Hidden indicators
Modal Particles in modern Dutch fiction and fiction translated from English: a quantitative approach

Ton van der Wouden (Meertens Instituut), Saskia Lensink (Leiden University), Karina van Dalen-Oskam (Huygens Instituut)

Introduction

The bulk of the fiction books produced and read in the Netherlands is translated, most often from English. Our paper investigates the possibilities of using stylometric analyses of certain non-content words, more specifically, modal particles, to distinguish translated prose from books originally written in Dutch.

Just like other continental Germanic languages but unlike English, Dutch has and uses a considerable amount of small, unaccented, adverb-like elements known in the literature as 'modal particles' (van der Wouden 2002). Their contribution to the meaning of the utterance in which they are used, is often more pragmatic than semantic, compare (1) and (2) below:

(1) *Ga zitten* 'sit down'
(2) *Ga **maar** zitten* 'please sit down'

The bare imperative form in (1) is too impolite in many situations. More normal is (2), where the particle *maar* (related to *maar* 'but') makes the utterance more acceptable.

There is a growing body of qualitative evidence that prose translated from English into Dutch often contains significantly less of these modal particles than original Dutch prose (e.g. Jakobsen 1986, Spies 2009, Niemegeers 2010, Kool 2014), but there also exist high-quality Dutch translations of English literary texts that abound with these types of elements. In our corpus of literary texts, we investigated the use of the modal particles *dan, eens, even, maar, misschien, nou, soms* and *toch*, particles Vismans (1994) used in his seminal study on modal particles in directives. For comparison, we also investigated two focus particles, words typically used to highlight a part of the utterance while at the same time suggesting an alternative (König 1991).

(3) ***Alleen** FRED kocht een nieuwe auto*
     Only FRED bought a new car
     Nobody other than Fred bought a new car

We included *alleen* 'only' and *zelfs* 'even'. Both English and Dutch use focus particles in a comparable way (cf. König 1991).

Our aim was to analyze the usage of this small set of words in a corpus of novels originally written in Dutch between 2007 and 2012 and to compare this with the usage in novels written in English and translated into Dutch in the same period. Good translations in the Netherlands are expected to be fluent, while at the same time reflecting the foreign author's personality and intention. A good translation is supposed to read as if it is an original – the translator is invisible, the author and her intention are visible, but not the original linguistic and stylistic peculiarities (Venuti 2008). We hypothesize therefore that the relative frequencies of the selected modal particles in Dutch translations from English will in general agree with the distribution of these words in originally Dutch novels. However, we do expect possible differences in relative frequencies of modal particles in translations that do not aim for fluency ('domestication') but rather opt for 'foreignization', or in translations that are perceived as being

of low quality. Likewise, we do not expect any differences in the use of focus particles in Dutch and translations from English.

Methods and results
We used a part of the corpus of novels in the project The Riddle of Literary Quality (van Dalen-Oskam 2014). This corpus consists of the 401 bestselling titles marketed as fiction (and 'best-lending' from public libraries) from the period 2010-2012. The corpus mainly contains titles from three genres: suspense, romantic fiction, and literary fiction.

Genre	Literary fiction	Romantic fiction	Suspense
Originally Dutch	73	5	58
Translated from English	44	36	89
Total: 305	**117**	**41**	**147**

Table 1. Coverage of different genres in the corpus of novels

We applied a hierarchical cluster analysis to the relative frequencies of aforementioned particles, and extracted four groups of particles (see Figure 1).
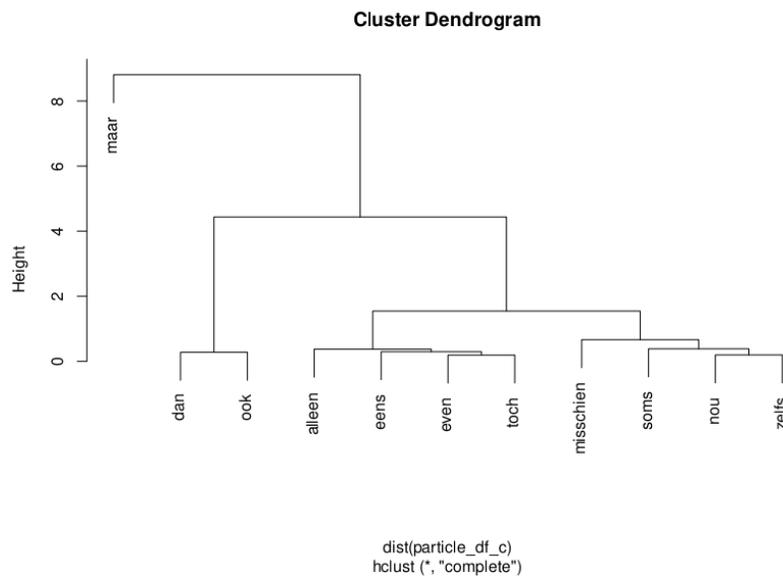

Fig. 1: Results of the hierarchical clustering model

We subsequently used these clusters as predictors in a logistic regression model, where we explored the extent to which the presence of particles from each of the four clusters is an indicator of an original Dutch work or of a translated novel. Results from the logistic regression

model (Figure 2) show that the more particles from clusters 1 and 2 are present in a novel, the more likely it becomes that it was originally written in Dutch, whereas a large number of particles from the other two clusters are an indicator of a translated novel. There is moreover an interaction between the particles from cluster 2 book genre: In a literary novel, these particles are not particularly good indicators of whether a book has been translated or not, but in suspense or romantic novels, these particles are better indicators.
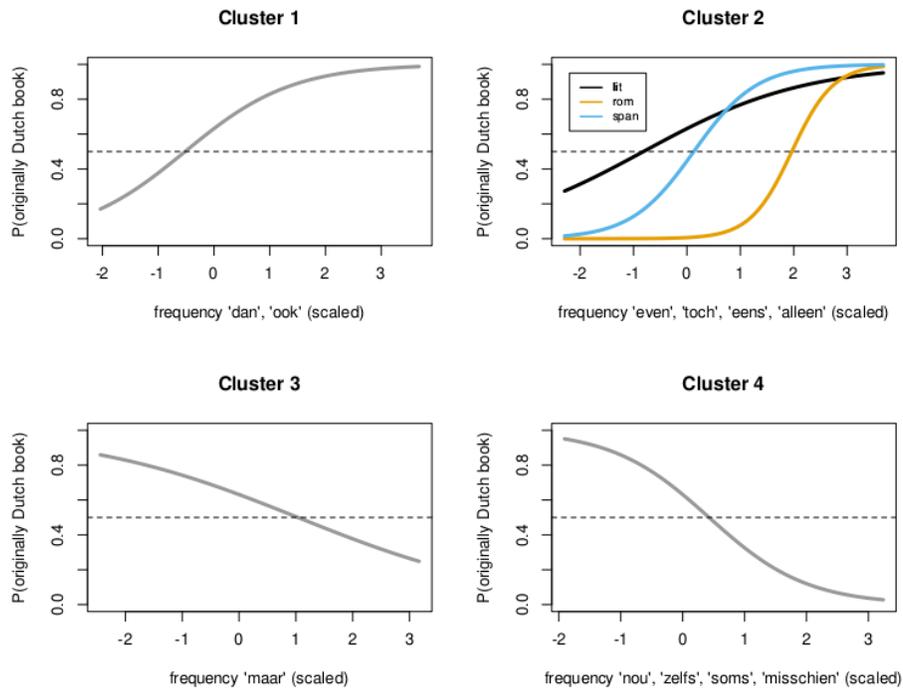

Fig. 2: Particle frequencies distinguishing between translated books and books written in Dutch.

We can therefore conclude that the relative number of particles present in a novel is a marker of whether or not the novel has been translated. Moreover, it is hardest to guess, on the basis of the prevalence of particles from cluster 2 (*even, toch, eens, alleen*), whether a literary work has been translated. This suggests that translators achieved the most authentic Dutch particle use for books in their translations of literary novels.

We double-checked this approach with a Principal Components Analysis (correlation) based on the eleven selected particles in all selected novels, using the Stylo Package for R, in order to see if the genre distribution agreed with the results of the previous cluster analysis. This proved to be the case (Figure 3).
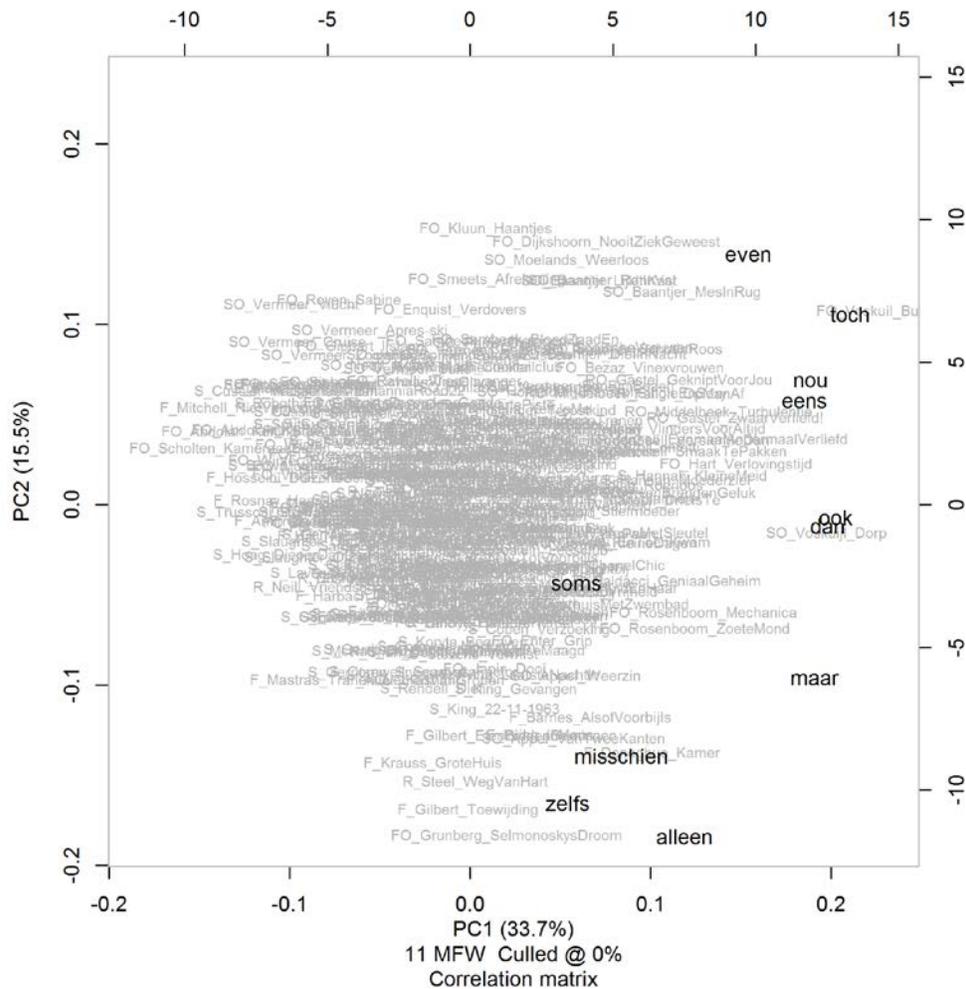
Fig. 3: PCA (correlation) of the use of the 11 particles using the Stylo Package for R.
Legend: F: Literary fiction, translated into Dutch from English; R: Romantic fiction, translated
into Dutch from English; S: Suspense fiction, translated into Dutch from English. FO, RO, SO:
the same genres, but originally written in Dutch.

Discussion and conclusion
From the perspective of translation studies, this suggests that the translated romantic and suspense novels tend to be less fluent in their use of modal particles than the literary translations. This could be the result of a deliberate choice of 'foreignization' by translators in these genres, or it could be related to a lower quality of a translation that was aimed to be fluent. Lower payment for translations of genre fiction, stricter deadlines, or the assumption that the readers would not notice anyhow may contribute to this situation. Examining these possibilities is beyond the scope of this contribution.

Bibliography

M. Eder, J. Rybicki, and M. Kestemont: Stylometry with R: a package for computational text analysis. *R Journal*, 8(1): 107-121, 2016. url: https://journal.r-project.org/archive/2016/RJ-2016-007/index.html.

A.L. Jakobson: Lexical selection and creation in English. *The third Nordic Conference for English Studies*, Stockholm, Almqvist, Wiksell, 1986.

E. König: *The meaning of focus particles. A comparative perspective.* London and New York, Routledge, 1991.

A.F.M. Kool: *Comparing particle frequency in original and translated Dutch*. MA thesis, Utrecht University Faculty of Humanities, 2014.

S. Niemegeers: *The Dutch modal particle 'wel' and its English counterparts. A corpus-based contrastive and translation study*. PhD Thesis, Ghent University, 2010.

J.-J. Spies: *Translating Faulkner's The Sound and the Fury How Translation Can Benefit from Stylistic Analysis*. MA Thesis, Utrecht University Translation Studies, 2009.

K. van Dalen-Oskam: Prehistory of The Riddle ('The Riddle of Literary Quality: The search for conventions of literariness', transl. of: The Riddle of Literary Quality. Op zoek naar conventies van literariteit', *Vooys: tijdschrift voor letteren* 32 (2014), 3, p. 25-33.) http://literaryquality.huygens.knaw.nl/?p=537#more-537

L. Venuti: *The translator's invisibility. A history of translation.* Second edition. London and New York: Routledge, 2008 (1st ed. 1995).

R. Vismans: *Modal particles in Dutch directives: A study in Functional Grammar.* PhD thesis, Free University Amsterdam, 1994.

T. van der Wouden: Partikels: naar een partikelwoordenboek voor het Nederlands. *Nederlandse Taalkunde* 7, 20-43, 2002.