

The effect of prescriptivism on comparative markers in spoken Dutch*

Ferdy Hubers and Helen de Hoop
Radboud University Nijmegen

Dutch prescriptive grammar rules dictate that the complementizer *dan* ‘than’ should be used in comparative constructions of inequality. This has been an issue for grammarians from the sixteenth century onwards when *als* ‘as’ started to be used as an alternative form in this type of context. In order to find out why and when people choose one comparative marker over the other, we examined the use of these markers in the Spoken Dutch Corpus (CGN). We found that the use of *dan* is overall more common than *als* in comparative constructions of inequality, even though from a linguistic point of view *als* might be favoured. The choice between *als* and *dan* turns out to be strongly correlated with the level of education. Although this factor has been assumed to be of influence for a long time, as far as we know it has never been quantitatively tested before. We conclude that the effect of the level of education we found reflects the strong influence of the prescriptive rule taught in schools, repressing the use of *als* in comparatives of inequality.

Keywords: Dutch, prescriptivism, comparatives, Spoken Dutch Corpus, education

1. Introduction

Native speakers of Dutch are taught in school that they should use the comparative marker *dan* ‘than’ instead of *als* ‘as’ in a comparative of inequality, similar to English *John is taller than Bill* versus *John is as tall as Bill*. This prescriptive rule has to be taught explicitly because there is a tendency in Dutch to use the comparative marker *als* in both types of comparatives, which would lead people to use the ‘incorrect’ construction in (1):

- (1) Jan is groter als Willem
Jan is taller as Willem
“Jan is taller than Willem”

- (2) Jan is even groot als Willem
 Jan is equally tall as Willem
 “Jan is as tall as Willem”

The use of *dan* is still dominant in present-day Dutch, which might be the result of the prescriptive rule taught in schools, but there is clearly a tendency to use *als* in comparatives such as (1). This raises the question where the ‘incorrect’ use *als* comes from. Can we identify language-internal or language-external factors that favour the use of *als* and if so, what is the role of prescriptive grammar in repressing this use? In order to answer these questions, Section 2 summarizes the state of the art in prescriptive grammar and the argument that has been put forward in favour of *dan*. Section 3 investigates the contexts of use of *als* and *dan* and concludes on the basis of the functions that *als* and *dan* have in Dutch, that *als* would be a more natural choice in comparative constructions of inequality than *than*. In Section 4 we present our corpus study and the statistical analysis that shows the influence of the prescriptive grammar rule.

2. The choice between *als* and *dan* according to prescriptive grammar

In the second half of the sixteenth century the rise of the comparative marker *als* is seen in Dutch, but already in the first half of the seventeenth century a countermovement emerged. This was led by Balthazar Huydecoper, who strongly opposed the use of *als* as a comparative marker (Stroop 2011). Stroop (2011) reports on a recent Google search he conducted from which it became clear that the countermovement has been very successful, at least in written language on the internet: *Dan* was used far more often than *als* in comparatives of inequality. It is well known that prescriptive grammars and language advisory councils strongly recommend the use of *dan* in this context. The online language advisory council of the *Nederlandse Taalunie* ‘Dutch language union’ states that only *dan* is the correct form here, but also notes that *als* is often used, especially in spoken Dutch (<http://taaladvies.net/taal/advies/vraag/354>). The online *Taaladviesdienst* ‘Language advisory council’ of the popular magazine *Onze Taal* ‘Our language’ formulates it as follows, after acknowledging that linguists have claimed that there is nothing wrong with the use of *als*:

“Maar in de praktijk kan toch het best voor *groter dan* gekozen worden, omdat *groter als* nog altijd veel weerstand oproept — en dat zal nog wel een tijdje zo blijven.”

“In practice it is better to use *groter dan* ‘bigger than’, because *groter als* ‘bigger as’ still generates a lot of resistance — and this will be the case for some time.”

<http://www.onzetaal.nl/taaladvies/advies/groter-als-groter-dan>

On the same site we find a reference to an article in the *Woordenboek der Nederlandsche Taal* (WNT) 'Dictionary of the Dutch Language' (1884) in which the influence of linguists is mentioned as a source for maintaining *dan* even when *als* might have been dominant in spoken language (especially in "*volkspraak*" 'common speech') since the 17th century. Although the WNT first argues that *als* and *dan* are both justified from a historical perspective, it continues as follows:

"Zal men daarom in de schrijftaal beide gelijkelijk erkennen, of wel, alleen op de spreektaal lettende, *als* voor *dan* in de plaats stellen? Het ware niet te wenschen. (...) *Als* moge op zich zelf ook onberispelijk zijn, maar het is niet gerechtigd om *dan* in de plaats, die het volkomen goed bekleedt, te verdringen. (...) En werkelijk is *dan* bruikbaar en juist *dan als*."

"Should written language acknowledge both variants equally, or should it, following spoken language, prefer *als* over *dan*? This is not desirable. (...) *Als* might, in and of itself, be impeccable, but it is not equipped to supersede *dan* from its justified place. (...) And *dan* is more useful and more correct than *als*."

<http://gtb.inl.nl/iWDB/search?wdb=WNT&actie=article&uitvoer=HTML&id=M003534>

The WNT further compares the development in Dutch to the development in German, in which *denn* has been replaced by *als*, and ends with a passionate appeal not to let that happen in Dutch and to save *dan* from being replaced with *als*.

From this the question follows as to why *dan* is to be preferred over *als* in comparatives of inequality. No satisfactory answer has been found. The 'argument' usually given is that *als* is disapproved of and therefore should be avoided, whereas *dan* is considered 'correct'. The only other argument prescriptive grammars give in favour of *dan* is that from a semantic point of view it is better to have two different forms in order to be able to distinguish between comparatives of equality and inequality. Evidently, the marker itself does not contribute to the meaning of equality or inequality. The interpretation of (1) above is not altered to equality by the use of the marker *als* instead of *dan*. Similarly, if *dan* were used in a comparative of equality such as (2) (which is sometimes encountered as an instantiation of hypercorrection),¹ this would not trigger an interpretation of inequality. A distinction between comparative markers of equality and inequality is therefore not needed from a semantic point of view. Moreover, constructions such as in (3) involve a comparison of inequality semantically, but are syntactically similar to comparatives of equality, and therefore require the comparative marker *als*.

- (3) Jan is twee keer zo groot als Willem
 Jan is two times so tall as Willem
 "Jan is twice as tall as Willem"

Hence, the main argument to use *dan* in a comparative of inequality instead of *als* is that *dan* is the prescribed form. Since only *dan* is considered correct and since this rule is also explicitly taught at schools, the question arises why *als* still exists as an alternative form in comparatives of inequality. In Section 3 we will review the contexts in which *als* and *dan* are used in Dutch, and argue that this provides us with an explanation for the tendency to use *als* instead of *dan* in comparatives.

3. Functions of *als* and *dan* in context

Several functions of *als* and *dan* can be distinguished in Spoken Dutch, in accordance with the different word classes they belong to. Below are the types of contexts in which *als* is used in Dutch (all taken from the spontaneous speech components a and b of the Spoken Dutch Corpus (CGN)):

- (4) Maar *als* je 't vak echt op de rails wil zetten
but if you the subject really on the rails want put
dan moet je d'r ook wat van weten
then must you there too something of know
“But to properly set up the curriculum, you really need to know something about it”
[CGN fn000109.69]
- (5) Want ik stel me voor je bent er *als*
because I imagine me for you are there as
boer uit of niet uit
farmer out or not out
“Because I can imagine, as a farmer you either made up your mind or not”
[CGN fn000096.288]
- (6) Nou dat is vast net zo erg *als* The Bold and the Beautiful
well that is likely just so bad as The Bold and the Beautiful
“Well that is probably just as bad as The Bold and the Beautiful”
[CGN fn000005.319]
- (7) Nou, iemand koopt sneller een slaapzak *als* een tent
well somebody buys faster a sleeping.bag as a tent
“Well, someone will more easily buy a sleeping bag than a tent”
[CGN fn000897.240]
- (8) Maar zijn wel een beetje zachter *als* dat ze moeten zijn
but are PART a bit softer as that they must be
“But they are a little softer than they need to be”
[CGN fn000293.121]

The utterances in (4–8) illustrate the type of contexts in which *als* is used. We see that *als* can have either of two functions: It is either a complementizer as in (4) and (8) or a preposition as in (5)–(7). Examples (6)–(8) are comparatives, of which (6) is a comparative of equality, and (7) and (8) are comparatives of inequality. In a comparative of inequality *als* can be a complementizer, as in (8), or a preposition, as in (7). Napoli (1983), Hoeksema (1983), and Hendriks (1995) convincingly argue that comparative markers such as *than* in English are either complementizers, when their complement is a clause as in (8), or prepositions, when their complement is a phrase as in (7). Note that phrasal comparatives cannot always be analysed as reduced clauses. Crucially, a reduced clausal analysis is impossible for the complements in (9) and (10) below, which show that *than* and *dan* function as prepositions here (Napoli 1983; Hendriks 1995). Similarly, (11) is an example of the Dutch preposition *als* whose complement could not be analysed as a reduced clause underlyingly.

- (9) Jane is taller than six feet
- (10) Niemand schaatst harder dan zichzelf
nobody skates harder than himself
“Nobody skates faster than himself”
- (11) Hij is verkleed als mij
he is dressed.up as me
“He is dressed up like me”

Across languages prepositions and complementizers are often morphologically and semantically related and differ only in the type of complement they take, a phrase or a clause. This was illustrated for *als* above. Below we give two more examples of preposition-complementizer pairs in Dutch, *voor* ‘before’ and *tot* ‘until’.

- (12) a. Er is een parkeerplaats **voor** het huis
there is a parking.place before the house
“There is a parking space in front of the house”
- b. Hij woonde in Utrecht **voor** de oorlog
he lived in Utrecht before the war
“He lived in Utrecht before the war”
- c. Hij woonde in Utrecht **voor** hij ziek werd
he lived in Utrecht before he ill got
“He lived in Utrecht before he got ill”
- (13) a. Hij fietste **tot** de brug
he cycled until the bridge
“He cycled until the bridge”

- b. Hij werkte door **tot** zijn pensioen
he worked on until his retirement
“He worked till his retirement”
- c. Hij werkt door **tot** hij erbij neervalt
he works on until he thereby down.falls
“He works until he drops dead”

The prepositions *voor* ‘before’ and *tot* ‘until’ get a spatial reading in (12a) and (13a) and a temporal reading in (12b) and (13b). The complementizers *voor* ‘before’ in (12c) and *tot* ‘until’ in (13c) get a temporal interpretation. Other examples of complementizers in Dutch that are also prepositions are *sinds* ‘since’, *om* ‘round’ and *na* ‘after’. Other complementizers are compounds of a preposition and a complementizer, e.g., *voordat* ‘before.that’, *totdat* ‘until.that’, *nadat* ‘after.that’, *doordat* ‘through.that’, *alsof* ‘as.if’. Note that in English, prescriptive grammarians only allow the use of *like* as a preposition (e.g. *He is like my little brother*) and not its use as a complementizer (e.g. *He sounds like he is from Boston*), although the latter use is very common as well. The prescriptive rule, as discussed by van Gelderen (2002: 131), states: “[*l*]ike is a preposition and not a complementizer. This means that it can introduce an NP but not a clause. Instead of *like*, *as* is used to introduce a sentence.”

It is not surprising that *als* ‘as’ in Dutch can also be a preposition or a complementizer in comparatives in Dutch, dependent on whether its complement is a (noun) phrase or a clause, respectively. Importantly, *als* is not just a preposition or a complementizer in comparatives, but also in the other contexts in which it occurs, as illustrated in (4) and (5) above. At this point, let us consider the various contexts in which *dan*, the other comparative marker in Dutch, can occur (van Bergen 2010). Some examples, again extracted from the Spoken Dutch Corpus, are given below.

- (14) Maar als je ’t vak echt op de rails wil
but if you the subject really on the rails want
zetten **dan** moet je d’r ook wat van weten
put then must you there too something of know
“But to properly set up the curriculum, you really need to know something
about it” [CGN fn000109.69]
- (15) Ik kan **dan** bij m’n ouders slapen
I can then at my parents sleep
“Then I can sleep at my parents” [CGN fn000838.256]
- (16) **Dan** heeft ie twee banken een tafeltje en **dan**
then has he two sofas a table.DIM and then

nog een eettafel
yet a eat.table

“He then has two sofas, a small table and then also a dinner table”

[CGN fn000626.193]

- (17) Hoe ver **dan**?

how far then

“How far is it then?”

[CGN fn000983.123]

- (18) En tegen de tijd dat hij veertig is verdient

and against the time that he forty is earns

ie veel meer **dan** deze vent

he much more than this guy

“And by the time he is forty, he earns more more than that guy”

[CGN fn000509.80]

- (19) Ik durf te wedden dat jouw ouders makkelijker

I dare to bet that your parents easier

boeken lezen **dan** dat jij dat doet

books read than that you that do

“I bet that your parents read books more easily than you do”

[CGN fn009180.13]

We can conclude from the examples above that *dan* belongs to more word classes than *als*. In (14–17) *dan* is used as a temporal adverb or a modal particle, translated as *then* in English. Only in comparative constructions such as (18) and (19), *dan* is used as a preposition or a complementizer, comparable to *als* in Dutch, and translated as *than* in English. In other words, whereas *als* is a preposition or a complementizer in all contexts of use, including comparatives, *dan* is a preposition or a complementizer *only* in comparatives. The most frequent function of *dan* is its function as a temporal or modal adverb or particle, and its use as a preposition or complementizer is restricted to comparatives of inequality. This became clear after searching half of the Spoken Dutch Corpus (CGN). In about 40.000 cases *dan* was used as an adverb or particle, whereas only in about 2000 cases it was used in comparative contexts.² By contrast, *als* is always a preposition or a complementizer, and therefore its use in comparatives seems more natural than the use of *dan*.

Postma (2006) offers a different explanation for the tendency to use *als* instead of *dan* in comparatives. He argues that *dan* as a comparative marker in Middle Dutch originated from merging complementizer *dat* with a negative clitic *en*. This negative element is responsible for the monotone decreasing behaviour of comparatives, notoriously contexts which license the occurrence of negative polarity items (Hoeksema 1983). For example, in English the negative polarity item *anyone*

can occur in a comparative of inequality: *John is taller than anyone else* or *John is taller than anyone expected*. Postma (2006) argues, however, that *dan* is no longer marked for its monotone decreasing properties in contemporary Dutch, which is why it has been replaced by other complementizers in certain dialects of Dutch. These other complementizers, to wit *als*, *of*, and *wie*, are marked for being monotone decreasing in comparatives, according to Postma (2006), because they are also used in other monotone decreasing contexts. By contrast, *dan* never leads to a monotone decreasing context except for comparatives, and this is why Postma argues that *dan* is being replaced by other complementizers. Postma's hypothesis about *dan* originating from merging the complementizer *dat* and the negative clitic *en* is very attractive and plausible, but in our view it does not explain why *dan* is being replaced by other complementizers. The fact that *dan* is not monotone decreasing outside the domain of comparatives of inequality does not support the claim that it has lost its monotone decreasing value, simply because *dan* is not used as a complementizer at all outside the domain of comparatives. Therefore, we can only compare *dan* and *als* in comparatives and in that context they are equally monotone decreasing.

The predictability of *als* as a comparative marker is thus higher than *dan* which could explain the tendency to use *als* instead of *dan* in a comparative of inequality. If there were only one comparative marker, *als*, for both types of comparatives, there would be no uncertainty about which marker to select for comparatives (De Lange, Vasic & Avrutin 2009). In technical terms, when *als* would have the probability of 1 and *dan* the probability of 0 as a complementizer or a preposition, the absolute entropy would reach the minimum level, which would facilitate processing (production and interpretation) of a comparative. By contrast, if the choice between *als* and *dan* is not so clear, and if the probability distribution between the two elements were more homogeneous, it becomes more difficult to select the right comparative marker, and the uncertainty level increases (De Lange et al. 2009). Therefore, from a linguistic processing point of view, we would expect that, all else being equal, *als* would be a better choice for a comparative marker in Dutch than *dan*.

4. The comparative markers *als* and *dan* in the Spoken Dutch Corpus (CGN)

This section presents our corpus study of the comparative markers *als* and *dan* in the Spoken Dutch Corpus (CGN). We chose the CGN because it is big (nearly 9 million words), it contains spoken Dutch only (clearly, violations of prescriptive rules barely occur in written language), and it is easy to use.

4.1 Method

We searched for *als* and *dan* as comparative markers, making use of the following tags: The markers *als* and *dan* were tagged as complementizers (VG2) (also erroneously when used as prepositions) or, only in the case of *dan*, erroneously tagged as an adverb (BW); the comparatives were tagged as adjectives ending in *-er* (ADJ10) such as in *groter als* ‘taller than’, or ending in *-s* (ADJ7) such as in *anders als* ‘different than’, or they were tagged as pronouns (VNW24 and VNW26), such as in *meer als* ‘more than’ or *minder als* ‘less than’. We limited the distance between the comparative and the comparative marker to maximally three words, as illustrated in example (20) below, as the search would otherwise produce many false hits in which *als* and *dan* do not function as comparative markers.

- (20) Vroeger kregen de kinderen veel **meer** huiswerk naar
 formerly got the children much more homework to
 huis **dan** tegenwoordig
 house than presently
 “In the old days children got more homework than nowadays”
 [CGN fn009223.49]

Thus, we collected 4565 utterances from the Spoken Dutch Corpus. We removed all data from the less spontaneous data, in which as expected the use of *als* as a comparative marker of inequality was rare. We only kept data from the spontaneous speech components, i.e. 2929 utterances. We randomly selected half of this set, while retaining the relative proportions of *als* and *dan*, their distribution over the various components and their distribution over two language groups, Dutch and Flemish. We reduced the amount of data to a set of 1465 occurrences. This set of data was annotated for the linguistic variables: Type of comparative (phrasal or clausal) on the basis of complement type of *als* or *dan*; the distance between the comparative and the comparative marker (maximally three words) and the comparative combination that indicates whether the comparative is used in combination with a noun, as in *meer huiswerk* ‘more homework’, an adjective, as in *minder dronken* ‘less drunk’, or independently, as in *beter* ‘better’ or *anders* ‘different’. The data were annotated independently by two annotators. The interannotator-agreement concerning the type of comparative was almost perfect (Cohen’s Kappa = 0.907). The linguistic variable *distance* also turned out to have an almost perfect interannotator-agreement (Cohen’s Kappa = 0.952), as well as the variable *comparative combination* (Cohen’s Kappa = 0.897). Sociolinguistic information about the speakers of the utterances was taken from the CGN. We used the sociolinguistic variables *gender*, *age*, *birth region*, *region of residence*, and *education* for our analysis.

4.2 Results and analysis

From crosstabs examination it appears that only two of the variables introduced above, namely *education* and *region of residence*, are likely to play a part in predicting the choice between *als* and *dan*. All other variables, the linguistic as well as the non-linguistic ones, turned out not to relate to the choice between *als* and *dan*.

The variable *region of residence* is coded in six categories, five regions — north Netherlands (Friesland, Groningen, Drenthe, Overijssel and Noord-Holland), middle Netherlands (Zuid-Holland, Utrecht, Gelderland), south Netherlands (Limburg, Noord-Brabant, Zeeland), north Belgium (Flanders), south Belgium (Wallonia) — and one category unknown. The latter category is ignored in the analysis. Table 1 gives the crosstabulation in which *comparative marker* and *region of residence* are compared.

Table 1. Crosstab Comparative marker — Region of residence

		Region				
		North-N	Middle-N	South-N	North-B	Total
Comparative Marker	<i>als</i>	26 (13.5%)	66 (13.8%)	51 (40.2%)	69 (12.8%)	212
	<i>dan</i>	167 (86.5%)	412 (86.2%)	76 (59.8%)	470 (87.2%)	1125
Total		193 (100%)	478 (100%)	127 (100%)	539 (100%)	1337

In Table 1 the regions of residence are compared with the use of *als* and *dan* in comparative constructions. In the northern region of the Netherlands people do not often use *als* in comparatives (n=26; 13.5%); they clearly prefer *dan* (n=167; 86.5%). For the middle region of the Netherlands and the northern region of Belgium (Flanders) this is similar. Only the southern region of the Netherlands, consisting of the provinces Limburg, Noord-Brabant and Zeeland, shows more variation. In this region *als* (n=51; 40.2%) is more frequent than in the other regions, although it is still less frequent than *dan* (n=76; 59.8%). To conclude, whereas in the other regions *als* occurs only in about 13% of the comparatives, in the southern region of the Netherlands it occurs about 40% of the time.

The variable *education* is coded in four categories, adopted from the CGN, to wit *high* (finished university for professional or academic education), *middle* (finished secondary education or senior secondary vocational education, *low* (finished primary education), and *unknown*. The unknown cases are not taken into account. Table 2 presents the crosstabulation that compares the type of comparative marker with the level of education.

Table 2 compares the dependent variable *comparative marker* and the independent variable *education*. From this crosstab it becomes clear that people with a high education in general use more *dan* (n=1006; 89.7%) than *als* (n=115; 10.3%).

Table 2. Crosstab Comparative marker — Education

		Education			
		High	Middle	Low	Total
Comparative Marker	<i>als</i>	115 (10.3%)	82 (36%)	23 (62.2%)	220
	<i>dan</i>	1006 (89.7%)	146 (64%)	14 (37.8%)	1166
Total		1121 (100%)	228 (100%)	37 (100%)	1386

Middle educated people use more *dan* (n=146; 64%) than *als* (n=82; 36%) too. The number of *als* used by middle educated people is higher than the number *als* used by high educated people. People with a low education, in contrast to the people from the other categories, use more *als* (n = 23; 62.2%) than *dan* (n = 14; 37.8%).

4.3 Discussion

The reason for setting up this corpus study was to find out what factors motivate the choice between the comparative markers *als* and *dan* in comparatives of inequality in Dutch. We have found that the overall use of *dan* greatly exceeds the use of *als* in these comparatives. One important factor is the region of residence. People from the southern part of the Netherlands, specifically the provinces Noord-Brabant, Limburg, and Zeeland, use *als* far more often than people from the rest of the Netherlands and Belgium; although they still use *dan* more often than *als*. This effect of region on the use of *als* rather than *dan* can be explained by the existence of dialects of this area in which *als* is the comparative marker in a comparative of inequality, such as in the Brabantian dialect of Dutch. In fact, almost all dialects of Dutch predominantly use *als* in comparative constructions of inequality, but the dialects spoken in Noord-Holland and Zuid-Holland are more closely related to Standard Dutch than the southern dialects (Barbiers, Bennis, De Vogelaer, Devos, van der Ham 2005: 1.3.1.4). Therefore, it might be the case that the influence of the dialects on Standard Dutch in these provinces is less strong than the influence of the southern dialects on Standard Dutch.

Another important factor is education. We found that the use of *dan* highly correlates with the level of education. Only low educated people use more *als* than *dan* and highly educated people almost exclusively use *dan* with comparatives of inequality. Apparently, the prescriptive rule that is taught at school overrules any other factor that may favour the use of *als*. Because highly educated people use *dan* far more often than lower educated people, and because this is independent from other factors such as the distance between the comparative and its marker, the type of comparative, gender, age, etc., we may conclude that the prescriptive rule is taught in the educational system very successfully, and that at least for highly educated people the use of *dan* in comparatives of inequality has become part of

their internal grammar. This is reminiscent of the approach of van Bergen, Stoop, Vogels & de Hoop (2011) who argue that for people who have as part of their grammar the prescriptive constraint that prohibits the use of *hun* 'them' as a subject, this constraint outranks all other language-internal constraints that would favour the use of *hun* 'them' as a subject.

5. Conclusion

The aim of this article was to seek an explanation for the tendency in Dutch to use *als* in comparatives of inequality instead of *dan*. We argued that on the basis of the functions of *als* and *dan* in context, *als* would be a better candidate than *dan* in comparatives, because *als* is always used as a complementizer or a preposition, also in comparatives, while *dan* is never used as a complementizer or a preposition, except for in comparatives. We conducted a corpus study in order to identify which factors can be used to predict the choice between *als* and *dan* in comparatives of inequality. It was found that region of residence is a predicting factor, in the sense that people from the south of the Netherlands use *als* more often than people from other regions of the Netherlands and Belgium. The most striking result was the correlation with the level of education. It turned out that highly educated people almost always use *dan* while low educated people use more *als* than *dan*. Middle educated people use more *dan* than *als*, but to a lesser degree than highly educated people. This suggests that the prescriptive rule taught at school that dictates the use of *dan* in comparative constructions of inequality has a major impact on the use of *dan* in Dutch. It is to be expected that without this strong normative rule imposed on language use, *als* might have replaced *dan* as a comparative marker already.

Notes

* We would like to thank the audience at the TIN-dag 2013, two anonymous reviewers, the editors of this volume, and our colleagues Ad Foolen, Lotte Hogeweg, Sander Lestrade, Kees de Schepper, Peter de Swart, Thijs Trompenaars, Ruti Vardi, and Puck Wildschut, for helpful comments and discussion. A special word of thanks goes to Vera van Mulken for translating the Dutch fragments to English.

1. Ad Foolen provided us with the following example of this type of hypercorrection, taken from the Dutch quality newspaper *NRC Handelsblad*, April 6, 2013: *Volgens schattingen lekt 4 procent van het opgepompte gas uit een veld weg naar de lucht en methaan is een 25 keer zo sterk broeikasgas dan CO2*. 'Approximately 4 per cent of the pumped gas escapes through the air and methane is 25 times a strong greenhouse gas as CO2.

2. These frequency counts do not reflect the absolute numbers, but still give a good indication of the proportions.

References

- Barbiers, Sjef, Hans Bennis, Gunther De Vogelaer, Magda Devos & Margreet van der Ham. 2005. *Syntactische atlas van de Nederlandse dialecten Deel 1*. Amsterdam: Amsterdam University Press.
- Bergen, Geertje van. 2010. Dutch *dan* in discourse. Presentation Semantics in the Netherlands. SiN-day 8, November 5, Radboud University Nijmegen.
- Bergen, Geertje van, Wessel Stoop, Jorrig Vogels & Helen de Hoop. 2011. Leve *hun*! Waarom *hun* nog steeds *hun* zeggen. *Nederlandse Taalkunde* 16. 2–29.
- De Lange, Joke, Nada Vasic & Sergey Avrutin. 2009. Reading between the (head)lines: A processing account for article omissions in newspaper headlines and child speech. *Lingua* 119. 1523–1540.
- Gelderen, Elly van. 2002. *Introduction to the grammar of English: Syntactic arguments and socio-historical background*. Amsterdam/Philadelphia: John Benjamins.
- Hendriks, Petra. 1995. Comparatives and categorial grammar. PhD dissertation, University of Groningen.
- Hoeksema, Jack. 1983. Negative polarity and the comparative. *Natural Language and Linguistic Theory* 1. 403–434s.
- Napoli, Donna Jo. 1983. Comparative ellipsis: A phrase structure analysis. *Linguistic Inquiry* 14. 675–694.
- Postma, Gertjan. 2006. Van *groter dan* naar *groter als* — structurele oorzaken voor het verval van het comparatieve voegwoord *dan*. *Nederlandse Taalkunde* 11. 2–22.
- Stroop, Jan. 2011. De lange arm van de grammatici. In Zdenka Hrnčířová, Ellen Krol, Kees Mercks, Jan Pekelder & Jesse Ultzen (eds.), *Praagse perspectieven 7. Handelingen van het colloquium van de sectie Nederlands van de Karelsuniversiteit te Praag, op 24 en 25 maart 2011*, 137–153. Prague: Prague University Press.

Author's addresses

Helen de Hoop
Dept. of Linguistics/ CLS
Radboud University Nijmegen
P.O. Box 9103
6500 HD Nijmegen, The Netherlands
h.dehoop@let.ru.nl

Ferdy Hubers
Dept. of Linguistics
Radboud University Nijmegen
P.O. Box 9103
6500 HD Nijmegen, The Netherlands
f.hubers@student.ru.nl