conversational partner some questions about her personal situation (having children, living together, being married), and when one of them asked her these questions, she answered that she has children and

¹¹ The speaker confuses the expression *met andere ogen kijken* 'look from a different view' (literally: 'look with different eyes') with *met extra ogen kijken* (literally: 'look with extra eyes'). This results in the nicely striking imaginary picture of someone being all eyes, which is exactly what the speaker is saying.

that she is not married. Apparently, the realisation that the teacher was an unmarried mother came as a shock for the student, which the speaker expresses by uttering (31).

- (31) en **je** ziet het gezicht betrekken hè.
'And **you** see the face fall.'
[C.I, INT, fn000206.216]

Although the story starts as a description of a specific event which happened in the past, and although the speaker was the actual person who saw her student's face fall, she presents the end of the story as a generalisation. We can infer that this event represents a tendency, and that anyone in this kind of situation would see how the shock made the students' faces fall. The way of making a generalisation in (31) is more subtle than the one in (30): the utterance is not an explicit conditional. However, the mechanism is the same: the speaker generalises a personal experience by implicitly reasoning as follows: I have seen this student's face fall, and I have been in similar situations more than once, so if someone else faced immigrants with such cultural differences, she could see their faces fall too. Again, a personal experience is presented as a generalisation, for which the generic pronoun *je* is suitable.

The move from specific to general can also happen more gradually, as shown in (32).

- (32) SP1 wat schilder **je** dan?
'And what do **you** paint?'
SP2 in aquarel, ja, en in de periode dat **ik** het nog wel deed nu staat 't op een laag pitje dan volgde **je** een cursus en met een groep ging **je** dan een stilleven schilderen en een enkele keer met mooi weer wordt er een plekje uitgekozen waar **je** met z'n allen naartoe gaat en dan ga **je** buiten schilderen.
'In water colours, yes, and when **I** used to do it – it's on a low profile now – then **you** followed a course, and with a group, **you** went painting a still life, and if the weather is good, a place is chosen to go with all of **you** and then **you** are painting somewhere outside.'
[C.I, INT, fn000190.137 - fn000190.140]

Again, there is an interviewer question containing the deictic pronoun *je*, which refers to the interviewee. In the initial part of the answer, the speaker is indeed giving specific information about herself: she uses the first person pronoun *ik* and she uses the past tense to describe a period in the past. After *dan* ‘then’ she switches to *je*, but we know that she still refers to herself, because *dan* refers to the specific period in which the speaker used to paint in water colours and because there is no reason why this would go for people in general. Next, she also switches from past tense to present tense, which makes the situation even more generalised. Again, we can see very clearly that self-referring *je* is used in generalisations, as a generic pronoun, whereas we know at the same time that the actual referent is the speaker. Thus, self-referring *je* is in fact generic *je*, which the hearer considers to be specific in reference.

3.3.3 Subjectivity and the role of the hearer

Now that we know that self-referring *je* is in fact generic *je*, in an utterance that is presented as a generalisation, the question arises how we know whether a generalised expression is a genuine generalisation, or only presented as such while actually being specific. The answer explains why it is impossible to make a clear distinction between self-referring generic *je* and non-self-referring generic *je*: it partly depends on the hearer’s *belief* in the generalisability of the situation.

After making a more precise annotation, which, as said before, did not work out well, a second annotator made a precise annotation too. I categorised 102 tokens as ‘referring to the speaker in the first place’. The second annotator only categorised 41 tokens as self-referring, nine of which I did not annotate as self-referring. Although the overlap of 32 tokens should not be neglected, it is obvious that the comparison does not show perfect agreement. This is exactly what we would expect if the interpretation of *je* is hearer-dependent: although some factors cause more agreement among listeners (specific questions, pronoun switches), our belief in the generalisability depends on our own frame of reference.

This mechanism makes the situation possible that the hearer believes that a speaker using *je* is actually talking about herself,

while the speaker knows that this is not the case. This is possibly the case in (33), which is uttered after a silly remark that the speaker just made.

- (33) ‘t is na vijf geweest en **ik** denk van negen tot vijf en dan daarna dan denk **je** niet meer zeg maar hè
‘It is after five now and **I** think from nine till five and after that **you** don’t think anymore, right.’
[C.II, TEL, fn008027.178]

I perceived *je* in this utterance as self-referring, firstly because of the pronoun switch, and secondly because I did not recognise this generalisation and therefore did not immediately believe that people in general only think from nine till five. However, the speaker may know of several people who often have the same experience, which makes it perfectly generalisable for him. This might explain why soccer players are thought to use self-referring *je* more than other speakers do: soccer players generalise their own situation, but the people watching them on television do often not recognise their situation as generalisable and may yell at the television that it is not ‘people in general’ who missed the goal, but only the speaker himself.

It is also possible that I as a hearer do not perceive *je* as self-referring when it actually refers to the speaker only, because I do believe that the situation is generalisable. This is shown in (34), in which the speaker is describing the specific function that she has as a teacher of Dutch for immigrants.

- (34) maar **ik** ben cursistenbegeleidster van de nieuwkomers dus de inburgeraars die een contract met de gemeente hebben en die moet **je** nauwlettend in de gaten houden met presentie en voortgang
‘But **I** am a supervisor of the new students, who are immigrants having a contract with the municipality, and **you** have to keep a close eye on them with respect to their presence and progress.’
[C.I, INT, fn000206.70 - fn000206.71]

Although we can infer from the context that the speaker is the one who has to keep a close eye on the new students, I do not perceive *je* in (34) as typically self-referring, because my common knowledge tells me that anyone in the position that the speaker more or less coincidentally has, would have to keep a close eye on those students.

3.3.4 Conclusions

Self-referring *je* turned out to be a non-clear-cut variant of generic *je*. Even if the context provides clear indications that the speaker is talking about a specific situation, it is presented as a generalisation about people in general. Whether *je* is perceived as self-referring partly depends on the listener, and therefore, considering *je* to be self-referring is a highly subjective matter. As a hearer, we use more and less clear cues for our interpretation: sometimes, the sentence structure gives us an idea about it, and often, our personal frame of beliefs, expectations and experiences is an important factor. If we recognise the generalisation (for example, because we experienced something similar once), we will be inclined to believe that the generic pronoun really refers generically; if not, we will suspect the speaker of talking about herself.

The obvious question now is: why do speakers use a generic pronoun to refer to themselves? In order to find an answer to this question, we need to know more about generic *je*. This issue will be discussed in the following section.

3.4 Generic *je* and self-referring *je*

What or whom does generic *je* refer to? The reference of generic pronouns is usually considered to be roughly equivalent to ‘people’, ‘everyone’, ‘anyone’ (Egerland 2003), ‘someone’, or ‘the typical person’ (Moltmann 2006). The following sentences show how generic *je* can indeed be paraphrased using such words.

(35a) **Je** moet werken om te kunnen leven.

‘**You / one** must work to live.’

[adapted from Egerland 2003:73 (1a)]

(35b) People / everyone must work to earn a living.

(36a) Als **je** een neus hebt, kun **je** ruiken.

‘If **you** / **one** have/has a nose, **you** / **one** can smell.’

[adapted from Moltmann 2006:277 (64b)]

(36b) If someone has a nose, she can smell.

or: Anyone who has a nose can smell.

The sentences above are statements concerning people in general, which almost have the status of a universal law: worldwide, there are only few exceptions to the rule in (35a), and as far as I know, having a nose even is a strict condition for people to be able to smell.

My data show that in natural conversation, however, *je* mostly does not refer to all people in the world, but to a subset thereof. This phenomenon is not limited to Dutch *je* or second person pronouns. Ramat & Sansò (2007), in their paper on the developmental stages of the impersonal pronoun derived from the noun meaning ‘man’, show that such ‘*man*-elements’ refer to ‘all human race’ or ‘mankind’ in their initial stage, but grammaticalise into a pronoun that can refer to a more restricted subgroup of relevant human people. Although *je* is etymologically very different from *man*-elements, its behaviour as a generic pronoun is apparently quite similar to those elements.

The main question is now: how do listeners know which people are relevant, according to the speaker? Again, the context in which the utterance occurs is decisive. Apparently, listeners are perfectly able to combine all explicit and implicit cues that are hidden in the context with their knowledge of the world. In what follows, I will discuss the explicit and implicit restrictions that occur in my data.

3.4.1 *Explicit restrictions*

Explicit cues seem to occur in the shape of adverbials. In (36a) in the previous subsection, we have already seen an example of this process: the adverbial *if*-clause of a conditional can function as a restrictor of the referent, and the second *je* thus refers to ‘those people having a nose’. In (37), I give an example of condition-restricted *je* from my data. This utterance was taken from a

conversation about a telephone service that people can call when they feel that they are in trouble.

(37) voor als **je** zeg maar homo bent en dan **je** weet niet hoe **je** met je ouders moet 't erover hebben of zoiets weet je wel dan kun **je** hun bellen.

‘[It’s] for if **you** are, like, gay and then, **you** don’t know how **you** should tell it to your parents or something like that, you know, then **you** can call them.’

[C.II, CON, fn000541.298]

In (37), *als je (zeg maar) homo bent* restricts the actual referent of the following three occurrences of *je*, and the referent of the last *je* is further restricted as ‘homosexual people who do not know how to talk about it with their parents’.

Explicit restrictors are not necessarily conditionals, but sentences with a restricting adverbial phrase are usually paraphrasable as a conditional. We see this in the following examples.

(38a) In Spanje heb je dat ook daar heb **je** twee achternamen zeg maar de naam van je moeder en van je vader.

‘You have that in Spain too: there **you** have two surnames: your mother’s name and your father’s name’.

[C.II, CON, fn000630.191]

(38b) If you / one live(s) in Spain, then you have / one has two surnames.

(39a) Als student heb **je** een studentenbaantje.

‘As a student, **you** have a student job.’

[C.II, INT, fn000288.10]

(39b) If you are / one is a student, then you have / one has a student job.

The examples above show how a prepositional phrase can determine the actual subset to which *je* refers: in (38a), *je* is restricted by the adverbial *daar* and its antecedent *in Spanje*, and in (39a), *je* is restricted by the phrase *als student*. In the latter sentence, the unrestricted *je* that would have occurred in the condition part of a conditional is even absent; the only *je* present is already restricted by the adverbial phrase.

3.4.2 *Implicit restrictions*

Even in those cases where explicit restrictors are absent, *je* does often not refer to ‘the typical human being’, but rather to ‘those people relevant in this conversation’. As pointed out before, making explicit the restrictions on relevant people is extremely difficult, whereas interpreting the pronoun correctly is seldom a problem. I will give some examples to show how this works.

At first, *je* often occurs in sentences that seem to be the consequent of a conditional: although an explicit *if*-clause is missing, there is a *then*-clause. This is shown in (40a), in which the speaker tells that he does not celebrate carnival very enthusiastically anymore since he moved to a city where people do not celebrate carnival at all. The conditional in (40b) expresses the speaker’s supposition hidden in (40a).

(40a) maar toen woonde **ik** net één of twee jaar in Utrecht en dan begin **je** d’r al een beetje van af te kicken hè.

‘But **I** had lived in Utrecht for one or two years by then, and then **you** start kicking it, don’t you.’

[C.II, CON, fn000375.115]

(40b) If you have lived in Utrecht for a while, then you start kicking carnival.

Of course, the condition inferred in (40b) is in fact quite explicit in (40a). However, there is a second condition hidden in (40a): *je* refers to people who come from a Dutch region in which carnival is celebrated, that is, the south. And we can even go one step further and argue that there is a third condition: *je* refers to people who used to celebrate carnival before they moved out from the south. The hearer can infer all this partly from the linguistic context (the speakers tell each other about their own carnival experiences) and partly from (Dutch) common knowledge (in southern regions in the Netherlands, people celebrate carnival, and in the remaining regions, they do not).

3.4.3 Self-referring *je* as an instance of generic *je*

Now that we know that *je* is highly flexible in reference, we seem to be able to explain how self-referring *je* can emerge. We have seen that the reference of *je* can vary from ‘all humanity’ to a very limited subset of this group: those people relevant in the specific situation. If this subset can be as small as we want it to be, we expect that it can also be just one relevant person, and thus, it can refer to the speaker only. The underlying conditional would then be something like ‘if one was me / if one was in my specific situation, then ...’.

However, this representation is problematic. Firstly, it is important to remember that self-referring *je* has, according to the literature, the *appearance* of generality, that is, the speaker presents the situation as a generalisation, but it is not really a generalisation. Thus, self-referring *je* presents the situation differently from what it is like in reality. This mismatch between the real situation and how it is represented was also found in my data. In contrast, when *je* refers to a subset of all people in the world, it is not perceived as a representation which deviates from the situation in the real world. Apparently, listeners naturally accept cases of generic *je* in which the referent is only a small group (e.g., ‘teachers of Dutch’, or ‘people from the south’), whereas they do not automatically accept instances of self-referring *je*, in which the referent is restricted to only one person.

Secondly, when we examine (41) in some detail, we can see that self-referring *je* does not really presuppose the underlying conditional as formulated above. In the conversation where it was taken from, the speakers are talking about driving a car when there is a lot of traffic, and a serious accident that they read about in the newspaper, in which someone drove with high speed into a traffic jam.

(41) en dan ben **ik** ook zo bang hè als **je** dan moet stoppen en dan kijk **je** in de spiegel en dan komt er iemand achter **je** aan...

‘And then **I**’m so scared; if **you** have to stop and **you** look in the mirror and then someone is driving behind **you** / is pursuing **you**...’

[C.II, CON, fn000344.106]

The pronoun switch from *ik* 'I' to *je* clearly indicates that the speaker is talking about her own experiences when driving a car, as it is unlikely that she is scared in any case in which someone else is driving. Put differently, the paraphrase 'if you are driving a car and you have to stop, and you see someone driving behind you when you look in the mirror, *I* am always very scared' does not make sense; (41) is rather paraphrased as 'If I am driving a car and I have to stop, and I see someone driving behind me when I look in the mirror, I am always very scared'. Yet, a conditional is present: an implicit one, determining the referent as 'people who drive a car', and an explicit one, determining the referent further as 'people who drive a car and who have to stop and look in the mirror and see someone driving behind them'. However, if the referent is determined in this way, it is still a subset that, at least potentially, includes more persons than only the speaker, as it does not follow from the context that the speaker is the only person in the world who is sometimes driving a car and sees someone behind her when looking in the mirror before stopping. Moreover, 'if you were me, and you were driving a car...' is not an appropriate paraphrase of (41): using *je*, the speaker does not claim that only people who are like her would be scared in a driving situation, but she claims (implicitly) that anyone would be scared in this situation.

In the following examples, we see that self-referring *je* can, like 'genuine' generic *je*, vary in the size of its referent. In (42), the speaker is talking about his daily habit of jogging in the forest. The speaker of (43) is a teacher of Dutch for low literate adults, who is telling the interviewer that she likes reading thrillers because it enables her to stop thinking about her work when she is at home. The speaker of (44) is probably losing his job. He argues that it is a hard time for finding a new job in the building trade and he utters (44) to found this claim.

- (42) daar voel **je** je goed bij
 'That makes **you** feel good' / 'Then **you** will feel good'
 [C.II, INT, fn000138.232]
- (43) heel vaak ben **je** thuis ook nog over dingen aan het doordenken hè
 '**You** often keep reflecting on things at home, you know.'
 [C.II, INT, fn000209.323]

- (44) kijk **je** zit vlak voor de kerst **je** komt nergens meer binnen.
'Look, it will be Christmas soon; you won't be employed anywhere.'
(lit: 'Look, **you** are just before Christmas; **you** won't come in anywhere.')
- [C.II, CON, fn007856.330 - fn007856.331]

The claim made in (42) is simply that people in general feel good when they take a daily exercise. However, it is the speaker himself who jogs, and who feels good when he does. (43) is, literally, a statement about the speaker and her colleagues: the job is emotionally demanding, which results in the situation that it is hard to stop thinking about it in the evening. At the same time, we know that the situation described in (43) is so specific that it is probably merely the speaker who keeps thinking about it. And (44), finally, is remarkable in that the first part of the utterance refers to a certain point in time, and at the moment that (44) was uttered, the 25th of December was coming for all people in the world (although a lot of people would not call this day 'Christmas', or 'the 25th of December'). At the same time, the second *je* token can only refer to a small group of building workers in a specific country or even a specific city, as our common knowledge tells us that it is by far not a universal truth that nobody will be employed just before Christmas. And as the situation of being fired as a building worker at that particular point in time is very specific, we assume that *je* refers to the speaker.

The examples above suggest that although generic *je* can refer to a very limited subset of human people, self-referring *je* cannot simply be analysed as an extremely limited subset. Rather, self-referring *je* is exactly like the real generic *je* in that its referent can be, and often is, contextually restricted, but unlike 'normal' generic *je*, self-referring *je* is only presented as a generalisation, whereas contextual and non-linguistic cues tell the listener that it probably refers to a specific person in the real world.

3.4.4 A first motivation for the use of self-referring *je*

Now, we have found out what self-referring *je* is: a variant of generic *je* that is used in statements that are presented as

generalisations, while they in fact refer to specific situations. Even though we have reasons to assume that *je* refers to the speaker, it is presented as if it does not. Apparently, speakers sometimes feel the need to generalise from their own situation. Now we come to an important question: why do speakers feel this need?

At first, speakers can use a generic pronoun in order to avoid a first person pronoun because they want to weaken the link between themselves and their message. This mechanism has a distancing effect, which a speaker can use when she feels insecure or uncomfortable about something, that is, when she does not (want to) identify with something. This makes a strategic use of *je* possible, which was noted by Van Wassenaar (1994), who points out, as we saw in Chapter 2, that speakers often use self-referring *je* in case of personal failure.

However, the reader should note that the mechanism of generalising specific experiences, feelings and the like is not as uncommon as it seems. Several authors (e.g., Scheibman 2007) have shown that generalisations are omnipresent in conversations. Speakers do not typically utter well-founded, inductively established generalisations; rather, they reason intuitively, often on the basis of their own experiences. Apparently, people are inclined to believe that what goes for themselves, probably will go for anyone else, or at least for those people they identify with. The following example illustrates this. The speaker is a teacher of Dutch as a second language for immigrants.

(45) maar de laatste twee jaar zit ik vooral met de hogeropgeleiden en da's wel heel leuk tuurlijk daar kun **je** echt goede gesprekken mee voeren

'But I have mainly been teaching the more highly educated students for the last two years, which is very nice of course; with them, **you** can have really nice conversations.'

[C.I, INT, fn000206.74]

Of course, it is not true that people in general typically have nicer conversations with highly educated people than with more lowly educated people; therefore, we assume on the basis of common knowledge that it is the speaker herself who likes to talk with the highly educated immigrants. However, the same sentence with a

first person pronoun has a very different connotation, as we see in (45').

(45') ... daar kan **ik** echt goede gesprekken mee voeren
'... with them, **I** can have really nice conversations.'

In (45'), the speaker does not only seem more immodest, as if she is emphasising her own high educational status, but more importantly, the focus is different from (45). In (45'), the focus is on the speaker, whereas (45) focuses on the highly educated immigrants. This is not surprising: it has often been pointed out that impersonal pronouns are, like passives, defocusing strategies (e.g., Sansò 2006). It is probably not a coincidence that such an unmarked form as *je* is used. The reason that the speaker uses *je* here is probably not that she wants to avoid the first person pronoun, but rather, (45') is simply not the message that she wants to express, because this part of the sentence is not meant to tell something about herself, but rather about the immigrants. And as her conversational partner is highly educated too, he will probably recognise and accept this generalisation. According to Scheibman (2007), generalisations indeed have an inclusive effect: the conversational partners use them in order to emphasise their mutual agreement. This role of the addressee brings us to the final piece of the puzzle: the link with deictic *je*.

3.4.5 Conclusions

In this section, it was shown that generic *je* refers to people in general, roughly paraphrased as 'anyone' and the like, but that it usually refers to a smaller subset of people relevant in the specific situation. Listeners use all kinds of cues to interpret *je* correctly. *Je* seems to presuppose an underlying, more or less explicit conditional, which determines the relevant subset. It was therefore suggested that self-referring *je* presupposes a conditional that limits the referent to only one relevant person, which is the speaker. However, we saw that this analysis of self-referring *je* is not satisfying: self-referring *je* has the appearance of generality, whereas generic *je* with a restricted referent is perceived as a

genuine generalisation. Moreover, the linguistic referent of *je* is still a more or less restricted subset of ‘people in the world’.

3.5 Deictic *je*, generic *je* and self-referring *je*

At this moment, deictic *je* re-enters the scene. In my final analysis, I will combine all types of *je* that we have seen, and account for the fact that it is the second person pronoun which is, in its generic use, sometimes used with first person reference. Therefore, I will start this section by examining those tokens that have not received any attention before, while they will turn out to have a key-role in this story: the tokens that can have both a deictic and a non-deictic interpretation.

3.5.1 *The ambiguous cases*

In Corpus II, I annotated 37 tokens as ‘ambiguous’. A closer examination shows that in these cases, its exact reference (deictic or generic) is often irrelevant, and that the ambiguity is thus usually unproblematic. (46) is an example.

(46) als **je** zoiets regelt dan voel **je** je daar tuurlijk ook verantwoordelijk voor.

‘When **you** arrange something like that, then **you** will feel responsible for it of course.’

[C.II, TEL, fn008210.273]

In the conversation from which (46) was taken, one speaker is telling her conversational partner about an excursion the day before, which she arranged. She tells her how some things went wrong and how she tried the best she could to solve those problems. (46) is her conversational partner’s reaction to this story. As a conditional, (46) is a generic sentence: by using the present tense, the speaker generalises from the specific situation in question. However, we have seen in section 3.2.1 that a deictic interpretation of *je* is not ruled out in a generic sentence. The context does not give the definite answer in this case: a deictic reading is possible because the addressee is the central person in the discourse, and a generic reading is possible because we know that feeling responsible for

something one arranges is natural. However, the two readings are perfectly compatible with each other and the ambiguity does not lead to interpretational problems.

3.5.2 *From personal to impersonal pronoun*

In principle, it is possible that the link between second person and generic reference is purely accidental. In Chapter 4, however, evidence will be given for the opposite: there is a natural – though not a forcing – link between second person and generic reference. For now, we will simply assume that the possibility of *je* to refer generically is not a coincidence. On the basis of this assumption, the presence of ‘ambiguous’ tokens as we saw in (46) is not surprising: it is natural to suppose that the development towards non-deictic *je* started with this kind of cases. I suggest that the source of the generic use of *je* is the invitation to the addressee to imagine herself to be in a certain situation. That self-referring *je* implies such an invitation was already suggested by Van Wassenaar (1994) and Onrust (1995), as we saw in Chapter 2. This invitation is explicit in two types of phrases: conditionals and the imperative *stel je voor* ‘imagine’. In the following sentences, the small step from deictic to generic *je* is illustrated: in (47a), *jij* and *je* receive deictic interpretation, whereas *je* in (47b) will rather be interpreted generically.

(47a) Stel je voor: als **jij** zo moe zou zijn, zou **je** nu ook naar bed willen.

‘Imagine **yourself** being so tired: **you** would want to go to bed now too.’

(47b) Als **je** heel moe bent, wil **je** naar bed.

‘If **you** feel very tired, **you** want to go to bed.’

Of course, this is only speculative; historical data can give the final answer. However, it would explain why sentences with generic *je* are often naturally paraphrasable with a conditional construction: the speaker takes an imaginary addressee as ‘the typical person’, who is the starting point for the generalisation. In principle, *je* thus refers to anyone, represented by this imaginary addressee. Restrictions in the context can specify the actual situation that the

speaker is talking about. The imaginary addressee is put in this situation, and the actual referent seems to be a subset of contextually relevant people.

This relation between deictic and generic *je* explains a difference in interpretation that exists within the class of generic pronouns in Dutch, to which I already referred in Chapter 2. This difference is illustrated in the following sentences.

(48a) In Frankrijk drinken **ze** wijn bij het eten.

‘In France, **they** drink wine with dinner.’

(48b) In Frankrijk drink **je** wijn bij het eten.

‘In France, **you** drink wine with dinner.’

Both (48a) and (48b) are generic sentences, but while (48a) can be paraphrased as ‘French people have wine with dinner’, (48b) would rather be ‘When someone is in France, she will have wine with dinner’. Thus, *ze* does not presuppose a conditional, whereas *je* again does, as we have seen several times now. This difference is not due to the fact that *ze*, having third person reference, excludes the speaker and the addressee in its referents, as the first person plural *we*, which can also refer generically, behaves like *ze* in this respect, as we can see in (48c).

(48c) In Nederland eten **we** vroeg.

‘In the Netherlands, **we** have dinner early.’

or: ‘(We) Dutch people have dinner early’.

Apparently, pronouns like *ze* and *we* describe the world how it is, and a pronoun like *je* describes the world how it might be from a certain point of view or in a certain situation.

An example from the corpus which illustrates this mechanism is (49), in which the speaker tells that she has trimmed a bush which was partly hanging over the pond in her garden.

(49) en ik denk van jee nou dat daar heb je eigenlijk niets aan die vijver **je** ziet ‘m helemaal niet.

‘And I think/thought, now you don’t have use of that pond: **you** don’t see it at all.’

[C.II, TEL, fn006924.86]

Clearly, (49) does not say that people in general do not see the pond in question, although this is in fact true, but rather that someone would not see the pond if she was standing in front of it. This *je* does not refer to a group of people that exists in the actual world, but to a hypothetical person: an imaginary addressee, standing in front of this specific pond. In fact, the referent of *je* is not really restricted: it could be anyone.

3.5.3 A final motivation for the use of self-referring *je*

As was pointed out in Chapter 2, it has been suggested that self-referring *je* seems to appeal to the addressee. Now that we found out about the link between generic and deictic *je*, it is clear why this can happen with the impersonal pronoun *je*: even in its generic use, *je* presupposes an imaginary addressee. This mechanism invites the addressee to take the perspective of the speaker: the situation is seen from within instead of being seen from outside. The generalising force of *je* gives the impression that the addressee, as well as anyone else, would have had the same experience, thoughts or feelings as the speaker if she had been in her position. This explains the idea of several Dutch authors that *je* is used in order to evoke imagination and empathy in the addressee: imagining oneself to be in someone else's position is the basis of empathic feelings.

Importantly, generic *je* can be used specifically without referring to the first person. In my data, I have also found some, though not many, instances in which it refers to a third person. The following two utterances are examples.

- (50) en dat is wel moeilijk om **die jongens** aan de gang te krijgen want ja of **die** zijn met hun brommers aan het sleutelen of **ze** hebben een krantenwijk of **ze** zitten in de supermarkt of **ze** hebben een andere baan en dan komt school in feite heel slecht uit omdat **je** d'r dan nog bij aan het werk bent.
'And it is hard to motivate **those boys**, for **they** are either tinkering with their mopeds, or **they** have a paper round, or **they** work in the supermarket, or **they** have another job, and then school is in fact quite inconvenient because **you** are also working.'

[C.I, INT, fn000189.56]

(51) maar het was even de vraag van of 't inderdaad zou gaan lopen want in 't begin is 't toch wel spannend natuurlijk hè **je** doet toch een uh ja **je** moet iets gaan opstarten ja **hij** moest ergens een ruimte hebben

'But for some time, it was the question whether it [the business] would be successful indeed, for that is precarious in the beginning of course: **you** are doing a, uh, well, **you** must set up something, and well, **he** had to find a room somewhere.'

[C.II, TEL, fn006868.138 - fn006868.142]

In (50), a teacher of Dutch on a secondary school for professional training is talking about his pupils' motivation problems. (51) was taken from a conversation in which one of the speakers tells the second speaker about the way his father set up his own business. In both utterances, *je* does not refer to the speaker or to a subset in which the speaker is included: in (50), *je* refers to *die jongens* 'those boys', and in (51), both tokens refer to the speaker's father. Interestingly, the present tense is used in (51), although it clearly refers to a situation in the past. Both utterances are generalisations from a strong empathising perspective: the speakers present the respective situations as totally imaginable. Intuitively, the same mechanism as in self-referring *je* is working here; the only difference is that *je* does not refer to the speaker.

That specific *je* is not necessarily self-referring suggests that the reason why it often *is* self-referring is a pragmatic one. It is well-known that speakers mainly talk about themselves, which results in omnipresence of first person pronouns (e.g., Dahl 2000). It is also the speaker who is the most in need of the addressee's empathy. That speakers usually take themselves and not other people as the point of departure for a generalisation is not really remarkable either.

Besides the possibility of evoking the addressee's empathy, we saw that *je* also has the possibility of defocusing from the subject or agent. We saw that this makes it possible to formulate a message that would be different (e.g., in whom is focused on) when *ik* 'I' was used, even though it might actually refer to the speaker herself. The defocusing and generalising force of *je* also give the speaker the

possibility of creating distance between themselves and their message. Together, these aspects of *je* can have different effects, dependent on the context: the utterance can get a modest flavour, as was found by Bennis et al. (2004), or a flavour of false modesty, as perceived by Appel (1993), but it can also conceal the gravity of failure, as pointed out by Van Wassenaar. The meaning of *je* is so flexible that there is not one speaker's motivation to use it instead of a first person pronoun.

In the light of Hogeweg (2009), the flexibility of *je* is not very surprising. She found that word meanings in general are flexible, and that the interpretation of a word depends on the context in which it occurs, which is exactly what we saw happening with *je*. She argues that speakers start with an intention to express something. This intention consists of semantic features. The speaker will choose the word from her lexicon that shares the most semantic features with her intention. In this chapter, we have seen why *je* is in some situations more appropriate for reference to the speaker (or a third person) than *ik* (or *s/he*). As hearers are usually perfectly able to interpret *je* correctly, the ambiguity of *je* is not (always) problematic enough to choose *ik*. Moreover, using *je* is still an economical way to express this intention, because it is not a new form, but a form that already existed for addressee- and generic reference.

3.5.4 Conclusions

In this section, generic *je* has been connected to its source: deictic *je*. It was suggested that the conditional may be the first step from purely deictic *je* towards more generic *je*. In this context, the use of self-referring *je* is not as strange as it seemed at first sight. Speakers tend to generalise on the basis of their own experiences, and the impersonal *je* is even more suitable for this goal because it implicates an appeal to the addressee, who is invited to take the speaker's perspective. If the addressee accepts this perspective, the effect is a feeling of empathy. This effect was already noticed in the literature. Strong evidence that the speaker-oriented nature of generic *je* has a pragmatic reason was found in the fact that *je* can also refer specifically to a third person. *Je* thus turns out to be a highly flexible form; not a substitute for the first person pronouns *ik*

‘I’ and *me* ‘me’, but an economical form which can present specific instances for several reason as a generalisation.

3.6 General conclusions

In this chapter, I have shown on the basis of two subcorpora from the CGN that the interpretation of *je* is highly context-dependent. This goes for the distinction between deictic and generic *je*, but also for the detection of self-referring *je*. As a matter of fact, the border between self-referring *je* and non-self-referring generic *je* turned out to be non-existent: self-referring *je* is in fact generic *je*, but contextual cues and the hearer’s perspective can lead to a speaker-referring interpretation. Therefore, it is a highly subjective matter.

Next, it was shown that generic *je*, although implying a generalisation, does not always refer to all humanity. In contrast, it usually refers to only a small subset thereof. The hypothesis that this mechanism can be extended to self-referring *je*, that is, that self-referring *je* refers to a subset of only one person, the speaker, was examined and rejected. Importantly, self-referring *je* always suggests that the situation, how specific it may be, can be generalised to other people than the speaker. Thus, the statement has the appearance of generality, which is or is not accepted by the hearer.

In the final section, the role of the hearer was made explicit. I suggested that generic *je* evolved from deictic *je* in conditional constructions. Therefore, the hearer is, even with generic *je*, addressed and invited to take the speaker’s perspective. It was argued that *je* refers to an imaginary addressee and does not show the world how it is, but rather how it might be from a certain point of view. The effect of this mechanism is the addressee’s empathy (that is, if the addressee accepts the speaker’s point of view). I showed that speakers do not only use *je* specifically to refer to themselves, but sometimes also to refer to a third person. This clearly indicates that *je* is not a substitute for *ik* ‘I’; rather, we could say that the reason why self-referring *je* occurs more often than third-person-referring *je* is pragmatic: speakers mostly talk about themselves and they need the addressee’s empathy more for themselves than for others. The different aspects of *je* (evoking empathy, defocusing, generalising) lead to different effects in

different contexts. Speakers can use these effects to give their message a nuance that would be absent when using a first person pronoun.

Finally, I pointed out that the flexibility of *je* is not as surprising as it might seem; in fact, it is a very economical way of expressing a message for which the first or third person pronoun is sometimes less appropriate.

Impersonals and seconds: a crosslinguistic comparison

In the previous chapter, I presented an analysis of Dutch generic *je*, and in particular of those instances that seem to refer to the speaker. Pragmatics turned out to be of major importance for the use and interpretation of *je*. This makes it plausible that in other languages than Dutch, second person pronouns and other generic constructions can refer to the speaker as well. In this chapter, I will therefore focus on the phenomenon from a crosslinguistic perspective. In section 4.1, I will discuss the possibility of self-reference of generic constructions that are not originally second person pronouns. Second person pronouns will be the topic of section 4.2. In the first part of section 4.2, I will examine the generic reference possibilities of second person pronouns. Next, the occurrence of self-referring second person pronouns will be discussed. In section 4.3, the possible motivations for the use of self-referring generic constructions and its effects will be discussed, as well as a comparison between second person pronouns and other generic constructions. Finally, I will draw conclusions in section 4.4.

4.1 Generic constructions and self-reference

If pragmatic factors indeed determine the use of self-referring *je*, as I have argued above, it is likely that self-reference of a generic pronoun is not a typically Dutch phenomenon. Thus, the first question that will be answered here is: can generic constructions in other languages than Dutch also refer specifically? If they can, this is evidence that the generic nature of *je* makes self-reference possible, and that speakers are apparently inclined to generalise their own situation.

In fact, self-reference and other specific reference of generic constructions with human reference¹² widely occurs throughout languages, regardless of their form. The impersonal pronouns with human reference, which are highly common in Europe, can be used specifically, and the same goes for zero constructions as found in Finnish and Mandarin Chinese, and the morphemes with generic human reference that occur in rGyalrong (Sino-Tibetan) and Athapaskan languages like Dogrib and Slave. However, there is variation in how direct or unambiguous this first person reference is, which I will show in the following subsections.

4.1.1 *Generic and arbitrary*

In order to understand what makes a generic construction refer unambiguously to ‘I’ or ‘we’, we first need to go back to the distinction between episodic and generic sentences, which was referred to earlier in section 3.2. In contrast to generic *je*, the majority of the impersonal pronouns in Europe can occur in non-generic contexts as well¹³. In those episodic contexts, however, they do not refer to ‘people in general’ and neither to a subset thereof, but rather to a particular person or group which is, for some reason, not further specified. I will use the term ‘arbitrary’ for this second interpretation¹⁴. The first who related this semantic difference to the syntactic structure of the sentence was Cinque (1988), who focused on Italian *si*¹⁵. As many of the constructions in question can have

¹² The impersonal constructions as common in Europe are usually claimed to have obligatorily human reference. Although this is virtually true, I would rather say that they refer to those entities that we identify with, which are usually human people but sometimes also (digital) pets, for instance.

¹³ Of course, *je* can occur in non-generic contexts, but then it is a personal pronoun with deictic reference.

¹⁴ The terminology in this research domain is often confusing. I have chosen to follow Egerland’s (2003) terminology, as it occurs to me as the most transparent one.

¹⁵ Initially, episodicity was assumed to be the determining factor. Later, alternative accounts have been presented; for example, the speech act and aspect have been argued to be a better predictor for the right interpretation. In my thesis, the differences among these accounts are not of any relevance and for the sake of clarity, I will maintain the initially proposed link with episodicity. For more details on the subject, the reader is referred to Condoravdi

both interpretations, I will use the covering term ‘impersonal constructions’ from now on.

The Dutch impersonal pronoun *men* is an example of an impersonal pronoun which can have both a generic and an arbitrary interpretation. The difference is illustrated in (52a) and (53a).

(52a) **Men** leeft maar een keer.

‘**One** lives only once.’

[Weerman 2006:7 (6a)]

(53a) Vanmorgen heeft **men** de beslissing genomen.

‘**They** have made the decision this morning.’

(52a) is a generalisation: as there is no contextual restriction, it refers to people in general. Thus, *men* has a generic interpretation in this sentence. In (53a), however, *men* refers to a specific, but undefined group of people. There is no implication of generality here. The speaker may know exactly who she is talking about, but they may be completely unknown as well. Importantly, generic reference of a pronoun suggests (potential) inclusion of the speaker and the addressee, whereas arbitrary reference suggests the exclusion of both.

Some impersonal pronouns only occur in generic contexts, as is the case with generic *je*. An example of such a pronoun is English *one*. The examples below show that *one* can occur in the English counterpart of (52a), but not in the counterpart of (53a).

(52b) **One** lives only once.

(53b)***One** made the decision this morning.

In the generic sentence (52b), *one* is fine, and has generic reference. The ungrammaticality of (53b) shows that the English impersonal pronoun *one* cannot occur in episodic sentences.

(1989) and Alonso-Ovalle (2002) for the problems with episodicity as a determining factor, and to Egerland (2003) for an alternative account.

4.1.2 Specific reference in episodic contexts

Now, we will see why the discussed distinction is relevant for the self-reference of impersonal pronouns. It has been observed that, at least in some languages, the arbitrary interpretation is more limited than the generic interpretation: while the latter seems to be compatible with all kinds of predicates, the former is not. I will not discuss the factors that determine the predicate's compatibility with an arbitrary interpretation here, but see Cinque (1988) and Egerland (2003) for details. For now, it is only relevant that there is a third interpretation that impersonal pronouns can have: the specific interpretation. If an impersonal pronoun occurs in a context in which both the generic and the arbitrary reading are excluded, it refers unambiguously to the speaker or to a group including the speaker.

According to Ramat and Sansò (2007), the arbitrary and the specific reading of *man*-elements (impersonal pronouns that are derived from the noun meaning 'man') presuppose a generic reading, that is, there are no *man*-elements that can refer arbitrarily or specifically but not generically. The arbitrary and the specific reading do not presuppose each other. Thus, besides the *man*-elements which have generic reference only (e.g., Hungarian *az ember*, Sigurðsson & Egerland 2009:162) and *man*-elements with generic and arbitrary reference only (e.g., Dutch *men*, Weerman 2006; German *man*, Zifonun 2001), there are also *man*-elements with generic and specific reference only (e.g., Icelandic *maður*, Sigurðsson & Egerland 2009:163), and *man*-elements with all three possibilities (e.g., French *on*, Cinque 1988; *man* in some varieties of Swedish, Egerland 2003). The following examples illustrate the three different interpretations in Swedish.

(54a) Till för femtio år sedan dog **man** vanligtvis
until for fifty year ago died MAN usually
 hemma.
at.home

'Until fifty years ago, one usually died at home.'

[Egerland 2003:80 (10b)]

(54b) Klockan fem sjöng **man** sånger i trappuppgången
the.clock five sang MAN songs in the.staircase
 (det måste ha varit några tonåringar som vanligt).

‘At five o’clock, they were singing songs in the staircase
(it must have been some teenagers as usual).’

[Egerland 2003:77 (5a)]

(54c) I går på eftermiddagen blev **man**
in yesterday on the.afternoon was **MAN**
fast anställd.
permanently employed

‘Yesterday afternoon, I was employed.’

[Egerland 2003:76 (4b)]

According to Egerland (2003), (54a) has an obligatorily generic interpretation: an arbitrary interpretation is excluded because of the predicate, and a specific interpretation is dispreferred for pragmatic reasons. As (54b) and (54c) are episodic sentences, a generic reading is impossible. In (54b), the specific interpretation is again odd for pragmatic reasons; without the sentence between brackets, it would have been perfectly possible. In (54c), however, it is the only possible interpretation because the predicate rules out an arbitrary interpretation. Thus, in this sentence, *man* unambiguously refers specifically, to the speaker. Note that (54c) is ungrammatical for speakers of a variety of Swedish in which a specific interpretation is not possible.

Egerland suggests that his analysis can probably be extended to the counterparts of *man* in Romance languages. There is an interesting difference, however: while Swedish *man* refers to the first person singular, its counterparts in Romance languages refers to a group of people including the speaker: ‘we’. The use of a *man*-element is exceptional in Romance languages; French *on* is etymologically related to *homme* ‘man’, but many other languages use the reflexive for impersonal reference¹⁶. According to Cinque (1988), the interpretation of Italian *si* and Spanish *se* is, however, very similar to the interpretation of French *on*, in spite of their different origin. This similarity is shown below for French in the a-sentences, Italian in the b-sentences, and Spanish in the c-sentences.

¹⁶ Languages often have several strategies for impersonal reference, and so have the Romance languages. Only the reflexive is relevant here.

- (55a) **On** mangeait bien dans ce restaurant.
ON eat.IMP well in that restaurant.
- (55b) In quel ristorante **si** mangiava bene.
in that restaurant SI eat.IMP well
- (55c) **Se** comía bien en este restaurante.
SE eat.IMP well in that restaurant
 ‘One used to eat well in that restaurant.’
 [D’Alessandro & Alexiadou 2003:6-7 (17), (11), (15)]
- (56a) **On** a bien mangé dans ce restaurant.
ON has well eat.PF in that restaurant
- (56b) In quel ristorante **si** è mangiato bene.
in that restaurant SI is eat.PF well
- (56c) **Se** comío bien en este restaurante.
SE eat.PF well in that restaurant
 ‘We have eaten well in that restaurant.’
 [D’Alessandro & Alexiadou 2003:6-7 (18), (12), (16)]

Again, the pronouns refer directly to the first person, but in contrast to Swedish *man*, they have plural and not singular reference.

A use quite similar to the Romance languages was found in the isolate Yuchi (Helmbrecht 2002), and in the Athapaskan languages Slave and Dogrib (Saxon 1993). I give an example from Dogrib. This language has a morpheme (*ts’e*) which originally had impersonal reference, both generic and arbitrary, and which can have specific reference too, as shown in (57).

- (57) Gik’ada-**ts’e**-redè.
3-TS’E.IMP-tease.PL
 ‘We are teasing them.’
 [Saxon 1993:344 (9)]

According to Saxon, this example is potentially ambiguous; however, she does not make explicit which extra interpretation it can get. As (57) is an episodic sentence, I assume that it will be an arbitrary and not a generic interpretation.

The most well-known example of an impersonal construction with specific reference is French *on*, which is probably due to the fact that it has largely replaced the original first person plural

pronoun *nous* in subject position¹⁷ (Laberge & Sankoff 1979; Coveney 2000). This development reflects the fact that the specific use of *on* has become stylistically neutral for many speakers of French, both in France and in francophone Canada.

4.1.2 *The suggestion of a generalisation*

The data presented suggest that the link between impersonal constructions and first person reference is not accidental, as specific reference occurs in both related and unrelated languages. The amount of languages in which specific, direct reference to the speaker occurs is not very large, but we will now turn to the languages in which more indirect reference of impersonal constructions to the speaker is possible.

4.1.2.1 Swedish *man* and French *on*: pragmatics

In his analysis of the three different interpretations of Swedish *man*, Egerland (2003) formulates two restrictions that determine whether an interpretation is excluded or not. The generic reading is incompatible with perfective aspect (which roughly corresponds to episodicity), and the arbitrary reading is incompatible with some particular types of predicates. Neither of the restrictions rules out a specific reading, which implies that this reading is always available. In a way, this is indeed the case, as illustrated with (54a) and (54b), here repeated as (58a) and (58b).

(58a) Till för femtio år sedan dog **man** vanligtvis
until for fifty year ago died MAN usually
 hemma.
at.home

‘Until fifty years ago, one usually died at home.’

(58b) Klockan fem sjöng **man** sånger i trappuppgången
the.clock five sang MAN songs in the.staircase
 (det måste ha varit några tonåringar som vanligt).

‘At five o’clock, they were singing songs in the staircase
 (it must have been some teenagers as usual).’

¹⁷ *On* is, as common for *man*-elements, restricted to subject position.

The reason why a specific reading is ruled out in the sentences is a purely pragmatic one. If *man* in (58a) were interpreted as referring to the speaker, (58a) would say that the speaker usually died at home until fifty years ago, which does not make sense. In (58b), *man* is disambiguated by the sentence between brackets; without this sentence, a specific interpretation is fine. However, the restrictions on a generic and an arbitrary interpretation are stronger than pragmatic rules, which is shown in (59).

(59) #Klockan fem dog **man**.
the.clock five died MAN
 ‘I died at five o’clock.’
 [Egerland 2003:78 (6a)]

Although a specific interpretation of *man* in (59) is odd, it is still the optimal one, because the formulated restrictions definitely exclude the other interpretations of *man*.

Apparently, *man* is strongly speaker-oriented. We see this again in (60), which was taken from a broadcast interview with a local worker on a factory being closed down.

(60) SP1 Och hur upplever **du** företagsledningens agerande?
 ‘and how **you** feel about actions taken by the board?’
 SP2 **Man** blir besviken.
MAN becomes disappointed
 ‘One gets disappointed.’
 [Ragnarsdóttir & Strömqvist 2005:146 (7)]

This little dialogue is strongly reminiscent of self-referring *je*: on the basis of contextual cues (an addressee-referring interviewer question), we can infer that the interviewee is talking about himself. According to Ragnarsdóttir & Strömqvist (2005:146), this instance of *man* differs from those cases in which it occurs in an episodic contexts: while the latter simply means ‘I’, the former still “retains the flavor of a generic expression”. Thus although Swedish *man* can refer directly to the speaker, it does not do so in generic contexts, and like Dutch *je*, it gives the statement the appearance of a generalisation, whereas it might refer to a specific situation in the

actual reality. The same happens in the following example from French, in which the speakers discuss a car accident.

- (61) SP1 **J'** ai cassé le steering avec mon genou parce que **j'** étais assis de côté.
 'I broke the steering wheel with my knee because I was sitting on the side.'
- SP2 Et puis **vous** êtes correct?
and then you.POL are okay
 'And you're okay?'
- SP1 Ah oui **on** s' en vient petit vieux mais
ah yes ON REF on come little old but
on est correct.
ON is okay
 'Oh yes, one arrives as an old little man, but one is okay.'
 [Laberge & Sankoff 1979:420 (2)]

Again, we know from the context that *on* indirectly refers to the speaker. Thus, even though French *on* is the most common pronoun for first person plural reference, it can still refer to the speaker, that is, to a singular person, in certain generic contexts.

4.1.2.2 More indirect speaker-reference

Some languages are claimed to have a construction that reminds of French *on*. In the examples given in the literature, however, pragmatic rather than syntactic or semantic factors seem to determine the interpretation and the reference seems to be ambiguous between a generic and a specific reading. Yet, the examples illustrate the strong relation between impersonal reference and first person reference. I give an example from the Athapaskan language rGyalrong in (62). I also give two examples from the Slavonic language Polish, which has both a *man*-element, as shown in (63), and a reflexive for impersonal (and specific) reference, as in (64).

- (62) ko? tálo mə-jata-tsom? káde **ká**-thi
this milk NEG-take-thither in.a.while KÁ-drink
 re?¹⁸
be.necessary
 ‘Don’t take this milk away, for one (i.e., I/we) will drink it later.’
 [Sun 2005:15 (49a)]
- (63) **Człowiek** chce odpocząć cieszyć się chwilą...
CZLOWIEK wants rest cherish REF moment
 ‘One (but also: I/we) want(s) to rest, cherish the moment...’
 [Ramat & Sansò 2007:106 (34)]
- (64) Proszę nie przerywać. Mówi **się**.
please not interrupt speak SIE
 ‘Please don’t interrupt. I’m speaking.’
 [Siewierska 2007:21 (44)]

In several other languages, the impersonal construction is not claimed to refer directly or unambiguously to ‘I’ or ‘we’, but rather, the reference is indirect. Usually, first person reference of such a construction is considered to be a strategy to avoid explicit speaker-reference for some reason. The following examples illustrate this use: in (65), we see German *man*, in (66), prince Charles of England is quoted using *one* for self-reference, and (67) is an example of the Finnish zero construction, which has third person singular verb agreement.

- (65) Ein junger Mann [...] besteht darauf, meinen Whisky zu zahlen, weil er Vater geworden ist [...].
 ‘A young man insists on paying my whisky, because he has become a father.’

¹⁸ In Sun (2005), all instances of ‘a’ in this sentence are represented upside down. The ‘h’ is small and represents aspiration. Unfortunately, I was not able to insert the correct symbols.

Er stellt sich vor und will wissen, wie
he introduces REF PART and wants know how
man heißt, wieviel Kinder **man** hat,
MAN is.called how-many children MAN has
 vor allem Söhne; ich sage „Five“.
before all sons I say five

‘He introduces himself and wants to know one’s name, how many children one has, particularly sons; I say “Five”.’

[M. Frisch, *Homo Faber*, 217, cited in Zifonun 2001:241 (30)]

- (66) It was a sad moment leaving **one’s** family on the tarmac, waving **one** goodbye.

[Wales 1996:82, cited in Altenberg 2004/5:95 (8)]

- (67) Illalla ei uskaltanut nukahtaa, kun
evening NEG dare fall.asleep because
 pelkäs-i että nukku-u pommiin,
be.afraid-3.SG COMP sleep-3.SG bomb
 Jääskeläinen kerta-a
Jääskeläinen reiterate-3.SG

‘In the evening one didn’t dare to fall asleep, because one was afraid that one would oversleep, Jääskeläinen reiterates.’

[Helasvuo & Vilkuna 2008:231 (17)]

In all three examples, contextual cues tell the listener or reader that the referent must be the speaker: in (65), the young father is talking to the speaker only, and in (66) and (67), the speaker refers to one specific event in the past. However, the situations are presented as generalisations and are thus, like the Swedish example (60) and the French example (61) in the previous subsection, reminiscent of the use of Dutch *je*. According to Paardekooper (1991), the Dutch *men* also had the possibility of indirect first person reference, but this usage has disappeared. This might be due to the fact that *men* is almost completely restricted to written language now (Weerman 2006) whereas first person reference of impersonal constructions seems to be a typical spoken language phenomenon.

4.1.2.3 Non-first person specific reference

The impersonal constructions discussed are fairly similar to Dutch *je*: whereas only some of them can have direct first person reference in episodic contexts, all of them can be used in statements that are presented as generalisations, pragmatic factors telling the listener or reader that they actually refer to a specific situation in reality.

That impersonal pronouns have a special relation with the speaker is not a new observation. Dahl (2000) takes Swedish *man* as an egophoric expression, that is, an expression that typically refers to the speech act participants. Moltmann (2006) argues that British *one* is a particularly speaker-oriented pronoun, as generalisations with *one* are usually based on a first person experience. However, Malamud (2006:118) points out that “[s]ince the speaker must have some grounds for uttering generalizations, the generalizations with *one* are often taken to be made on the basis of the speaker’s own experience” and argues that this is rather a pragmatically-driven tendency than a grammatical requirement. This is in line with my own finding that *je* sometimes refers specifically to a third person instead of the first person. This was also explained as a pragmatic influence: speakers are quite egocentric and generally talk about themselves most of the time. In case that they do use *je* to refer to someone other than themselves, they will strongly empathise with this person and take her perspective. As pragmatics play an important role in the self-reference of non-Dutch impersonal constructions too, we would expect that non-first person specific reference also occurs in those languages. And indeed, there are some indications that this is the case.

At first, Paardekooper (1991) does not only say that Dutch *men* used to have speaker-reference in certain contexts, but also that it could refer to the addressee or a third person. Similarly, Zifonun (2001:242) argues that German *man* can, dependent on the context, have all referents of the pronominal paradigm. Importantly, she adds to this observation that *man* is never really synonymous to the personal pronouns, as “mit der Verwendung von *man* stets der Effekt der Typisierung oder Anonymisierung verbunden ist, der mit dem Gebrauch entsprechender Personalpronomina nicht erzielt würde” (‘the use of *man* is always connected to the effect of typification or anonymisation, which would not be achieved by the

use of the corresponding personal pronouns'). The same has been observed for French *on*. In the following example, the speaker uses *on* to refer to her addressee, who is a habitual fellow cross-puzzler.

- (68) Comme ça **on** fait des mots croisés
like that ON does DET words crossed
 sans nous attendre?
without us wait
 'So one is doing crossword puzzles without waiting for us?'
 [Laberge & Sankoff:421 (3)]

Siewierska (2007:21) argues that in the case of more direct specific reference of impersonal reflexive constructions, speaker-reference is typically associated with declaratives, while questions trigger an addressee-referring interpretation. The following examples from Polish illustrate this mechanism.

- (69) Proszę nie przerywać. Mówi **się**.
please not interrupt speak SIE
 'Please don't interrupt. I'm speaking.'
 [Siewierska 2007:21 (44)]
- (70) Co **się** robiło na przerwach?
what SIE did on breaks
 'What did you usually do during the break?'
 [Siewierska 2007:19 (38)]

Again, pragmatic factors determine how the construction is interpreted.

4.1.3 Conclusions

The examples that I gave in this section illustrate that speaker-reference of impersonal constructions is crosslinguistically common. This implies that the link between generic and self-reference cannot be language-specific and coincidental; rather, we should say that speakers have, apparently, a preference for a less direct way of referring to themselves, at least in certain situations.

We have also seen that in some languages, the impersonal construction can be used in episodic contexts and have more direct

reference to the first person (singular or plural). In this use, it can even become the unmarked pronoun for ‘I’ or ‘we’, losing its stylistic connotation. This has happened in French, and to a much lesser degree in some varieties of Swedish. Still, *on* and *man* keep their generic flavour in generic contexts, and this use strongly reminds us of self-referring *je*. This similarity suggests that the conclusion drawn in Chapter 3, that self-referring *je* is in fact generic *je*, is justified. The following question is what happens with second person pronouns in other languages than Dutch. This issue will be discussed in the next section.

4.2 Second person pronouns, genericity, and the first person

In Chapter 3, I presented an account of self-referring *je* which was based on data from the CGN. Self-referring *je* turned out to be a variant of generic *je*. At the same time, a deictic flavour is still present, which is not surprising in the light of the idea that generic *je* must, at some point in time, have evolved from the originally addressee-referring pronoun. I argued that a second person pronoun is suitable for generic reference on the basis of the conditional, which makes it possible to reason from the addressee to anybody else. One piece of evidence for the natural link between the second person and generic reference was found in the data: 8% of the *je* tokens were not clearly deictic or generic, but seemed to be ambiguous instead. An important question now is whether it is crosslinguistically common for second person pronouns to refer generically. And if it is, we will examine these pronouns more closely in order to find out if self-reference of second person pronouns is also common.

4.2.1 Second person pronouns and generic reference

Let us first consider some probability implications. If the generic use of second person pronouns does not occur in other languages than Dutch, the generic use of *je* is possibly accidental. If it is common in related languages whereas it does not occur in unrelated languages, it might be a consequence of language contact, and it can still be accidental. However, if it is crosslinguistically common,

there must be something about second person pronouns that makes them suitable for generic reference.

Siewierska's (2004) typological study on person reveals that it is quite common for second person pronouns to have the possibility of generic reference. In Europe, it is even extremely common: it occurs both in Indo-European languages and in non-related languages like Finnish and Hungarian. However, it occurs outside Europe as well, e.g. in Hindi and Kurdish, and also in completely unrelated languages like Mandarin Chinese, Modern Hebrew, Godié, Tuvaluan, and Macushi. We can thus safely conclude that the generic use of Dutch *je* is not a quirk of Dutch.

There are, however, several comments to be made on the generic use of seconds. At first, according to Siewierska, only the second person singular tends to be used generically; if a non-singular form can have generic reference, it is generally a polite singular form as well. This is the case, for instance, in French and Russian. In both languages, the second person singular pronoun (*tu* and *ты*, respectively) can have generic reference, and its polite counterpart (*vous* and *вы*, respectively), which is a plural form as well, can also have generic reference. As far as I know, the generic use of the Dutch polite second person pronoun *u* has never been attested. Still, it may be used sometimes, but it is probably infrequent.

Secondly, it must be noted that the languages in which seconds can have generic reference vary in the applicability of this usage. For instance, generic seconds are, in some languages, restricted to certain types of discourse or specific (often informal) situations. There are also differences in frequency of usage. Siewierska claims that it is more frequent in English than in any other language. However, we have seen that in Dutch, *je* is used deictically in only half of the cases. Siewierska does not give any evidence for her claim, and as far as I know, no other countings of the deictic and generic use of *je* are available; possibly, the generic use of English *you* is not really more frequent than the generic use of Dutch *je*. Still, it is not unlikely that English and Dutch are extreme in this respect, since many languages have an extra impersonal pronoun that is frequently used, whereas both English *one* and Dutch *men* are highly formal. Thus, generic seconds may be used in more contexts in English and Dutch than in other languages, because English and

Dutch do not have another impersonal pronoun which is appropriate while many other languages do.

Interestingly, the use of seconds has been increasing in European languages during the 20th century. Although this increase is often attributed to the influence of English, many authors who examined this development more closely have rejected this explanation, or pointed out that the influence is not as big as believed (e.g., Helasvuo & Laitinen 2006 for Finnish; Weerman 2006 for Dutch; Jensen 2009 for Danish). One of the explanations for the rise of generic seconds, which often goes hand in hand with a decrease in the use of the original impersonal pronoun, is that seconds are syntactically more flexible than impersonals: as pointed out in Chapter 2, seconds have oblique and possessive forms and occur with all kinds of predicates, whereas impersonal constructions are often restricted in these respects. Another explanation, which focuses on the semantic difference between the two types of constructions, will be discussed in section 4.3.2.1. One may suppose that seconds have the disadvantage that they cannot refer arbitrarily, because they are restricted to generic contexts; however, third person plural pronouns usually take over the arbitrary function that impersonals often have in episodic contexts, so this is not really a disadvantage.

Finally, we must keep in mind that although occurring in genetically and areally unrelated languages, generic seconds are not *extremely* common, that is, in several languages they do not occur, e.g., Japanese and Korean (Kitagawa & Lehrer 1990). Thus, although seconds pronouns are pretty suitable for generic reference, they certainly do not automatically have this possibility.

4.2.2 *Generic use of emphatic forms*

The generically referring second is not necessarily overt. It has even been claimed that it is obligatorily dropped in pro-drop languages; an overt second in such a language would automatically receive deictic interpretation (e.g., Weerman 2006:18). Similarly, it has been suggested that generic seconds cannot be stressed (e.g., Whitley 1978), and for a language like Dutch, which has a distinction between strong and weak personal pronouns, that strong pronouns cannot be used generically, that is, *jij* would only refer

deictically (Malamud 2006:48). In the light of the function of impersonal reference, these claims are not really surprising. As was suggested in Chapter 3, impersonal constructions can be used as a defocusing strategy. If this is the case, it seems unlikely that the emphatic form of the pronoun is chosen. Yet, none of the claims is true.

Whereas it may be true that third person plural pronouns with impersonal reference are obligatorily dropped in pro-drop languages (Cabredo Hofherr 2006), this does not automatically go for seconds as well. This is shown for Hungarian in (71) and for Italian in (72). In Greek, however, pro-drop is obligatory in case of generic reference, as shown in (73).

- (71) (**te**) jól tudsz enni abban a
 (you) well can.2.sg eat in.that the
 vendéglöben.
 in.restaurant
 ‘You can eat well in that restaurant.’
- (72) (**tu**) reagisci instintivamente in casi come
 (you) react.2.sg instinctively in case like
 questi.
 such
 ‘You react instinctively in a case like that.’
 [Kitagawa & Lehrer 1990:754 (46)]
- (73) (***esu**) mporis na fas kala se ekino
 (*you) can.2.sg to eat.2.sg well in that
 to estiatorio
 the restaurant
 You can eat well in that restaurant. (* generic)

In both (71) and (72), the overt pronoun is optional in both the deictic and the generic interpretation. In (73), it is optional in the deictic, but precluded in the generic reading. Thus, the possibility of an overt generic second in a pro-drop language is language-dependent.

Whitley (1978:24) illustrates his claim of what he calls the ‘stressophobia’ of impersonal pronouns with example (74a). In

(74b), the Dutch counterpart of this sentence is given (as *je* cannot have stress¹⁹, the strong counterpart is used here).

(74a) It's amazing what even a short break can do for ***yóu** / you.
(* generic) [Whitley 1978:24]

(74b) Het is ongelooflijk wat zelfs een kleine pauze voor ***jóu** / je kan doen.
'It's unbelievable what even a small break can do for **yóu** / you.' (* generic)

A generic reading of the stressed *you* and *jou* is not possible, whereas a deictic reading of all variants is fine. However, Bolinger (1979) shows, in a reaction to Whitley, that in English, generic *you* can be stressed in contexts of contrasts. I give two of his examples and their Dutch counterparts.

(75a) **You** can't expect sympathy if **yóu** do it, only if **somebody else** does it.
[Bolinger 1979:195]

(75b) **Je** kunt geen sympathie verwachten als **jij** het doet, alleen als **iemand ánders** het doet.
'**You** can't expect sympathy if **yóu** only do it, if somebody **else** does it.'

(76a) I've felt the same way sometimes. It's all right for the professor to ignore the nobodies, but when the one he ignores is **yóu**, that means he's unfair.
[Bolinger 1979:195]

¹⁹ In very specific contexts, *je* can have stress; an example is given in Klaasse (2005):

- (ii) **Jé** hebt niet altijd behoefte aan gezelschap? Spreek voor jezelf!
'**Yóu** do not always feel the need for being in company? Speak for yourself!'

Here, we see that an apparent instance of self-referring *je* is not accepted by the speaker of (ii). The effect of this stressed *je* is ironical or mocking. However, this use is marginal.

(76b) Ik voel me soms ook zo. Het is niet erg als de professor de sukkels negeert, maar als **jij** degene bent die wordt genegeerd, dan is het niet eerlijk.

‘I feel the same way sometimes. It’s all right when the professor ignores the dorks, but when it one who is ignored is **yóu**, it’s unfair.’

(75b) and (76b) already show that the Dutch strong pronoun *jij* can be used generically. However, for the generic use of the unstressed *jij*, the context does not have to be as emphatically contrastive as is the case in the examples above. Weerman (2006:8) already noted that some speakers, e.g., soccer players and their coaches, sometimes use *jij* for generic reference, but he states that this use is substandard. In order to test my intuition that generic *jij* is more widespread than assumed, I have taken all *jij* tokens from the interviews with teachers of Dutch in the CGN. The tokens were annotated for basic reference, as happened earlier for the *je* tokens from Corpus II. The results are represented in Table 4.

Table 4. Basic reference of *jij* in interviews with teachers of Dutch

	Absolute	Relative
Deictic	50	51%
Non-deictic	33	34%
Ambiguous	15	15%
Total	98	100%

Apparently, generic *jij* is fairly frequent. I give two examples: in (77), *jij* is clearly contrastive and in (78), there is no clear contrast.

(77) ja **jij** wordt een jaar ouder maar de leerlingen niet.

‘Well, **you** are getting old, but the pupils are not.’

[INT, fn000108.170]

(78) in dat vakje word **je** dus geplaatst en dan kun **jij** heel hard roepen van ik ben niet geschikt voor die vorm van onderwijs maar als daar uren zijn en ergens anders zijn geen uren dan heb **je** te kiezen of helemaal geen uren of daar inderdaad dus lesgeven.

‘**You** are forced into that kind of job and then **you** can call out loudly ‘I don’t suit in that type of education’ but if they have hours for you and there aren’t any hours somewhere else, then it’s up to **you**: either having no hours at all, or teaching there.’
[INT, fn000096.225]

Interestingly, both *jij* tokens could even be considered as self-referring.

In order to compare the amount of non-deictic tokens to the *je* tokens from Corpus II, Table 2 is repeated here as Table 5.

Table 5 Basic reference of *je* in Corpus II

reference	absolute	relative
deictic	247	53%
non-deictic	183	39%
ambiguous	37	8%
total	467	100%

We now see that generic *jij* is not only quite frequent, but also that the proportions of basic reference of *jij* and *je* are in fact rather similar to each other. It should be noted, however, that eleven *jij* tokens were uttered by one speaker, so this use may be more idiosyncratic than the generic *je*. Moreover, *jij*, deictic or generic, is overall much less frequent than *je*. Still, generic *jij* is used by 18 out of 80 speakers. Importantly, the generic use of *jij* usually seems odd when used out of the blue, but in context, it is not that weird at all.

4.2.3 *Second person pronouns and the first person*

It has turned out that generic reference of second person pronouns is non-accidental. As first person reference of generic constructions is crosslinguistically common, it is unlikely that the self-referring second is a typically Dutch phenomenon.

Indeed, it is not. I have found several observations of self-referring seconds in English (Scheibman 2001, 2007; Hyman 2004; Mildorf 2006), but also in French (Laberge & Sankoff), Finnish (Laitinen 2006), and Modern Hebrew (Perez & Tobin 2009). In all of these languages, seconds are frequently used with generic

reference. In the following sentences, I give an example from each of these languages.

- (79) When **I** got [to Oxford] **I** think the first thing **I** learned was that for the first time in **my** life **you** were totally divorced from **your** background.

[English; Margaret Thatcher, cited in Hyman 2004:161 (1)]

- (80) La famille **je** les reconnais presque tous
the family I them recognise almost all
 parce qu' ils viennent au bureau **tu** les
because they come to.the office you them
 reconnais tsé **tu** les vois arriver puis
recognise you.know you them see arrive next
tu dis ca ca c'est de la famille.
you say that that that.is of the family

'The family – I recognise almost all of them because they come in to the office. You recognise them, you know, you see them coming and you say, that, that's gotta be family.'

[French; Laberge & Sankoff 1979:437 (37)]

- (81) **minä** esimerkiksi suuressa pihmisjoukossa ni mul herää aina semmoset niiku niiku hirveet varmistelureaktiot ja semmoset reaktiot oikeastaan mitkä niinku et **mie** niinkun hirvittävästi oikeastip pelkään em mie nyt ninkur rehellisestis sitä välttämät itelleni aina myönnä

'for example **I**, in a big crowd of people, I always sense such, like, awful feelings of insecurity and in fact such reactions like that; **I** am awfully, I am truly afraid, I don't always admit it, like, honestly to myself, necessarily,'

mut loppujel lopuks se om pelkoa siitä
but end end it is fear it
 et itej **joutu-u** naurunalaseksi ja et
COMP self get-3.SG laughable and that
 se porukka osotta-a **sinua** niinku
the crowd point-3.SG you like
 sormella et ha et **sä oot**
finger COMP ha COMP you be.2.SG
 tyhmä.
stupid

‘but after all, it is the fear that one oneself will be laughed at and that the crowd is, like, pointing their finger at you, that hah that you are stupid.’

[Finnish; Laitinen 2006:222-225 (20a-b)]

(82) SP1 Hmm, okay. And how, how did **you** feel, what were **you** thinking in the moment that **you** got onto a bus the first time afterwards?

SP2 The truth, in, on that line specifically, when **I** switched over from that, a little, **you** know, **you** are looking, **you** are on edge, on edge.

[Hebrew; Perez & Tobin 2009:94]

(original Hebrew text not given in source)

In all examples, we can use the same cues as in the case of Dutch *je*: we see pronoun switches in (79), (80), and (81), a specific interviewer question in (82), and on the basis of common knowledge, we know that in all cases, a specific situation is generalised. In (81), the speaker initially switches from a first person pronoun to a zero construction, and next, she uses a second, probably because she needs an oblique form here, for which the zero construction cannot be used.

In short, it can be concluded from the examples given that self-reference of generically used seconds in different languages is both similar to indirect self-reference of other impersonal constructions and to self-reference of Dutch *je*: the speaker generalises over her own specific experience, abstracting from the particular aspects of her own situation.

4.2.4 Conclusions

We have seen that the generic use of *je* is not some weird quirk of Dutch. On the contrary, seconds can refer generically in several languages, both related and unrelated to Dutch. This use is virtually restricted to singular forms, but polite forms can also have generic reference. The frequency and applicability of generic seconds are language-dependent. The same goes for the possibility of emphatic forms with generic reference. In Dutch, the strong second person

pronoun *jij* turned out to have this possibility, though this use is less common than the generic use of *je*.

It was also shown that self-referring seconds are not typically Dutch: they were found in English, French, Finnish, and Hebrew. We have seen that the mechanism at work those constructions is the same as the mechanism in Dutch.

4.3 Impersonals versus seconds

In Chapter 3, I analysed Dutch self-referring *je* as an instance of generic *je* with traces of the original second person meaning. The generic aspect makes a distancing effect possible, whereas the more covered second person reference invites the addressee to imagine herself to be in the speaker's position, which will lead to empathy. Now that we know that self-reference of generic constructions is crosslinguistically common, it would be interesting to see if the intuitions on this use in other languages are similar to those of the Dutch authors and myself. If they are, it is expected that impersonals have this distancing effect, whereas they lack the involving effect that seconds have. In this section, I will focus on the suggested motivations and effects and the possible differences between impersonals and seconds. At first, I will discuss the effects of self-referring impersonal constructions. Next, the effects of generic and self-referring seconds will be discussed.

4.3.1 Impersonal constructions: effects of genericity

Generally, the use of impersonal constructions is associated with creating distance between the speaker and the message. However, both the assumed speaker motivation and the perceived effect of this mechanism vary: they may depend on the language, on the context, and presumably also on the hearer.

When an impersonal pronoun is felt as an instance of self-reference, it is often perceived as a way of avoiding responsibility (e.g., Jisa & Vigué 2005:128 for French *on*; Zifonun 2001:243 for German *man*; Ragnarsdóttir & Strömquist 2005:147 for Swedish *man*). Not surprisingly, the reactions to this use are often quite negative. For instance, self-referring English *one* is considered to be arrogant and pompous (Siewierska 2004; Altenberg 2004/5), and

Zifonun (2001:242) assumes German *man* to be “eine (manirierte, verhüllende oder distanzierende) Bezeichnungsalternative zu *ich*” (‘an (affected, concealing or distancing) alternative for *I*’), which makes a strategic or even highly manipulative use possible.

The use of impersonal constructions for self-reference is also associated with contexts of modesty, which is based on the idea that the avoidance of the first person pronoun is iconic. According to Egerland (2003:74), Swedish *man* has, at least for some Swedish speakers, a flavour of modesty; they consider *man* to be a communicative strategy “when the speaker does not wish to emerge as the focus of attention”. Modesty is even the main explanation for the self-referring use of the zero construction in Mandarin Chinese (Siewierska 2004). As for Finnish, it has been suggested that the zero construction for self-reference is a politeness strategy (Leinonen 1983; Hakulinen 1987). More recently, Laitinen (2006) and Sorjonen (2001) have argued that rather, this use should be considered as a manifestation of simultaneous specific and non-specific reference. In other words, the speaker merges the personal and the general, which she does in order to create subjectivity and personal involvement. This account is more in line with my intuitions on self-referring *je*.

It is remarkable that the more negative effects (avoiding responsibility, manipulation, arrogance) are associated with impersonal pronouns, while the more positive effects (modesty, politeness, involvement) are mainly associated with zero constructions. As the use of a pronoun is probably more marked than the absence of an overt form, the former may suggest more emphatically that a generalisation is made here, which can make the hearer more ‘suspicious’ (“is this really generalisable?”), whereas the absence of a pronoun does not, or to a lesser extent, have this implication. However, more information on the effects of both overt and zero pronouns in impersonal constructions is needed, and this remains a question that could find an answer in the future.

In sum, self-reference of impersonal constructions is generally claimed to be a defocusing strategy, which leads to a more distanced way of talking about oneself, but whether this strategy is considered to be positive or negative varies among hearers. These intuitions, though apparently contradicting, are in line with the intuitions on

Dutch self-referring *je* and with its analysis as a generic pronoun, which is possibly but not necessarily used strategically.

4.3.2 *Seconds: generalisation and involvement*

Now, we turn again to the seconds. First, I will discuss the intuitions on the addressee-referring origin of generic seconds. Then, the accounts of self-referring seconds in English and in Hebrew will be discussed, and they will be compared to the intuitions in self-referring impersonal constructions.

4.3.2.1 The origin of generic seconds

The idea that generic seconds retain something of their original deictic flavour is not new. Although some authors consider generic seconds and other impersonally used personal pronouns to be completely unconnected to their personal use (e.g., Whitley 1978; Laberge 1980), most others agree that there is in fact a clear connection. As for English *you*, I know of two authors who have tried to make the link between its generic and deictic use explicit. According to Bolinger (1979), the impersonal use of personal pronouns is in fact metaphorical, and the impersonal use of *you* is a courtesy device. In his example, the speaker instructs her addressee how to make a kite, saying ‘you do it like this’. Using *you* is, according to Bolinger, the most courteous way to give this instruction: with *I*, the speaker would insist on her own role too much, and with *one*, she would flaunt her knowledge. Although this may be the effect indeed, it does not really explain why speakers started using *you* generically, as it does not only, or even typically, occur in instructions.

Hyman (2004:166) gives two alternative explanations. One of them is similar to the one that I have suggested, which he calls “an elliptical or truncated conditional expression”. The other explanation is that *you* might be a kind of inclusive *we*, which fills the gap in English of a distinction between inclusive and exclusive *we*. However, he notes himself that generic *you* does not always naturally include the addressee. As for Dutch *je*, we have seen that it sometimes rather refers to a third person than having genuine generic reference. Moreover, the use of generic *je* and *you* is limited

to generic contexts and it would often not be possible to replace *we* for *you* in order to mark the inclusive meaning; that is, it would only fill a small part of the gap. Thus, the conditional as a source remains the only satisfying explanation.

Apart from Bolinger and Hyman, none of the authors who wrote about generic seconds have attempted to explain why a second person pronoun can refer generically. However, several authors have opposed the generic second in a certain language to its impersonal counterpart (e.g., Bolinger 1979 and Malamud 2006 for English *one*; Jensen 2009 for Danish *man*) and the intuitions are very similar: there is a perspective difference (seconds *show* the situation from within, while impersonals *tell* how it is from outside), which is caused by the inherent invitation to the addressee to put herself into someone else's shoes, and which appeals to the addressee's involvement and feelings of empathy. Jensen takes this difference to be the cause of the increase of *du* at the cost of *man*.

As for Finnish, the intuitions deviate: Laitinen (2006) characterises the zero construction perspective as more 'from within' than the generic second. This is again an indication that the zero construction is semantically different from impersonal pronouns.

4.3.2.2 The effects of self-referring seconds

Now, we focus on self-referring seconds again. In section 4.2.2, I gave an example of self-referring seconds from four languages: English, French, Finnish, and Modern Hebrew. As for French, no explanation for or effect of this phenomenon is discussed. For Finnish, the use of the zero construction and the generic second are not considered to be very different. Thus, Laitinen's (2006) explanation of merging the personal and impersonal, as discussed in section 4.3.1, can be extended to the Finnish generic second.

The accounts of English self-referring *you* are more detailed. Hyman (2004), who does not typically focus on self-referring seconds but rather on the non-deictic use of *you* in general, is the only one to suggest that modesty might play a role. Mildorf (2006) claims that self-referring *you* (which she calls 'doubly-deictic *you*') is generic in the first place, which creates the possibility for distancing, and, besides, keeps its deictic flavour, which creates

involvement, as the addressee is invited to identify with the speaker's situation. This ambiguity makes it possible to use *you* strategically, for example, in order to avoid criticism. Scheibman (2001, 2007) focusing on pronoun shift from *I* to *you*, suggests two motivations. On the one hand, she calls this use of *you* "iconic distancing", following Abney (1996)²⁰, who claims that pronoun shifting can be used as a distancing technique in conversations on uncomfortable topics. On the other hand, she claims that generalisations in discourse have a solidarity marking function. Generalisations in this context are assumed to be based, usually, on personal and social expectations and beliefs, instead of being inductively constructed. By switching from *I* to *you*, the speaker is generalising from her own stance: she appeals to societal norms, and thus to the beliefs of her conversational partners. Formulating a claim as a generalisation thus has a broadening and an inclusive effect. Such a shift can be regarded as a strategy "to build empathy with other participants" (Scheibman 2007:132). This effect is strengthened by the deictic origin of *you*, which results in implicit reference to the addressee. Thus, according to Scheibman, the involving effect that self-referring/generic *you* has is not only due to its second person origin, but also to its generalising force.

A highly interesting case is Hebrew, in which the second person singular pronoun *atah* is used for male referents, while *at* is used for feminine addressees. According to Perez and Tobin (2009), both forms are used for generic reference, but the feminine form is only used in clearly feminine contexts and situations, e.g., in a sentence like 'when you are pregnant, then you ...' or in a group consisting of women only. The masculine form is the default, which is reflected by the fact that it can be used even in 'when you are pregnant, then ...'.

Perez and Tobin discuss the use of seconds in interviews with bus drivers who experienced a terror attack. The interviewee in the examples is male; the interviewer is female. I give two examples. In (83), the feminine form of the generic second is used. The second example was already given before in section 4.2.2, here repeated as

²⁰ Unfortunately, I was not able to find this paper and therefore I cannot give more details of Abney's analysis.

(84). In both examples, the seconds are marked for gender, with ‘F’ for feminine and ‘M’ for masculine.

(83) **You.F** see the bus, after **I** went to the [bus company] branch, because my bag was still there, everything, they took my bag. Two weeks later, **I**, for no particular reason, am getting [my things] organized, **I** see tons of glass shards inside the bag. Tons of glass. In the end, and **you.F** see the bus, God help us, and it is full of holes...

[Perez & Tobin 2009:92]

(original Hebrew text not given in source)

(84) SP1 Hmm, okay. And how, how did **you** feel, what were **you** thinking in the moment that **you** got onto a bus the first time afterwards?

SP2 The truth, in, on that line specifically, when **I** switched over from that, a little, **you.M** know, **you.M** are looking, **you.M** are on edge, on edge.

According to the authors, the use of the feminine pronoun in (83) can be seen as a communicative strategy to draw in the interviewer to the account of the experienced event, probably to create a shared understanding of the experience. In other words, it is used to invite the addressee to be involved and have empathic feelings for the speaker, just as we saw before. However, the use of the masculine pronoun in (84) is perceived as more detached and universalised: the speaker probably wants to ‘collectivise’ his difficult feelings and experiences. According to the authors, it is unclear to what extent the speaker’s specific situation is generalised: it may go for all Israelis (including the interviewer), but also for male Israeli bus drivers (excluding the interviewer). As such, the interviewer is not, or at least to a lower degree, drawn into the story as was the case in (83).

We can draw two conclusions from the Hebrew examples. At first, generic seconds can indeed have the involving effect that is often assumed to be a consequence of its original deictic flavour. In Hebrew, this effect appears when a male speaker uses the feminine second in a conversation with a female addressee. Second, however, generic seconds seem to be able to lose their second person reference, at least to a certain extent.

4.3.2.3 From second to first person pronoun?

The last question is: can seconds undergo the same development as the French impersonal *on*, that is, can they turn into a real first person pronoun (singular or plural) that is used in episodic contexts? Or, put differently, is it possible that Dutch *je* will replace *ik*?

Of course, we can only speculate on the answer on the latter question and I will only share some considerations. It is remarkable that there are several languages in which the impersonal pronoun can refer unambiguously to an 'I' or 'we', whereas this does not seem to happen with seconds. According to Helmbrecht (2002), though, there is a language which has a first person plural pronoun which has evolved from a second: the Algonquian language Wiyot. This may be an indication that it can happen, but it is only a very small one: firstly, it is only one language, and secondly, Helmbrecht does not specify the development. Thus, we do not know whether the Wiyot second has gone through a stage of generic reference first.

A second consideration is the ambiguity problem: an impersonal can unambiguously refer to the first person, because there are syntactic restrictions on the generic and arbitrary interpretation; in contrast, the second has its original deictic reference, and therefore, a second would always be ambiguous between a first and second person interpretation if it occurred in episodic contexts. Yet, according to Lukman (2009), the pronoun *awak* in the Austronesian language Jambi City Malay can refer to both the first and the second person, dependent on context. Although ambiguity between a first and second person interpretation does not seem to be an optimal situation, the use of *awak* does not typically lead to problems. This is possibly due to a fact pointed out before in section 3.2.2: speakers do not typically tell their addressee much about herself, because the latter is supposed to know that better than the speaker does. However, it must be noted that Jambi City Malay has a large range of first and second person pronouns which differ in politeness, which makes the situation very different from the situation in Dutch and other European languages.

As for seconds and episodicity, we can say that self-referring seconds usually occur in non-episodic contexts, which seems to give the utterance the appearance of a generalisation. Still, in the

literature, there are some examples of an utterance with *je* that is rather episodic than generic. A very clear example is given in (85).

- (85) **Je** kreeg de bal van Huntelaar en toen schoot **je** op het doel.
'**You** (I) got the ball from Huntelaar and then **you** (I) targeted on the goal.'
[Zeijlstra 2010]

As (85) is episodic, it is hard, or even impossible, to interpret *je* generically here. Without context, a deictic interpretation is the most obvious one. However, the context reveals that the speaker refers to himself. Interestingly, the utterance still has the flavour of a generalisation: the speaker seems to suggest that the addressee, and anyone else, would have done the same.

4.3.3 Conclusions

The intuitions on the effects of generic seconds mainly support my analysis of self-referring *je*. However, it turns out to be hard to say which aspect of the pronoun leads to which effect on the listener. The Hebrew examples show that a more direct reference to the addressee (a feminine form in a conversation with a female addressee) has an involving effect, whereas this is less so in the case of a more indirect reference (a male form and a female addressee). Yet, there is a perspective difference between seconds and impersonals that is most naturally explained as a consequence of second person traces that are still present. A more systematic study on the production and perception of seconds versus impersonals would possibly reveal the mechanisms that underlie both types of generic pronouns.

Finally, I have focused on the possibility of direct self-reference of seconds and the occurrence of Dutch self-referring *je* in episodic contexts.

4.4 General conclusions

In this chapter, I placed the case of Dutch self-referring *je* in a broader context. It was shown in section 4.1 that impersonal constructions often refer to a specific person, who is usually the

speaker. This reference can be more or less direct. In the case of direct reference, the referent is unambiguously the speaker or a group including the speaker, because the generic and arbitrary reading of the pronoun are ruled out by syntactic factors. In French, the impersonal pronoun *on* has even largely taken over the function of the original first person plural pronoun *nous*. However, the group of languages having this possibility is limited. Still, a more indirect kind of reference can occur, which is very similar to self-referring *je* in that it gives the sentence the appearance of a generalisation, while contextual cues lead the hearer to the idea that the speaker is actually or mainly talking about herself. As both phenomena occur in related and unrelated languages, the conclusion was drawn that speakers tend to generalise their specific situation, apparently regardless of the particular language they speak.

In section 4.2, I focused on the generic and self-referring use of seconds. Generically-referring seconds turned out to be crosslinguistically common, although languages vary in frequency and applicability of this use. Self-referring seconds have been observed in a number of languages. This use was again similar to the use of self-referring *je*.

In section 4.3, I discussed the suggested motivations for and effects of self-referring generic constructions and the differences between impersonals and seconds. As expected, the effect of distance was observed for both types of constructions, but the effect of involvement was only found for seconds and it was generally ascribed to the addressee-referring origin. Although there are some indications that nuance this idea, the intuitions on the difference in perspective between impersonals and seconds are indicative, and it is likely that the appeal on the addressee to evoke empathic feelings is stronger in seconds than in impersonals. Finally, I discussed the possibility of seconds undergoing a development similar to French *on*. Although this question did not find an answer, it was pointed out that we cannot immediately rule out this possibility.

Conclusion

The starting point for this thesis was the observation that the second person pronoun *je* sometimes has first person reference. This phenomenon was called ‘self-referring *je*’. I have tried to answer the following questions: what is it in *je* that makes it possible to have first person reference, and why do speakers use this form?

In Chapter 2, I shortly discussed *je* as a deictic and as a generic pronoun. Next, I gave an overview of the accounts of self-referring *je* that are present in the Dutch literature. I showed that the accounts differ in several ways. At first, some authors argue that self-referring *je* is typical for soccer players. However, the assumption that soccer players started this use of *je* turned out to be false: self-referring *je* was found to occur before soccer players were commonly interviewed for radio and television. This also ruled out the rise of personal questions in the 80s and 90s of the 20th century as a possible explanation. Secondly, the authors differ in whether they assume deictic *je* or generic *je* as a starting point for self-referring *je*, and third, they differ in their suggestion of a possible speaker’s motivation for the use of self-referring *je*. Although none of the accounts turned out to be satisfying, one intuition, shared by three different authors, was found to be interesting: the idea that speakers use self-referring *je* to evoke the addressee’s empathy.

The first limitation of the Dutch literature on self-referring *je* is that none of the accounts were based on empirical data. I therefore examined this phenomenon on the basis of corpus data in Chapter 3. This study revealed that self-referring *je* is an instantiation of generic *je*: although an utterance with self-referring *je* refers to a specific situation and a specific person (namely, the speaker), it still has a generalising flavour. Thus, *je* can give a situation the appearance of a generalisation, while it actually refers to a specific situation. Whether the hearer interprets *je* as more generic or more self-referring is a subjective matter: it depends on contextual factors

and her own frame of reference. If a hearer recognises the situation from her own experience, she will be more inclined to accept the utterance as a true generalisation.

The specific use of *je* gives speakers the opportunity to avoid the first person pronoun *ik* for some reason. This usage can be, but is not necessarily a strategic manoeuvre: sometimes, *ik* is not appropriate because it lays more focus on the speaker than she intends to express (for instance, because she wants to focus on someone else). *Je* can then be a good alternative for *ik*, because it is a way of defocusing from the agent or subject. Generalisations are frequent in informal discourse, because they have an inclusive effect: they are used to mark agreement among conversational partners. That speakers often take themselves as the point of departure for a generalisation is not surprising, since speakers are known to talk mostly about themselves. However, *je* sometimes refers specifically without having first person reference, referring to a third person instead. Intuitively, the mechanism of self-referring *je* is at work here, which indicates that the reason why *je* refers to the speaker so often is a pragmatic one.

Generic *je* turned out to be highly context-dependent: it presupposes a conditional, either explicit or implicit, which determines the actual referent. Usually, it refers to only a small subset of ‘people in general’. The conditional was argued to be the historical step from deictic to generic *je*. This explains why speakers are thought to use self-referring *je* in order to evoke empathic feelings: generic *je* still has traces of its deictic origin, that is, it supposes an imaginary addressee, which makes the real addressee imagine herself to be in the speaker’s situation. If the addressee accepts this hidden invitation, empathy will arise.

The second limitation of the Dutch literature is that Dutch self-referring *je* has not been related to similar phenomena in other languages. In Chapter 4, I compared my findings to what is known about impersonal constructions in other languages than Dutch. Self-referring impersonals turned out to be crosslinguistically common, both in related and unrelated languages. This is evidence for my assumption that the reason that *je* often refers to the speaker is pragmatic. The generic use of second person pronouns is common too, although it is certainly not automatically possible. That Dutch *je*

can refer generically is thus not accidental. As expected, several languages were found to have self-referring seconds.

Self-referring impersonals and self-referring seconds are similar in that a specific situation is generalised, but they are also different. Firstly, in some languages, self-referring impersonals can occur in episodic contexts, referring unambiguously to the speaker. In this case, the generalising flavour that the construction normally has, diminishes or even disappears. In contrast, seconds are claimed to refer deictically, that is, to the addressee, in episodic contexts. Although there are reasons to doubt this claim, there are no seconds that are known to be used like, for example, French *on*. Secondly, it is claimed by native speakers of languages which have both an impersonal and a second for generic reference, that impersonals lack the involving effect that seconds have. Thus, seconds do invite the addressee to take the speaker's perspective, whereas impersonals do not. There is an obvious reason for this difference: the fact that seconds originally refer to the addressee, while impersonals are originally generic. Again, this idea must be nuanced a bit: firstly, the generic nature of seconds has been argued to evoke empathy too, and secondly, seconds do not necessarily have this involving effect, as was shown with data from Hebrew. We might rather say that seconds appeal more strongly to the addressee than impersonals do.

Generally, we can say that my own conclusions with respect to self-referring *je* strongly agree with crosslinguistic facts on the phenomenon of self-referring impersonals and seconds. Pragmatics turns out to play a key-role in both the production and the interpretation of *je*, whether it is deictic, generic, or self-referring.

References

- Abney, L. (1996). Pronoun shift in oral folklore, personal experience and literary narratives, or 'What's up with you?'. *SECOL Review*, 20, 203-226.
- Alonso-Ovalle, L. (2002). 'Arbitrary pronouns are not that indefinite'. In: C. Beyssade, R. Bok-Bennema, F. Drijkoningen and P. Monachesi (Eds.), *Proceedings of Going Romance 2000*, 1-14. Amsterdam/Philadelphia: John Benjamins.
- Altenberg, B. (2004/5). The generic pronoun in English and Swedish: Cross-linguistic correspondences of *one* and *man*. *Languages in contrast*, 5, 1, 93-120.
- Appel, R., Hubers, G. & Meijer, G. (1976). *Sociolinguïstiek*. Utrecht: Aula.
- Appel, R. (1993). 'Niet alleen voetballers gebruiken vaak "je" als ze "ik" bedoelen'. In: J. ten Thije (Ed.), "Waar komen de juiste ideeën vandaan?" *Opstellen over taal, wetenschap en maatschappij. Feestboek voor Gerard Hubers*, 95-99. Amsterdam: Instituut voor algemene taalwetenschap, Universiteit van Amsterdam.
- Asser, E. (1955). *U hoort er wel meer van*. Amsterdam: Arbeiderspers.
- Bennis, H., Cornips, L. & Van Oostendorp, M. (2004). *Verandering en verloedering: normen en waarden in het Nederlands*. Amsterdam: Amsterdam University Press/Salomé.
- Bolinger, D. (1979). To catch a metaphor: *you* as norm. *American Speech*, 54, 194-209.
- Cabredo Hofherr, P. (2006). "'Arbitrary Pro" and the theory of pro-drop'. In: P. Ackema, P. Brandt, M. Schoorlemmer, & F. Weerman (Eds.), *Agreement and arguments*, 230-260. Oxford: Oxford University Press.
- Cardinaletti, A. & Starke, M. (1999). 'Pronouns in Germanic and Romance languages: An overview'. In: H. van Riemsdijk (Ed.), *Clitics in the languages of Europe*, 33-82. Berlin/New York: Mouton the Gruyter.
- Cinque, G. (1988). On *si* constructions and the theory of arb. *Linguistic inquiry*, 19, 521-582.

- Condoravdi, C. (1989). 'Indefinite and generic pronouns'. In: E. Fee & K. Hunt (Eds.), *Proceedings of the eighth West Coast Conference on formal linguistics*, 71-81. Stanford: CSLI.
- Coveney, A. (2000). Vestiges of *nous* and the 1st person plural verb in informal spoken French. *Language Sciences*, 22, 447-481.
- Dahl, Ö. (2000). Egophoricity in discourse and syntax. *Functions of Language*, 7, 33-77.
- D'Alessandro, R. & Alexiadou, A. (2003). Inclusive and Exclusive Impersonal Pronouns. Manuscript, University of Stuttgart.
- Daniëls, W. (2006). 125 jaar naar Dik Trom. *Onze Taal*, 75, 2/3, 53-55.
- Egerland, V. (2003). Impersonal man and aspect in Swedish. *Venice working papers in linguistics*, 13, 73-91.
- Hakulinen, A. (1987). 'Avoiding personal reference in Finnish'. In: J. Verschueren & M. Bertucelli-Papi (Eds.), *The pragmatic perspective* (pp. 140-153). Amsterdam/Philadelphia: John Benjamins.
- Helasvuo, M.-L. & Laitinen, L. (2006). 'Person in Finnish: paradigmatic and syntagmatic relations in interaction'. In M.-J. Helasvuo & L. Campbell (Eds.), *Grammar from the human perspective: Case, space and person in Finnish* (pp. 173-208). Amsterdam/Philadelphia: John Benjamins.
- Helasvuo, M.-L. & Vilkuna, M. (2008). Impersonal is personal: Finnish perspectives. *Transactions of the philological society*, 106, 2, 216-245.
- Helmbrecht, J. (2002). 'Grammar and function of *we*'. In: A. Duszak (Ed.), *Us and others. Social identities across languages, discourses and cultures* (pp. 30-50). Amsterdam/Philadelphia: John Benjamins.
- Hogeweg, L. (2009). *Word in progress. On the interpretation, acquisition, and production of words*. PhD thesis. Utrecht: LOT.
- van Hout, R. (2003). 'Hun zijn jongens. Ontstaan en verspreiding van het onderwerp *hun*'. In: I. Stroop (Ed.), *Waar gaat het Nederlands naartoe? Panorama van een taal*, 277-286. Amsterdam: Bert Bakker.
- Hyman, E. (2004). The indefinite you. *English studies*, 85, 161-176.
- Jensen, T. (2009). Generic variation? Developments in use of generic pronouns in late 20th century spoken Danish. *Acta Linguistica Hafniensia*, 41, 1, 83-115.

- Jisa, H. & Vigiúé, A. (2005). Developmental perspectives on the role of French *on* in written and spoken expository texts. *Journal of pragmatics*, 37, 125-142.
- Johansson, A. (2004). Viewing languages through multicultural corpora, with special reference to the generic pronoun in English, German and Norwegian. *Languages in contrast*, 4, 2, 261-280.
- Kitagawa, C. & Lehrer, A. (1990). Impersonal uses of personal pronouns. *Journal of pragmatics*, 12, 739-759.
- Klaasse, B. (2005). *Spreek voor jezelf. Veranderingen in het gebruik van persoonlijke voornaamwoorden met een algemene referentie in het Nederlands in taalvergelijkend perspectief*. MA thesis, University of Amsterdam.
- Laberge, S. (1980). 'The changing distribution of indefinite pronouns in discourse'. In: R. Shuy & A. Shnukal (Eds.), *Language use and the uses of language*, 76-87. Washington: Georgetown University Press.
- Laberge, S. & Sankoff, G. (1979). 'Anything *you* can do'. In: T. Givón (Ed.), *Discourse and syntax*, 419-440. New York: Academic Press.
- Laitinen, L. (2006). 'Zero person in Finnish. A grammatical resource for construing human reference'. In M.-J. Helasvuo & L. Campbell (Eds.), *Grammar from the human perspective: Case, space and person in Finnish*, 209-231. Amsterdam/Philadelphia: John Benjamins.
- Leinonen, M. (1983). Generic zero subjects in Finnish and Russian. *Scando-Slavica*, 29, 143-161.
- Los, B. (2002). 'The loss of the indefinite pronoun *man*; syntactic change and information structure'. In: T. Fanego, J. Pérez-Guerra & M. José López-Couso (Eds.), *English historical syntax and morphology*, 181-202. Amsterdam/Philadelphia: John Benjamins.
- Lukman (2009). *The role of politeness in the competition of Jambi City Malay local pronouns*. MA thesis, Raboud University Nijmegen.
- Malamud, S. (2006). *Semantics and pragmatics of arbitrariness*. PhD thesis, University of Pennsylvania.
- van Marxveldt, C. (1919). *De HBS-tijd van Joop ter Heul*. Amersfoort: Valkhoff.
- Mildorf, J. (2006). 'Sociolinguistic implications of narratology: focalization and "double deixis" in conversational storytelling'. In: M. Hyvärinen, A. Korhonen & J. Mykkänen (Eds.), *Studies across*

- disciplines in the humanities and social sciences 1*, 42-59. Helsinki: Helsinki Collegium for Advanced Studies.
- Moltmann, F. (2006). Generic *one*, arbitrary PRO, and the first person. *Natural language semantics*, 14, 257-281.
- Onrust, M. (1995). Voetbal-je. *Tekst[blad]*, 3, 13-14.
- Paardekooper, P. (1991). Van taal naar tekst, van /mən/ naar /mɛn/. *Taal en tongval*, 43, 69-96.
- Perez, A. & Tobin, Y. (2009). Difficulties of translation from Israeli Hebrew to American English: An analysis of pronouns and power relations in interview with bus drivers who experienced terror attacks. *Graduate journal of social science*, 6, 3, 86-107.
- Ragnarsdóttir, H. & Strömquist, S. (2005). The development of generic *maður/man* for the construction of discourse stance in Icelandic and Swedish. *Journal of pragmatics*, 37, 143-155.
- Ramat, A. & Sansò, A. (2007). 'The spread and decline of indefinite *man*-constructions in European languages'. In: P. Ramat & E. Roma (Eds.), *Europe and the Mediterranean as linguistic areas. Convergencies from a historical and typological perspective*, 95-132. Amsterdam/Philadelphia: John Benjamins.
- Sansò, A. (2006). "Agent defocusing" revisited. Passive and impersonal constructions in some European languages.' In: W. Abraham & L. Leisiö (Eds.), *Passivization and typology: Form and function*, 229-270. Amsterdam/Philadelphia: John Benjamins.
- Saxon, L. (1993). A personal use for Athapaskan 'impersonal' *ts'e*. *International journal of American linguistics*, 59, 343-354.
- Scheibman, J. (2001), 'Local patterns of subjectivity'. In: J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*, 61-90. Amsterdam/Philadelphia: John Benjamins.
- Scheibman, J. (2007). 'Subjective and intersubjective uses of generalizations in English conversations'. In: R. Englebretson (Ed.), *Stancetaking in discourse*, 111-138. Amsterdam/Philadelphia: John Benjamins.
- Siewierska, A. (2004). *Person*. Cambridge: University Press.
- Siewierska, A. (2007). 'Ways of impersonalizing: Pronominal vs. verbal strategies'. In: M. de los Ángeles Gómez González, J. Lachlan Mackenzie & E. González Álvarez (Eds.), *Current trends in contrastive linguistics: functional and cognitive perspectives*, 3-26. Amsterdam/Philadelphia: John Benjamins.

- Sigurðsson, H. & Egerland, V. (2009). Impersonal null-subjects in Icelandic and elsewhere. *Studia linguistica*, 63, 158-185.
- Sorjonen, M.-J. (2001). *Responding in conversation. A study of response particles in Finnish*. Amsterdam/Philadelphia: John Benjamins.
- Sun, J.T.-S. (2005). *Linguistic coding of generic human arguments in rGyalrongic languages*. Paper presented at the 11th Himalayan Languages Symposium, December 6-8, Chualalongkorn University, Bangkok, Thailand.
- Vermaas, H. (2004). *Mag ik u tutoyeren? Aanspreekvormen in Nederland*. Amsterdam/Antwerpen: L.J. Veen/Het taalfonds.
- Wales, K. (1996). *Personal pronouns in present-day English*. Cambridge University Press.
- van Wassenaar, A. (1994). Je = je, ik, we, men. *Onze Taal*, 63, 9, 203.
- Weerman, F. (2006). “It’s the economy, stupid!” Een vergelijkende blik op *men* en *man*. In: M. Hüning, U. Vogel, T. van der Wouden & A. Verhagen (Eds.), *Nederlands tussen Duits en Engels. Handelingen van de workshop op 30 september en 1 oktober 2005 aan de Freie Universität Berlin*, 19-46. Leiden: Stichting Neerlandistiek.
- Whitley, M. (1978). Person and number in the use of *we*, *you* and *they*. *American Speech*, 53, 18-39.
- Zeijlstra, H. (2010). *Let’s talk about you and me*. Talk at the Workshop Local Pronouns, June 7-8, Radboud University Nijmegen, The Netherlands.
- Zifonun, G. (2001). “Man lebt nur einmal.” Morphosyntax und Semantik des Pronomens *man*. *Deutsche Sprache*, 28, 232-253.
- van der Zijl, A. (2004). *Anna. Het leven van Annie M.G. Schmidt*. Amsterdam: Nijgh & Van Ditmar.