

Pronoun Loss as a Form of Deflection

Suzanne Aalberse*

1 Introduction

In the course of the history of English, the forms of the second person singular were replaced by the second person plural, both in the pronominal system (*thou, thee* > *you*) and in the agreement system (*hast* > *have*) (Lass 1999). The use of a second person plural pronoun as a singular form of address was a widespread politeness strategy in medieval Europe (Betsch 2003, Haugen 1976: 303-304, Taavitsainen & Jucker 2003, Mühlhauser & Harré 1990: 145-150). Only in English and Dutch did this politeness strategy lead to the loss of the original second person singular. In this paper I focus on the question why English and Dutch lost the second person singular pronoun whereas other European languages did not. I argue that the loss is not only socio-pragmatically motivated, but also involves a language internal factor, namely deflection. In this paper I follow Brown & Gilman (1960) in the use of the symbol T (derived from Latin *tu*) as a generic marker for the informal singular pronoun and the use of the symbol V (derived from Latin *vos*) as a generic marker for the formal pronoun. In section 2 I evaluate socio-pragmatic explanations for T-loss. Section 3 discusses the effect of the loss of second person singular marking on the verbal paradigm. Section 4 provides independent evidence for the hypothesis that avoiding uneconomic inflection plays a role in T-loss. Finally, section 5 concludes the paper.

2 Socio-pragmatic factors in T-loss

A sociolinguistic factor in T-loss is *identity marking*. *Identity marking* refers to the observation that the choice of address forms reveals something about the social status of a speaker.

*I would like to thank Robert Cloutier, Fred Weerman and PLC 29 for helpful comments. The usual disclaimers apply.

The use of polite forms originated in the higher classes and was much more common in the cities than in rural areas (Brown & Gilman 1960). Wales (1983, 1996) suggests that a new middle class with aspirations copied the habits of the politest society including the use of a V pronoun. More and more people copied this habit of using polite forms to avoid association with a lower class or a rural background. The use of polite pronouns received an extra impulse when a standard spoken language arose. T-forms then became increasingly associated with social unacceptability (Wales 1983: 117), which led to their loss.

Additional pragmatic motivation for T-loss is *markedness reversal*. *Markedness reversal* refers to a change in T from a neutral to a negative address form. Wales (1983, 1996) relates this change to social mobility. The rise of new social groups made distinction of status or rank less easy to determine. In cases of doubt one would rather be polite than risk giving offence and every precedent widened the range of cases where V was used. More and more addressees expected to receive V; receiving T became associated with condescension. T was no longer a neutral address form but a negative one and therefore had to disappear.

A third factor often mentioned in combination with T-loss is *pronoun mixing*. *Pronoun mixing* refers to the observation that one addressee could be addressed by one and the same speaker with both T and V. Sometimes T and V were even used in tandem in one sentence. Berteloot (2001) interprets mixing as a sign that T and V were interchangeable forms. If forms are interchangeable loss of one of these forms follows naturally. Wales (1983: 119) remarks that the fluctuation between T and V could itself be a sign of using an informal register and she too asserts that if both forms are often used together it is easy to imagine that one form might become redundant in the long run.

Apart from the fact not one of the three factors is unique to English and Dutch (compare Brown & Gilman for social factors and Hunt 2003: 47, Simon 2003: 89 for pronoun mixing) and thus cannot explain the unique position of Dutch and English in Europe, there is another problem. *Pronoun mixing*, *identity marking* and *markedness reversal* ignore the domain specific nature of the interpretation of T. This domain dependency becomes more clear if we look at pronoun mixing situations. In contexts where T and V co-occur, the two pronouns are both meaningful. Berteloot (2001) shows that in the text *Legenda Aurea* by Jacobus de Voragine children are always addressed

with T. High ecclesiastic officials always address each other with V. Mary and Christ are addressed with T and V. Busse (2002) shows that in Shakespeare's work the nominal address form *Monsieur* always co-occurs with the pronoun V. The vocative *bully* (used as a term of endearment) always co-occurs with a T-pronoun whereas the term *husband* combines with T and V almost equally.

The situations where only V is used can be characterized as a prototypical V-situation: the speaker-addressee relationship is formal and expressing respect for rank and status is relevant. High officials and the term *monsieur* both fit into the prototypical picture of V-use. In the mixed cases we see a clash between two factors: the relationship with a husband can be characterized as [+intimate] which requires a T-form and as [+status] which asks for a V-form. The same argument goes for addressing Christ and Mary. In the situations where only T is used the speaker is in a higher hierarchical position than the addressee (in the case of the children) or rank is not relevant (in the case of *bully*) and the relationship can be characterized as intimate. It is difficult to imagine how the T pronoun fully disappeared out of these prototypical T-situations. The increased social unacceptability of the T-form in the public domain does not imply an increase in unacceptability in the intimate domain. Moreover addressing a child with T does not entail the same social implications as addressing strangers in the public sphere with this pronoun.

3 Economy in the verbal paradigm via T-loss

Although socio-pragmatic factors can explain domain loss for the T-pronoun, they cannot explain the loss of T out of all domains. The question remains why English and Dutch lost their T-pronoun whereas other European languages did not. In this section I show that one factor that sets Dutch and English apart from most of the other European languages is that in English and Dutch, T-loss results in economy in the verbal paradigm. In my definition of economy in the verbal paradigm two assumptions are essential:

- (1) The acquisition of the verbal paradigm is based on overt evidence; only features relevant in target language are learned (Pinker 1984)

- (2) Acquisition hierarchy moves from unmarked general features to more marked specific features. (Neeleman & Weerman 1999, Harley & Ritter 2002)

I propose that the hierarchy of features for pronouns as proposed by Harley & Ritter (2002) is also applicable to the verbal paradigm. In Harley & Ritter's feature hierarchy, third person singular is the most unmarked featureless pronoun. Third person is less marked than first and second person because - unlike first and second person - third person does not necessarily refer to a discourse participant. Within the group of discourse participants the feature [addressee] is more marked than the feature [speaker], because second person is acquired after first person (Harley & Ritter 2002: 500). For the same reason plural forms are more marked than singular forms. Data concerning acquisition order as well as overgeneralization patterns in first language acquisition of verbal inflection observe the same markedness patterns as suggested for the pronouns by Harley & Ritter. In the acquisition of the verbal paradigms, the singular is also acquired before the plural and overgeneralization patterns are unidirectional: the singular is overgeneralized to the plural, the reverse pattern hardly ever occurs (Leonard, Caselli & Devescovi 2002, Van Kampen & Wijnen 2003: 251). An observation that confirms the relatively marked status of the feature [addressee] is that second person is acquired late relative to first and third person (Schlichting 1996: 124-145, Katičić 2003:260). Lastly research by Leonard, Caselli & Devescovi (2002) shows that Italian children are more likely to overgeneralize third person to first person than vice versa. In the singular, 11% of first person references are marked by third person inflection. First person is never used to refer to the third person. In the plural, 29% of first person references are marked by third person inflection. First person marking is used in 1% of the total number of third person plural references.

In one aspect I depart from the analysis by Harley & Ritter (2002). I assume syncretisms between first and second person and between second and third person are both allowed by the system. Harley & Ritter propose that only the former syncretism is possible because of three reasons. The first motivation is theory-internal: their system does not allow for flexibility in the formation of syncretisms. This motivation is irrelevant here. Secondly, syncretisms between first and second person occur more frequently than between second and third person. This

difference in frequency does not imply lack of naturalness of a syncretism between second and third person marking; it only implies an ordering of assumptions for a language learner. A syncretism between second and third person is assumed only if the target language does not show evidence for the assumption that first and second person form a syncretism. Lastly Harley & Ritter (2002) refer to Forchheimer (1953) who lists a set of behaviors that set first and second person apart from third person. Third person is more subject to objective subdivisions such as class, gender and location than first and second person. Moreover closely related languages often have cognate first and second person pronouns but third person pronouns that are not obviously related. I interpret Forchheimer's observations as support for the idea that third person is the most open category in the system. Because of this openness many candidates are allowed into this residual category. If the feature [addressee] is not part of the restricted category participant, the residual category third person is an obvious alternative.

On the basis of the above-mentioned markedness relations I propose four paradigm types presented in order of complexity in table 1 and 2. The paradigms in table 1 are concrete examples that correspond to more abstract feature paradigms presented in table 5.

Type I	Type II	Type III	Type IV
1. heareth	1. heare	1. heare	1. heare
2. heareth	2. heare	2. heareth	2. hearest
3. heareth	3. heareth	3. heareth	3. heareth

Table 1: Paradigm types in order of complexity

Type I	Type II	Type III	Type IV
[finite] → -eth	[finite] → -eth	[finite] → -eth	[finite] → -eth
	[>part.<] → -e	[<part.>] → -e	[<part.>] → -e
			[add.] → -est

Table 2: Feature paradigm types in order of complexity

Paradigm type I consists of only one feature: the general feature [finite] in this hypothetical case linked to the third person inflection marker *-eth*. All three persons are marked by the inflectional suffix *-eth*. If a language learner hears evidence for more features in her input, she will assume a second feature participant [part.] in her paradigm, in this case linked to the inflectional marker *-e*. This results in a paradigm where both discourse participants namely [speaker] and [addressee] are marked with the inflection marker *-e*. The broad interpretation of the feature [participant], i.e., reference to the feature speaker as well as addressee, is marked with open angle brackets (> <). If the type II paradigm is not in accordance with the input, the language learner reinterprets the feature [participant] as referring only to the default value, namely [speaker]. The more limited interpretation of the feature [participant] is marked by the use of closed angle brackets (< >). In paradigm type III only first person is referred to by the inflectional marker *-e*, both second and third person are referred to by the least specific inflectional marker, namely *-eth*. If this paradigm is still not in accordance with the input of the language learner, she will assume a third feature, namely [addressee], as presented under paradigm type IV.

The hypothesis is that T is only lost if replacing second person singular inflection (2S) by second person plural inflection (2P) results in a more economic singular verbal paradigm. I will test this prediction by looking at the effect of T-loss on the singular paradigm in some European languages. The Spanish paradigm in table 3 represents languages with rich inflection (six inflectional suffixes marking first, second and third person and number). In all rich-inflection languages the singular paradigm is a type IV paradigm and if we replace 2S-inflection by 2P-inflection, the paradigm remains a type IV paradigm. We predict no T-loss in rich inflection languages and this prediction is borne out.

	Singular	Plural
1	-o	-amos
2	-as	-aís
3	-at	- an

Table 3: Middle Spanish (based on Penny 1991:151)

With 2S: Type IV [finite] → -at [<participant>] → o [addressee] → -as	With 2P : Type IV [finite] → - at [<participant>] →- o [addressee] → ais
---	--

Table 4: Feature paradigm singular verbal inflection Middle Spanish

In table 5 and 6 we see the verbal paradigm of Middle Scandinavian based on Haugen (1976). The singular paradigm in Middle Scandinavian is a type III paradigm. Replacing 2S-inflection by 2P-inflection would create a more complex singular verbal paradigm, namely a type IV paradigm. As expected, T is not lost in the Scandinavian languages.

	Singular	Plural
1	-a	-um
2	-aR	- iþ
3	-aR	-a

Table 5: Middle Scandinavian (based on Haugen 1976:302)

With 2S: Type III [finite] → -aR [<participant>] → - a	With 2P : Type IV [finite] → - aR [<participant>] →- a [addressee] → - iþ
---	---

Table 6: Feature paradigm singular verbal inflection Middle Scandinavian

Tables 7 and 8 represent the Middle Dutch verbal paradigm. Replacing 2S-inflection by 2P-inflection in Middle Dutch changes a type IV paradigm into a less complex type III paradigm, so T-loss is expected.

	Singular	Plural
1	-(e)	-en
2	-s	-t
3	-t	-en

Table 7: Middle Dutch (based on Van Gestel et al. 1992)

With 2S: Type IV [finite] → -t [<participant>] → -(e) [addressee] → -s	With 2P : Type III [finite] → -t [<participant>] → -(e)
--	---

Table 8: Feature paradigm singular verbal inflection Middle Dutch

Table 9 and 10 show the paradigm of Southern English from around 1500. The paradigm with 2S-inflection is a type IV paradigm. Replacing 2S-inflection by 2P-inflection leads to a type II paradigm. Economy in the verbal paradigm is thus a possible extra motivation for T-loss in Standard English as well.

	Singular	Plural
1	-Ø	- Ø
2	-st	- Ø
3	-th	- Ø

Table 9: Southern English since 1500 (based on Lass 1999)

With 2S: Type IV [finite] → -t [<participant>] → e [addressee] → -s	With 2P : Type II [finite] → -t [>participant<] → -e
---	---

Table 10: Feature paradigm singular verbal inflection Southern English since 1500 (based on Lass 1999)

Table 11 and 12 show the verbal inflection of English in Northern England. English in Northern England had a type III paradigm. Replacing 2S-inflection by 2P-inflection does not lead to more economy in the verbal paradigm. As we can see in (9), the paradigm remains a type III paradigm. The prediction is that T is not lost in Northern England. The dialect atlas by Upton & Widdowson (1996) shows that this prediction is borne out: English in Northern England still shows the original T-form.

	Singular	Plural
1	-(e)	-es
2	-es	-es
3	-es	-es

Table 11: English of Northern England since 1300 (based on Lass 1992)

With 2S: Type III	With 2P : Type III
[finite] → -es	[finite] → -es
[<participant>] → -(e)	[<participant>] → -(e)

Table 12: Feature paradigm singular verbal inflection English of Northern England

4 Independent evidence for the role of inflection

In section 3 we have seen that – unlike most other European languages – Dutch and English V combined with a more economic verbal suffix than T. The hypothesis is that verbal inflection combining with the T-pronoun was under pressure. Usually this pressure leads to a change in the verbal paradigm, without pronoun loss. In the case of Dutch and English second person singular, there was an alternative to this more standard form of deflection. Instead of using the T-pronoun in combination with a simpler form of the verb, language users could also resort to the use of a V-pronoun that already combined with a simpler form of the verb. This strategy enabled language users to avoid second person singular inflection without neglecting positive evidence for the relationship between the T-pronoun and T-inflection in the input. This section provides independent evidence for the role of inflection in total T-loss in Dutch.

If avoiding uneconomic inflection is a factor in the loss of a pronoun, how would we be able to tell? A first prediction is that the old pronoun is lost in subject position first, because agreement is only relevant in subjects. For Dutch I have tested this prediction comparing thirteenth and sixteenth century texts. The untagged texts were brought together by Jacqueline Evers-Vermeul and Ninke Stukker and are taken from the *CD-rom*

Middelnederlands, the CD-rom *Klassieke Nederlandse literatuur*, the Laurens Jansz. Coster-Project and the *Dutch revolt* website. The corpus includes rhyme and prose texts in a variety of genres such as judicial texts, literary texts, recipes, saint lives, romances, songs and moralistic books. Table 13 shows the size of the total corpus.

	# texts	# words
13 th century rhyme	39	587,489
13 th century prose	4	311,090
16 th century rhyme	17	245,824
16 th century prose	33	476,264

Table 13: Corpus size

In the thirteenth century the T-pronoun is still very much in use. In this period we expect a similar distribution of T and V pronouns in the subjects and non-subjects. In my analysis any pronoun that does not combine with a finite verb is counted in the category non-subjects. Table 14 shows that in the thirteenth century texts we find a total of 10,897 address forms including 4742 subjects and 6155 non-subjects. The proportion of T-forms in the subjects is relatively large. Taking the distribution of subject and non-subjects in the total number of address forms as a departure point, we would expect to see 1008 subject forms in the T pronouns; in fact we observe 1162 subject forms. The chi-square probability (adjusted with Yates' correction) of getting the observed results if subject and non-subject had been evenly distributed among T and V pronouns is 0.00. The large amount of T-subjects confirms the prediction that subject T is still widely used in the thirteenth century.

	subject	non subject	total
du (T)	1162 (1008)	1154 (1308)	2316
gi (V)	3580 (3734)	5001(4847)	8581
	4742	6155	10897

Table 14: Thirteenth Century Address forms in Dutch

In the sixteenth century (table 15), we see a total number of 16,617 address forms. These address forms include 7371 subjects and 9246 non-subjects. The total number of T-pronouns is 571. The overall percentage of T-forms has decreased from 21% in the thirteenth century to 3% in the sixteenth century. Within this general trend of a decrease in the T pronouns, we see that the number of T-subjects relative to the number of non-subjects is very small. The observed amount of T-subjects is 128. Taking the subject/non-subject distribution of the total number of address forms as a point of departure, we would expect to see 253 subjects if subjects and non-subjects had been evenly distributed between T and V forms. The chi-square probability (adjusted with Yates' correction) of getting the observed results if subject and non-subject had been evenly distributed between T and V pronouns is 0.00. The relatively small number of T-subjects is as predicted. In the sixteenth century the T-pronoun is under pressure. A factor in this pressure is avoiding second person singular inflection and avoiding inflection is only relevant for subject-T.

	subject	non subject	total
du (T)	128 (253)	443 (318)	571
gi (V)	7243 (7118)	8803 (8928)	16046
	7371	9246	16617

Table 15: Sixteenth century address forms in Dutch

The prediction that subject T disappeared before non-subjects in Dutch is also confirmed by sixteenth century grammars. These grammars list the T pronoun as a direct object, an indirect object and a possessive, but the T is excluded in the function of subject because of the problematic verb form that follows the subject (Van der Sijs 2004: 469).

A second prediction that follows from the hypothesis that pressure on second person singular inflection motivates total T-loss is that the traditional pronouns *du/thou* will co-occur with irregular high frequency verbs longer than with low frequency verbs because the need to avoid inflection under pressure arises later in combination with a high frequency verb than in combination with a low frequency verb. Inflection on high

frequency verbs is acquired relatively early and high frequency verbs are therefore more resistant to deflection than low frequency verbs. Coveney (2000) refers to this restriction on the co-occurrence of pronouns and relatively infrequent verbs as 'lexical diffusion in reverse'. Confirmation of the prediction that T combines with irregular high frequency verbs longer than with low frequency verbs comes from Kaajan (1914). In his dissertation on the translation of the *Statenbijbel* (Dutch authorized version of the Bible, a translation advisory board for this Bible gathered in 1618) Kaajan reports that Bible translators opposed to the use of *du* as an address form of God partly because the common people only used *du* with a very limited set of verbs, namely *hebben* 'have', *zijn* 'be' and *zullen* 'shall'. The Dutch corpus data show that these three verbs do indeed combine with T most frequently. In the thirteenth century these three verbs together already make up 51% of the total number of finite verbs combining with T. In the sixteenth century T almost exclusively combines with irregular verbs. Out of the eight texts that use subject T, only three texts show T in combination with a regular verb. What these three texts share is that they are reprints from fifteenth century texts.

5 Conclusion

In this paper I have argued that socio-pragmatic explanations can motivate why T was lost in many domains of speech but that they cannot motivate total T-loss. An additional factor in T-loss is the verbal inflection. Dutch and English belong to the small set of languages where replacing the T-pronoun by a V-pronoun resulted in a more economic singular verbal paradigm. The fact that the pronoun V that combined with economic inflection could refer to the second person singular as well as T enabled language users to avoid problematic inflection. We have seen independent evidence for the role of inflection in the V-pronouns in Dutch. The T-pronoun is lost most quickly in subject position, the only position where agreement is relevant. Moreover if T is used as a subject, it mostly combines with common irregular verbs and these verbs are most resistant to deflection and are thus least likely to trigger avoidance strategies for T-inflection.

References

- Berteloot, Amand. 2001. Anredepronomen in Brabant um 1400. In *Vulpis Adolatio. Festschrift für Hubertus Menke zum 60. Geburtstag.*, ed. Robert Peters, Horst Pütz & Ulrich Weber, 41-56. Heidelberg: C. Winter.
- Betsch, Michael. 2003. The system of Czech bound address forms until 1700. In ed. Andreas H. Jucker & Irma Taavitsainen, 125-146.
- Brown, Roger & Albert Gilman. 1960. The pronouns of power and solidarity. In *Style in language*, ed. Thomas Sebeok, 252-275. Cambridge, Mass: MIT-press.
- Busse, Ulrich (2002). *Linguistic Variation in the Shakespeare Corpus. Morpho-syntactic variability of second person pronouns.* Amsterdam, Philadelphia: John Benjamins (= Pragmatics and Beyond New series 106).
- CD-rom *Klassieke Nederlandse literatuur. Van de Middeleeuwen tot en met de Tachtigers.* 1999. Utrecht: Spectrum Software.
- Cd-rom *Middelnederlands.* 1998. The Hague/Antwerp: Sdu.
- Coveney, Aidan. 2000. Vestiges of *nous* and the 1st person plural verb in informal spoken French. *Language Sciences* 22: 447-481.
- Dutchrevolt, <http://dutchrevolt.leidenuniv.nl>
- Forchheimer, Paul. 1953. *The category of person in language.* Berlin: Walter de Gruyter.
- Harley Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: a feature geometric analysis. *Language* 78, 482-526.
- Haugen, Einar. 1976. *The Scandinavian languages. An introduction to their history.* Londen: Faber & Faber.
- Hunt, Tony .2003. The use of tu/vus in the *Anglo-Norman Seinte Resurrecion.* In ed. Irma Taavitsainen & Andreas Jucker, 47-59.
- Kaajan, Hendrik (1914). *De pro-acta der Dordtsche Synode in 1618.* Rotterdam: T. De Vries.
- Kampen, Jacqueline van & Frank Wijnen (2003). Grammaticale ontwikkeling. In *Kindertaalvererving. Een handboek voor het Nederlands*, ed. Steven Gillis & Annemarie Schaeerlaekens, 225,285. Groningen: Martinus Nijhoff.
- Katičić, Antigone (2003). Early verb development in one Croatian-speaking child. In *Development of Verb Inflection in First Language Acquisition*, ed. D. Bittner, W. Dressler & M. Kilani-Schoch , 239-267. The Hague: Mouton de Gruyter.
- Neeleman, Ad & Fred Weerman. 1999. *Flexible Syntax; A Theory of Case and Arguments.* Dordrecht: Kluwer Academics.
- Laurens Jansz. Coster-Project, <<http://cf.hum.uva.nl/dsp/ljc/>>
- Lass, Roger 1992. Phonology and morphology. In *Cambridge History*

- of the English language. Vol II. 1066-1476*, ed. Norman Blake, 23-154. Cambridge: Cambridge University Press.
- Lass Roger .1999. Phonology and Morphology. In *The Cambridge History of the English language. Vol III 1476-1776*, ed. Roger Lass, 56-106. Cambridge: Cambridge University Press.
- Leonard, Laurence B. , M. Christina Caselli & Antonella Devescovi. 2002. Italian children's use of verb and noun morphology during the preschool years. *First Language* 22, 287-304.
- Mühlhauser Peter & Rom Harré. 1990. *Pronouns & people. The linguistic construction of social and personal identity*. Oxford: Basil Blackwell LTD.
- Penny, Ralph. 1991. *A history of the Spanish language*. Cambridge: Cambridge University Press.
- Pinker, Steven 1984. *Language Learnability and Language Development*. Cambridge, Mass: Harvard University Press.
- Schlichting, Liesbeth 1996. *Discovering Syntax. An empirical study in Dutch Language Acquisition*. Nijmegen: Nijmegen University Press.
- Simon, Horst J. 2003. From pragmatics to grammar. Tracing the development of respect in the history of German pronouns of address. In ed. Irma Taavitsainen & Andreas H. Jucker , 85-123.
- Taavitsainen, Irma & Andreas H. Jucker 2003. *Diachronic perspectives on address term systems*. Amsterdam, Philadelphia: John Benjamins (= Pragmatics and Beyond. New series 107).
- Upton Clive & J.D.A. Widdowson 1996. *An Atlas of English Dialects*. Oxford: Oxford University Press.
- Van der Sijs, Nicoline 2004. *Taal als mensenwerk. Het ontstaan van het ABN*. Utrecht: Sdu uitgevers.
- Van Gestel, Frank, Jan Nijen Twilhaar, Tineke Rinkel en Fred Weerman.1992. *Oude zinnen; grammaticale analyse van het Nederlands tussen 1200-1700*. Leiden: Martinus Nijhoff.
- Wales, Katie 1983. Thou and you in early modern English: Brown and Gilman re-appraised. *Studia Linguistica* 37, 107-125.
- Wales, Katie 1996. *Personal pronouns in present-day English*. Cambridge: Cambridge University Press.

Faculteit der Geesteswetenschappen, NTK
 Universiteit van Amsterdam
 Spuistraat 134
 1012 VB Amsterdam
 s.p.aalberse@uva.nl