

Moi! Are you from Twente or Groningen?

Emiel Dost

S4120140

Length: 3-4 pages

Abstract

This study investigates the usage of the dialectal greeting 'moi' in the Netherlands, specifically in the regions of Twente and Groningen. Dialectal greetings play a crucial role in cultural identity, and understanding their usage provides insights into regional linguistic variations. Utilizing a filtered Twitter corpus as a source of language data, this research aims to determine whether 'moi' is more commonly used in Enschede to represent Twente or in Groningen (city) to represent the province of Groningen. The methodology involves collecting and analyzing tweets containing the greeting 'moi' from both regions. The study hypothesizes that 'moi' is more prevalent in Groningen compared to Enschede, attributing this difference to historical, cultural, or linguistic factors. A key consideration is the temporal dimension, as 'moi' is reported to have a longer historical usage in Groningen (Stroop, 2009). While the dataset used for analysis may have limitations in representing the entire population, it serves as a valuable starting point for exploring regional dialectal variations. The study's findings may contribute to our understanding of how dialectal greetings evolve over time and differ across regions, shedding light on the linguistic landscape of the Netherlands.

1 Introduction

'Moi!' is an esteemed greeting or farewell in the eastern and northern part of the Netherlands. It even transcends national boundaries. In Germany, particularly near the Dutch border, it is adopted in

a dialectal manner as a form of greeting (Stroop, 2009). Notably, in the Finnish language, 'moi' translates to 'hi.'

In this research, we will focus on the Dutch usage of 'moi' to discern where it is more commonly used. The region Twente and the province Groningen both employ the dialectal word 'moi'. Since there is an extensive collection of tweets available, in which it is possible to search for words in a tweet and the location of a tweet, it can be determined whether 'moi' is more used in the region of Twente or in the province Groningen.

Considering variations in population size between the province of Groningen and the region of Twente, our study will narrow its focus to the cities of Groningen in the province of Groningen and Enschede in the region Twente. To compensate for the difference in population, the research should be done in relative usage (in percentages) of 'moi' in relation to 'hallo' between the two cities. By doing so, our goal is to reduce the effect of population differences and understand the subtle ways people greet each other in these areas.

The next sections of the study will help us understand how the cities of Groningen and Enschede utilize the word 'moi' and uncover the connection between them. Also, this paper will provide interesting and relevant history and background on the term 'moi' in Groningen en Twente.

2 Related Work

According to Verbeek (2019), there are two regional dialectal languages in the Netherlands. The first one is Limburgish, which is spoken in the southern area of the Netherlands. The second one is Low Saxon, which is mainly spoken in the north-eastern part of the Netherlands. Buurke et al. (2021) present a comprehensive map illustrating the languages spoken in the Netherlands.

The three major language divisions are shown in Figure 1. From these articles, we can assume that the foundation of the dialect in Groningen and Enschede have similarities.

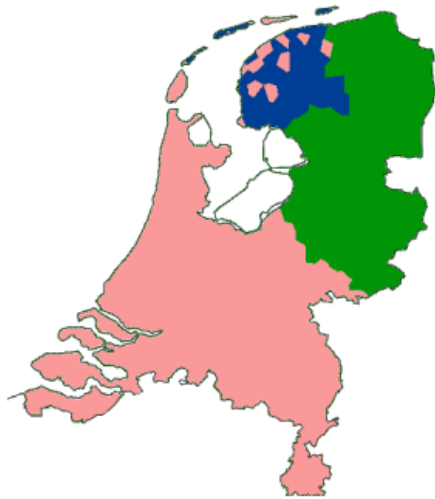


Figure 1: Major language divisions in the Netherlands. Pink: Low Franconian. Blue: Frisian. Green: Low Saxon (Buurke et al., 2021)

Stroop (2009) provides more insights into the origin and evolution of the greeting 'moi'. He states that the greet 'Moi' is a short amalgamation of 'Mooie morgen', translating to 'Nice morning'. Only the first part, 'moi', is left over. Since the 'morgen' disappeared from the expression, it was possible that 'moi' could be used as a more universal term. And so it evolves as a greet or farewell for every moment of a day, and not only for the morning.

The evolution and usage of the greeting 'moi' in the Netherlands offers an interesting linguistic narrative. The term does not have a long history in the Gronings dialectal vocabulary. Stroop (2009) states that the usage of the greet 'moi' was increasing in Groningen since around 1960. It is stated that it is mostly used in an informal-amicable atmosphere.

Also, Stroop (2009) observed that 'moi' is a widely used greeting in Twente, alongside other common expressions like 'moin,' 'moine,' '(t)jo,' '(h)eu(i),' and 'allo.'

Interestingly, despite its popularity, 'moi' did not originate in the vocabulary of Twente; instead, it made its way to the region from Groningen. This linguistic transfer underscores the connections of dialects within the Netherlands, showcasing how cultural and linguistic elements can traverse geo-

graphical boundaries and become integrated into new contexts.

In his study, Stroop (2009) explored and depicted the usage of various greetings and farewell expressions across the Netherlands. To gather information, he utilized data from a 1973 survey conducted by the dialectal center (Meertens-instituut). For the 2008 data, he relied on his own observations, marking locations on a map of the Netherlands whenever he heard someone using a greeting.

Interestingly, In Figure 2 we can see that in 1973 the greeting 'moi', marked with a '*', is only used in Groningen and for a bit in Drenthe. There are no occurrences in Twente.

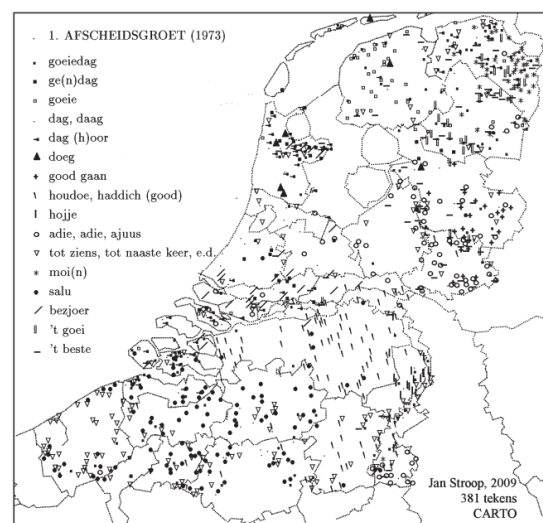


Figure 2: The usage of greetings linked to their location in 1973.

This differs from the results in Figure 3. Here we can see that in 2008 there are more occurrences of the greeting 'moi' in Twente in comparison to Groningen.

3 Data

The University of Groningen (RUG) has access to an large Twitter corpus, offering a substantial collection of tweets originating from individuals in the Netherlands. This corpus has been consistently updated since December 2010 and comprises a diverse range of tweet data and associated metadata.

For the purposes of this study, the specific time of tweet posting is not crucial. But since it is a large corpus, it is more efficient to take a sample of tweets out of the corpus. two distinct years can be chosen to get the information of tweets from.

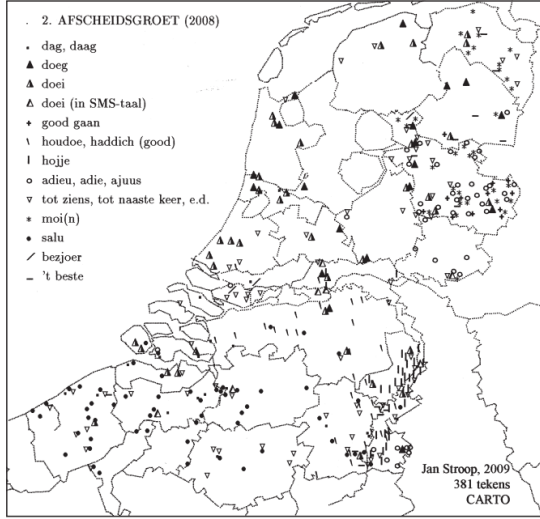


Figure 3: The usage of greetings linked to their location in 2008.

These yearly datasets can be downloaded from the servers of the University of Groningen using SSH and can be stored in a local directory in an Ubuntu environment. The filtered corpus could be subsequently processed with a bash program, which will focus on extracting tweets containing the words: 'moi', 'Moi', 'Hallo' and 'hallo'. These tweets will be stored in the Ubuntu environment.

Within the extracted tweets, a feature is the ability to search into the metadata and extract the location from which each tweet was sent. However, it is essential to acknowledge there are certain limitations in this approach. For instance, while a person may be tweeting from the city of Groningen or Enschede, their actual origin could be situated in a completely different part of the world. This nuance underscores the importance of interpreting location metadata with caution, recognizing that the physical location of a tweet's sender may not always align with the location indicated in the tweet itself.

Subsequently, the analysis will involve searching for occurrences of the words 'moi' and 'hallo'. The outcomes can be filled in in the Table 1, which is the contingency matrix for this research (with absolute values). In this context, the **independent variable** is the location the tweet is tweeted (Groningen, Enschede), while the **dependent variable** is the word to search (moi, hallo).

Pre-processing To account for variations in case usage (e.g., 'moi' or 'Moi,' 'hallo' or 'Hallo'),

	Groningen	Enschede	total
hallo	A	C	A + C
moi	B	D	B + D
total	A + B	C + D	A + B + C + D

Table 1: Template Contingency matrix with absolute values.

the search for these greetings in the tweets was conducted in a case-insensitive manner. This approach ensures that the analysis captures all relevant instances of the targeted words, maintaining the 2x2 structure of the contingency matrix.

Retweets are excluded from consideration as they do not reflect the original user's choice of words.

The filtered tweets can be saved in an output file, which then will counted using a bash script. Since each tweet occupied a single line, the number of lines in the output file corresponds to the tweet count.

All the data can be stored on the GitHub.¹

The outcomes can be displayed in Table 1, representing absolute values. To provide a relative perspective, these values can be transposed into Table 2.

	Groningen	Enschede
hallo	$\frac{A}{A+B} * 100$	$\frac{C}{C+D} * 100$
moi	$\frac{B}{A+B} * 100$	$\frac{D}{C+D} * 100$
total	100%	100%

Table 2: Template Contingency matrix with absolute values.

4 Predicted Results

In the context of this research, we anticipate that the greeting 'moi' will be more prevalent in Groningen compared to Enschede. This prediction is substantiated by the insights gathered from related articles discussed in the literature review.

Since 'moi' has been used in Groningen for several decades, it is highly plausible that it has firmly established itself in the linguistic landscape of Groningen. As indicated by Stroop (2009), the introduction of 'moi' to Twente occurred later, at least after 1973, as supported in Figure 2 and Figure 3 supports this claim.

An additional supporting argument is the linguistic diversity observed in Twente, where nu-

¹https://github.com/emield98/IRM_2024

merous words are employed for greetings and farewells, as opposed to the relatively smaller selection in the dialect of Groningen (Stroop, 2009). This suggests that when a person speaks in dialect, the choice between the words to greet or say goodbye there are fewer options and a person from Groningen is more likely to use 'moi'.

Discussion Our study aims to investigate the usage of the dialectal greeting 'moi' in the regions of Twente and Groningen in the Netherlands. Using a Twitter corpus that could be provided by the University of Groningen, we could conduct a quantitative analysis to examine the frequency of 'moi' compared to the more standard greeting 'hallo' in tweets from these regions.

The analysis involved filtering the Twitter corpus to extract tweets from Groningen and Enschede. We then searched for occurrences of the words 'moi' and 'hallo' in these tweets and recorded the absolute frequencies of each greeting in a contingency matrix.

The results can reveal interesting patterns in the usage of 'moi' and 'hallo' in Groningen and Enschede. Our analysis suggests that 'hallo' is the more commonly used greeting in both Groningen, compared to Enschede.

It is important to note that our analysis will be based on a specific sample of tweets from two years, and the results may not reflect the overall usage patterns of 'moi' and 'hallo' in these regions over an extended period. Further research with a larger and more diverse dataset would provide a more comprehensive understanding of the usage and cultural significance of these greetings.

In addition, it is important to consider potential inaccuracies in the metadata associated with the analyzed tweets. Factors such as VPN usage and people traveling between regions can lead to incorrect location information. These factors may impact the accuracy of our analysis and introduce biases in the results. Future research should explore alternative approaches to address these challenges and enhance the reliability of the findings.

5 Conclusion

In conclusion, this study aims to investigate the usage of the dialectal greeting 'moi' in the regions of Twente and Groningen in the Netherlands. Through the analysis of a Twitter corpus, we want to determine whether 'moi' is more commonly employed in Groningen or Enschede.

The expected results of our analysis suggests that 'moi' is relatively more frequently used in Groningen compared to Enschede. This is expected because the region Twente has multiple ways to greet someone in their dialect in comparison to Groningen. Also, Stroop (2009) investigated that in 1973, the greet 'moi' was only used Groningen, and not in Twente. In 2008, Stroop (2009) shows that Twente used the greet as well. So there is more history in Groningen with the usage of the greet 'moi'.

However, it is important to acknowledge the limitations of this study. The dataset used for analysis is limited to a specific timeframe and relied on location metadata associated with the tweets, which may contain inaccuracies due to VPN usage and traveling individuals.

To further enhance our understanding of the usage of 'moi' in these regions, the research should take place and future research should consider expanding the dataset to include a broader timeframe and explore alternative methods for verifying and validating the location information.

Despite the limitations, this study sheds light on the differences in the usage of 'moi' between Twente and Groningen. It serves as a foundation for further exploration and opens avenues for future research on the dialectal greetings in the Netherlands.

In summary, this study contributes to the understanding of regional dialects and greetings in the Netherlands, highlighting the significance of context and location in linguistic practices.

Lucky as we are with the double definition, we can end this report the same way it started. **Moi!**

References

- Buurke, R., H. Sekeres, W. Heeringa, R. Knooihuizen, and M. Wieling (2021). Estimating the level and direction of phonetic dialect change in the northern netherlands. Volume abs/2110.07918.
- Stroop, J. (2009). *Moi, adieë en salut. Groeten in Nederland en Vlaanderen*, Volume 10. Stichting Nederlandse Dialecten.
- Verbeek, E. (2019). The influence of dialects on the sociolinguistic perception of dutch accents.