Computational Political Landscape of the Netherlands and Prime Minister Schoof's Position

Wessel Ledder[1] and Iris Hendrickx[2]
[2]Centre for language and Speech Technology, Centre for Language Studies
[1,2]Radboud University, Nijmegen, The Netherlands
{wessel.ledder, iris.hendrickx}@ru.nl

Abstract

This study presents a computational model of the Dutch political landscape during the Schoof government period, constructed using debate speeches from the House of Representatives. We construct a two-dimensional representation of the Dutch political landscape by fine-tuning a BERT model on parliamentary debate speeches and applying dimensionality reduction techniques to the resulting embeddings. We evaluate the validity of this model by comparing it to an independently developed model from an external research institute, finding that both models reveal similar patterns along the socio-economic left–right dimension. We also examine content patterns and word frequency distributions in targeted samples located at distinct regions of the landscape to interpret the model. We further evaluate the stability of the landscape to ensure that the observed patterns are not driven by random variation. Finally, we position Prime Minister Schoof within this computational landscape. Schoof was intended to be a neutral Prime Minister without any party affiliation that would represent the coalition parties of the government equally. Our analysis will show whether Schoof was indeed neutral in his statements or not.

Keywords: political ideological landscape, parliamentary debate speeches, BERT embeddings

1. Introduction

Over the last years in the Netherlands, fragmentation of political parties is becoming more common in the Dutch political landscape (Sipma et al., 2021). The national election results of the last decade show a broad distribution of votes across multiple parties, with more parties getting a seat in the House of Representatives (HR)[1].

The latest government (July 2024 – June 2025) consisted of a coalition of four political parties (BBB, NSC, PVV and VVD) and the formation of this government required a considerable amount of time. As these parties could not agree on which party should deliver the prime minister, they opted for an external person that was not affiliated with any political party as to have a neutral prime minister to represent all coalition parties equally. The coalition appointed Dick Schoof, who had a background in civil service in the areas of Justice and Security, as Prime Minister (Van Holsteyn and Irwin, 2025).

In this study we aim to investigate the Dutch political landscape with computational methods and the role of Prime Minister Schoof in particular. Was Schoof really as neutral in his statements as was intended? We used Bidirectional Encoder Representations from Transformers (BERT) embeddings (Devlin et al., 2018), which carry semantic information of debate transcripts from the House of Representatives, as the basis for a computational political landscape of the political parties in the Netherlands. We aim to capture the underlying ideological political stances.

A political ideology is a set of ideas about how the society and economy should be organized. This definition is rather vague; it can be applied to a broad interpretation such as social-economic views of left versus right, or to much more narrow scope (e.g. in favor or against one United Europe). The ways in which political ideologies surface have been studied extensively, especially for the political system in the United States, e.g. (Poole, 2005; Diermeier et al., 2012). In this study we focus on the Dutch parliament. We investigated the following research questions:

- **RQ1:** How can we make a computational political landscape using debate speeches?
- **RQ2:** Where does Prime Minister Schoof fall on this landscape and is he truly neutral in his statements?

In this paper, we introduce a method to computationally create a model of the Dutch political landscape, based on speeches in the House of Representatives during the Schoof government. We do this by first fine-tuning a BERT model, after which we reduce dimensionality of the resulting BERT embeddings to create a two-dimensional landscape. We evaluate whether our political land-

[1]Recap of Dutch Parliament formation: during national elections, people vote for the political parties in House of Representatives (HR). Based on the number of votes, each party will get a number of seats in HR. As we do not have one party with the overall majority in the HR, a group of political parties agree on a coalition to form the active government consisting of the prime minister, the other ministers and the state secretaries.

Proceedings of ParlaCLARIN V Workshop on Interoperability, Multilinguality, and Multimodality in Parliamentary Corpora @ LREC 2026, pages 77–86
16 May 2026.

scape is interpretable and meaningful by comparing our landscape against a manually designed model that was created by an independent research institute (KiesKompas BV., 2023) and by inspecting the content and word frequencies in small samples in different corners of the landscape. We additionally test the stability of our semantic space to rule out random effects. Finally, we will place Prime Minister Schoof on our political landscape.

2. Related Work

Political debate speeches are a rich source to study the relation between language use and political ideologies. Schoonvelde et al. (2019) revealed structural differences in complexity of language use in political speeches from liberal versus conservative politicians. The claim that populist leaders use simple language in their speeches has also been investigated and this claim has been disputed (McDonnell and Ondelli, 2022; Zanotto et al., 2024). Neiman et al. (2016) studied language use of political speeches of Democrats and Republicans and their expressions of moral values. They did not observe systematic differences between the two parties in their study. Jordan et al. (2019) examined the language use of US Presidents and political leaders in other countries over a 20-year period to determine whether linguistic changes had occurred, particularly in analytical thinking and confidence. Using a method called LIWC (Tausczik and Pennebaker, 2010) to analyse patterns in word frequencies, they found an overall decline in analytical thinking and an increase in confidence during this period.

Furthermore, previous work has shown earlier attempts to create computational models of political landscapes. For example, Poole (2005) showed that voting decisions in the US congress can be modelled with a one-dimensional statistical model that expressed the dimensions of liberal versus conservative values. Diermeier et al. (2012) have shown credible evidence that political debate speeches from the US congress are expressing these ideologies in such a consistent way that one can train a machine learning classifier on speeches from conservative and liberal politicians to predict the position on this dimension for new unseen speeches. Their study also revealed that politicians tend to use certain terms that are characteristic for their ideology and that such terms are mostly related to culture rather than the economy.

Previous studies showed that word embeddings can be used effectively to capture underlying constructs such as political ideologies, to offer a unified framework for analysing political language (Rheault and Cochrane, 2020) or to model stances in political debates (Konjengbam et al., 2018).

These earlier findings form the underpinning for our approach where we aim to use parliamentary speeches as the source for modelling the underlying ideologies of Dutch political parties.

We did not use the obvious alternative source for getting insights in the ideologies from Dutch political parties, the political party program reports. This source has been used to construct a two-dimensional representation of the Dutch political landscape by the research institute Kieskompas (KiesKompas BV., 2023) in the run-up for the 2023 Dutch national elections. This landscape is created manually and consists of a social-economic left–right dimension and a cultural conservative–progressive dimension based on a list of stances that pertains relevant topics for the 2023 elections and can help voters to get insights in their own stances in light of the stances of the political parties on these topics (Wall et al., 2014). We will use the Kieskompas 2023 political landscape as an external source to validate the political computational landscape that we are creating in this study.

3. Method

We created a digital representation of the Dutch political landscape by fine-tuning a BERT embedding model on political debate speeches. We describe the model and fine-tuning in more detail in the next section. We evaluate whether this model indeed captures the political ideologies from the Dutch political parties with the following steps. We apply factor analysis (Spearman, 1904) to reduce the high-dimensional embedding vectors to two latent semantic dimensions that organize the embedding space. We compare this reduced space to an existing political model manually created by the company Kieskompas (KiesKompas BV., 2023). We validate the stability of the landscape as detailed in section 3.2. To answer our second research question on where Prime Minister Schoof falls on this computational political landscape, we collected the speeches of Schoof separately from the other debate speeches and project them into the embedding space to reveal which political parties are closest to his statements.

We first present the dataset that we use in our studies in the next section, followed by the details of the experimental setup to create and validate the political landscape.

3.1. Data

The data used in this study are the plenary reports of the House of Representatives (Tweede Kamer der Staten Generaal, 2024), in the period of the Schoof government (July 2nd, 2024 to June 3rd, 2025). These reports are the transcripts of the

#	Text	Translation	Party	Speaker
1	(...) Het is belangrijk dat we bij zorg rekening houden met de diversiteit in onze zorg. Daarom de volgende motie. De Kamer, gehoord de beraadslaging, constaterende dat mantelzorgers vaak zorg krijgen van hulpverleners met verschillende culturele achtergronden; overwegende dat deze diversiteit helpt om beter in te spelen op de behoeften in de ouderenzorg; verzoekt de regering om inclusiviteit en diversiteit in de zorg en mantelzorg actief te ondersteunen en mee te nemen in beleid, en gaat over tot de orde van de dag.	(...) It is important we keep in mind diversity in healthcare. Hence the next motion. The House, having heard the deliberation, noting that informal caregivers often work with healthcare professionals from different cultural backgrounds; considering that this diversity helps to better respond to the needs in elderly care; calls on the government to include inclusivity and diversity in their policy for healthcare and informal caregiving, and moves on to the order of the day.	DENK	De heer El Abassi
2	(...) De PVV vindt dat ons belastinggeld in de eerste plaats moet worden uitgegeven aan Nederland en aan de Nederlanders. We zijn er trots op dat er vandaag eindelijk gehoor wordt gegeven aan deze oproep. Dit kabinet, deze coalitie, zet nu echt forse stappen: een forse bezuiniging van 300 miljoen in 2025, die oploopt tot maar liefst 2,4 miljard in 2027. Daar komt in 2027 ook nog eens de bezuiniging van 1,6 miljard op de eerstejaarsopvang voor asielzoekers bij. Die wordt ook betaald uit ontwikkelingsgeld. Van het strengste immigratiebeleid ooit naar de grootste bezuiniging ooit op ontwikkelingshulp: dat is bij elkaar een prachtig mooie bezuiniging van 4 miljard op ontwikkelingshulp. Dat klinkt alle rechtse, hardwerkende Nederlanders, onze Henk en Ingrid, die een sterke overheid willen die hun belangen wél behartigt, als muziek in de oren.	(...) The PVV believes that our tax money should primarily be spent on the Netherlands and the Dutch people. We are proud that today, this call is finally being heeded. This cabinet, this coalition, is now taking major steps: a major cut of 300 million in 2025, rising to no less than 2.4 billion in 2027. In 2027, there will also be an additional cut of 1.6 billion for first-year asylum seeker reception, which is also funded from development aid. From the strictest immigration policy ever to the biggest cut on development aid: that together is a nice cut of 4 billion on development aid. That sounds like music to the ears to all right-wing, Dutch citizens, our Henk and Ingrid, who want a strong government that truly looks after their interests.	PVV	De heer Ram
3	(...) Afgelopen dinsdag sprak ik Jan uit Rotterdam, die zei: "Ik werk in de gehandicaptenzorg en die 310 miljoen euro bezuinigingen daarop ... Het kan gewoon niet!" En u gooit er deze kabinetsperiode nog een keer in totaal 4,6 miljard aan bezuinigingen op de zorg voor mensen bovenop. U doet dat, meneer Wilders. (...)	(...) Last Tuesday, I spoke with Jan from Rotterdam, who said: "I work in disability care, and these 310 million euros of cuts ... It is just not possible!" And during this cabinet's term, you add another 4.6 billion of cuts to people's healthcare. You are doing that, mister Wilders. (...)	SP	De heer Dijk
4	Ik heb vertrouwen in het kabinet-Schoof, de introductie die minister-president Schoof hier net hield, in zijn ethos, zijn wil en ook zijn trackrecord om de rechtsstaat te respecteren. En ja, ik denk dat hij een goede minister-president is en dat deze ploeg het in zijn geheel in zich heeft om die rechtsstaat te respecteren.	I have confidence in the Schoof cabinet, the introduction that Prime Minister Schoof just delivered here, in his ethos, his determination and also his track record in respecting the rule of law. And yes, I believe he is a good Prime Minister, and that this group as a whole is capable to respect the rule of law.	NSC	De heer Omtzigt

Table 1: Exemplar data. For each speech in the debate, the speech, speaker and respective party is stored. Note that the English translations were not part of the political landscape, and only used in our automatic analysis of the landscape.

political debates in the HR. The spokespersons from the political groups in the HR discuss with the ministers or state secretaries or with each other in these debates. We assume that the collection of debate speeches of all people from the same party together represent the stances of their political party and reflect their ideological stances.

The transcripts of the debates were manually typed out by transcribers from the government. This has the effect that even though these are transcribed speeches, these transcripts do adhere to the writing style conventions of using sentence boundaries and punctuation. Typical speech elements such as fillers or laughter occur sparsely. After scraping the data, we first filtered out any metadata so that only the transcripts of the report are kept. We removed any speeches from the chair of the HR, Martin Bosma, as the chair leads the debate, and never debates for his own opinion in the debate. Next, we matched each transcript to their speaker and their respective party. Any speeches containing less than 15 words were omitted to ensure that each speech sample has at least some semantic content. This resulted in a total of 42882 speech samples with an average length of 152.8 words.

Some speech examples can be found in Table 1 and demonstrate that some words can strongly show the ideological tendency. For example, in speech 1, words such as 'diversity' and 'inclusivity' tend to be uttered by left-wing parties, while in speech 2, words such as 'the Dutch people' and 'cut' are more associated with right-wing parties. However, these words alone do not show ideology, as the context is important: in speech 3, the word 'cuts' is mentioned twice, but the speaker is against the cuts. Finally, speech 4 shows an example where the ideology of the speaker is less clear from this speech utterance.

Due to the non-uniform representation of parties in the House of Representatives, the distribution of the number of speeches per party is skewed, as is visible in Figure 1. The political parties with the lowest amount of speech samples in Figure 1 have only a few speakers in the debates: JA21 (1 seat), Volt (2 seats), CU, SGP, FVD, PvdD and

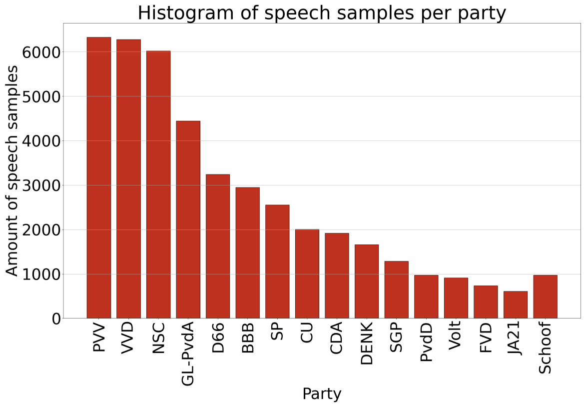

Figure 1: Distribution of total speeches per party.

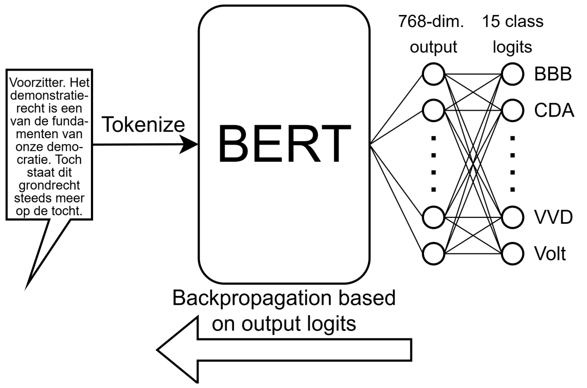

Figure 2: Graphical representation of the fine-tuning process. During training, the model must predict the party with each speech sample in its classification layer, allowing all weights to be fine-tuned accordingly.

DENK (3 seats) (Van Holsteyn and Irwin, 2025). We refer to Appendix A for the full name of each party, as well as the European Parliament group each party is affiliated to. The distribution of speakers within each party is also uneven, as party leaders, ministers and state secretaries generally have a larger role in the political debates.

3.2. Experimental Setup

For answering both research questions and to make a computational political landscape, we decided to use the Bidirectional Encoder Representations from Transformers (BERT) model (Devlin et al., 2018), which is a transformer neural network model, capable of transforming an input text sample into a numeric vector representation while capturing the semantics of the input text as a higher level embedding. These can in turn be used to create the landscape, as will be discussed later. In particular, we used a Dutch version of BERT, called BERTje (De Vries et al., 2019), which has been pre-trained to embed Dutch text.

For our task we fine-tuned the BERTje model on the stances of the Dutch political parties. The default embedding space that resulted from pre-training BERTje on a large and diverse training set tends to cluster words with similar meaning close together in the semantic space. However, for our digital political landscape we do not aim to cluster together on topics (like climate change or migration) but focus on ideological perspectives.

To achieve this, we fine-tuned the BERTje model to learn the stances and ideological perspectives of the political parties. First, we randomly split the dataset in a training set (80%), a validation set (10%) and a test set (10%), while keeping apart any speeches from Prime Minister Schoof in a separate dataset. To fine-tune the BERTje model weights, we first append a classification layer to the model, with one node for each party. Then, using the training set, we let the model predict the political party of the speaker of each speech. Weights are updated based on a cross-entropy loss function. The classification performance is evaluated with the validation set during fine-tuning, where after early stopping, we achieve an accuracy of 0.59 on the training set, and an accuracy of 0.42 on the validation set[2]. Note that for a complex linguistic 15-label classification task, the accuracy is expected to be low. The fine-tuning process is shown graphically in Figure 2.

Each sample from the training set is fed through the fine-tuned BERTje model (without classification layer), resulting in a 768-dimensional output vector embedding, which gives an aggregate representation of the full input, following the approach of Devlin et al. (2018). Factor analysis is used to project the high-dimensional vector embeddings of the training samples into two latent semantic dimensions that structure the embedding space. We then created the political landscape by projecting the samples from the held-out test set into the reduced space using the projection fitted on the training set. We computed the centroids for each political party based on the mean embeddings of speech samples from political party members in the reduced space.

3.3. Evaluation

We evaluate the resulting political landscape in various ways. First, we verify whether the represen-

[2]For any specifics on training or hyperparameters, please refer to code and data in our Github repository at github.com/wledderw/ComputationalPoliticalLandscape.

tations in the political landscape are stable. We did not only create the political landscape for the held-out test set, but created another version of the space using the training set data. We compare the resulting landscapes and the centroids per party of both the training and test set to see whether the centroid positions in relation to each other stay similar when using different data samples to create the landscape. Furthermore, we validate whether factor analysis provides an interpretable lower-dimensional representation of the parties, and hence a political landscape that indicates the underlying ideologies. Moreover, we validate our computational political landscape by comparing it to a manually-made political landscape by the company Kieskompas (KiesKompas BV., 2023).

We also inspected a random sample of the speech samples in different corners of the semantic space to interpret what that the underlying dimensions represent. While the first latent factor gave us a clear interpretation, the second latent factor was not so straightforward. We tried TF*IDF word ranking (Sparck Jones, 1972) to see if we could find words that might be important to distinguish this dimension direction, but that did not lead to any insights. Our manual inspection of the random sample did indicate that we saw more substantiated argumentations in the bottom speeches in the second dimension and more unfounded assertions or irrational oppositions at the top of the dimension. Linguistic Inquiry and Word Count (LIWC) is an automatic text analysis tool developed by psychologists to count words and group these in psychologically meaningful categories. Analysing someone's writing or speech style can give insights in emotions, social relationship and thinking styles (Tausczik and Pennebaker, 2010). Therefore, we applied LIWC-22 (Boyd et al., 2022) as this can be used to measure the amount of 'analytical thinking' in texts which seemed to fit with our intuition on how to interpret this dimension. We applied LIWC in the following way: we took the top 200 and bottom 200 speech samples from the second dimension of the test set, we translated speech transcriptions to English[3] and applied LIWC.

For answering research question 2, finding Prime Minister Schoof's place on this landscape, we use the same procedure as stated above to find Schoof's average embedding and hence his location on the political landscape. Note that we did not fine-tune the BERTje model to Schoof's speeches, allowing us to compare Schoof with different parties on the embeddings space.

[3]Translations adapted from Google Translate, accessed February 17, 2026, https://translate.google.com/.

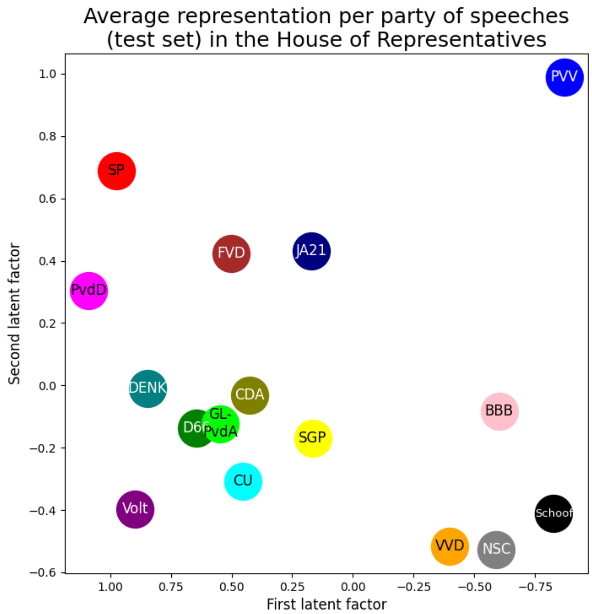

Figure 3: Computational political landscape based on samples of the test set.

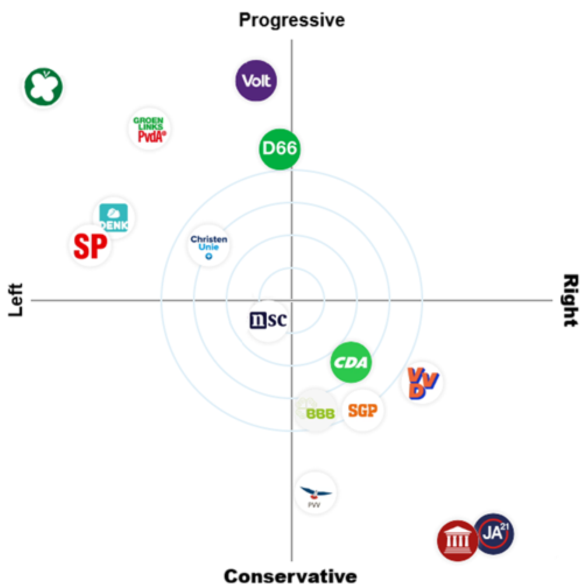

Figure 4: Political landscape created by Kieskompas for the 2023 Dutch elections (KiesKompas BV., 2023). For most parties, the acronym of the party is clear from the logo, except for the butterfly (PvdD) and the temple (FVD). Note that we removed parties from this political landscape that are not in the HR.

4. Results

We present the computational political landscape that is the result from the factor analysis on the fine-tuned BERTje model applied to the test set reduced to its two main dimensions, in Figure 3.

Let us start with the position of Prime Minister Schoof in this political landscape. Schoof is located closest to the coalition party NSC in the bottom right corner of the landscape. We can observe that Schoof is also close to coalition parties VVD and BBB.

First landscape dimension We observe in the first dimension (x-axis) that there is a clear separation between the coalition parties (BBB, NSC, PVV and VVD) on the right side and the other opposition parties on the left side. Since the coalition is right-leaning, this dimension also reflects a left–right spectrum.

When we compare our first dimension to the left–right dimension in the political landscape in Figure 4 that has been designed manually by KiesKompas BV. (2023), we can also see a large overlap in orientation. In both plots, parties with a social-economic left ideology (PvdD, SP, DENK, Volt) are on the left-most side while the right oriented parties PVV, VVD and BBB are on the right side of the first latent dimension[4]. We also observe differences between the two plots. The parties JA21 and FVD are much further to the right in the Kieskompas landscape then in our landscape, while for NSC we see the opposite as it is placed to the right in our landscape while the Kieskompas placed it in the middle of the left–right dimension.

Second landscape dimension The second latent factor is more difficult to interpret. It does not match with progressive–conservative axis in the Kieskompas political landscape shown in Figure 4. On the second latent factor (y-axis), we see a cluster of parties BBB, CDA, D66, DENK, GL-PvdA and SGP in the middle to bottom region of the axis. Parties FVD, JA21, PvdD, PVV and SP are located at the top of the second latent dimension, while CU, NSC, Volt and VVD are found at the bottom of the y-axis.

After manual inspection of a random sample on the y-axis, we noticed that the top speech samples contained more emotional arguments while speeches at the bottom seemed to consist of supported and critical arguments. To study this in a more systematic way, we computed the LIWC-22 (Boyd et al., 2022) 'analytical thinking' summary variable for the top 200 (higher latent factor values)

[4]Note that we have reversed the axis of the first latent factor in our landscape to match left and right parties to the left and the right of the plot.

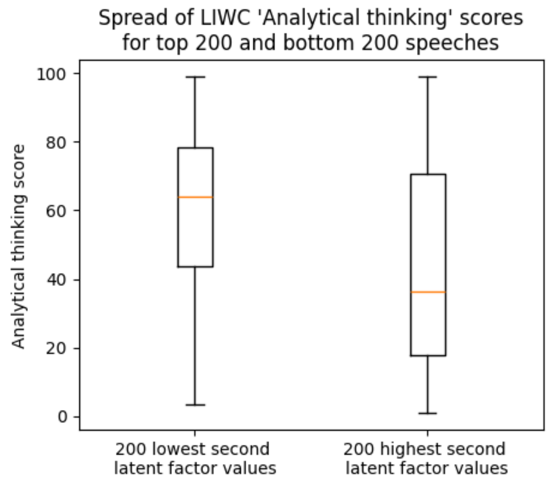

Figure 5: Spread of LIWC analytical thinking scores for the two extremes of the second dimension of the political landscape.

and bottom 200 samples (lower latent factor values) from the y-axis. The spread of the scores for each of the two sets of samples can be found in Figure 5. For the speeches of lower latent factor values, 'analytical thinking' had an average score of 59.8, while the speeches for higher latent factor values had an average score of 44.0, on a scale running from 1 (personal and intuitive) to 99 (formal and logical). Hence, we hypothesize that our second latent factor indicates the degree of analytical thinking, where a lower second latent factor means a higher degree of analytical thinking.

4.1. Stability of the political landscape

We investigated the stability of the landscape in two ways: we reviewed the stability of the semantic space based on different sets of speech samples, and we looked at the dimensionality reduction technique.

Stability over different samples First, we verified whether the semantic space created on the basis of the held-out test set samples keeps the relative positions between the political party centroids in similar positions when compared to the landscape created on the basis of the training set.

The semantic space has been trained using the training set, and the factor analysis dimensionality reduction also has been fitted on the embedded vectors of the training set. To analyse whether our model is stable, we plot the centroids for each party of the training set and the test set in the same plot, as is visible in Figure 6. The centroids from the training set are more transparent than the centroids from

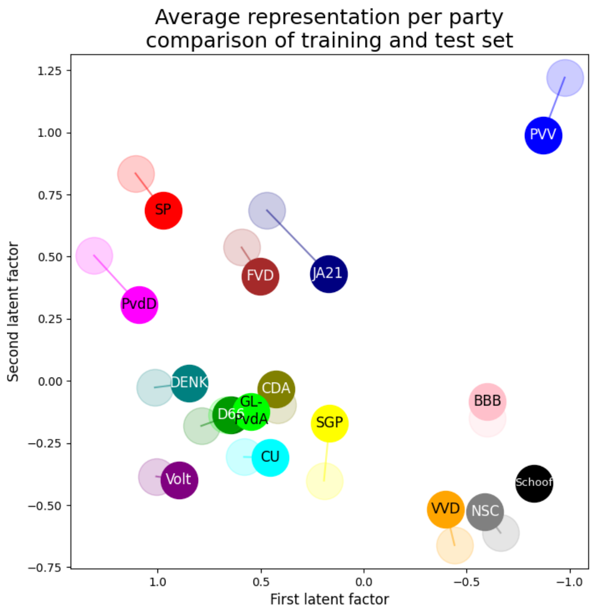

Figure 6: Comparison of computational political landscape for the training and test set. The transparent colours show the locations on the political landscape of parties by using speeches from the training set, while the opaque colours show the locations on the political landscape of the parties by using speeches from the test set.

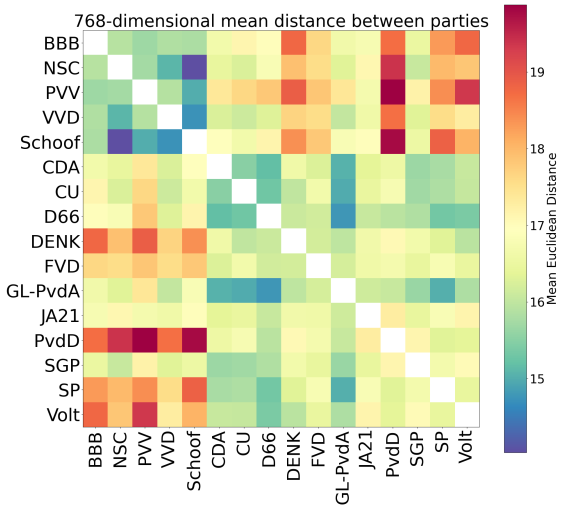

Figure 7: A Euclidean distance matrix between parties in the 768-dimensional semantic space. Note: coalition parties (BBB, NSC, PVV, VVD) and Prime Minister Schoof have been put together, the rest of the parties are sorted alphabetically.

the test set. In the plot, we see that the centroids for most parties move between the training set and the test set towards a central point. Due to this, the relative order of the parties does not change much. Hence, we declare our computational landscape to be stable from this perspective.

Stability of dimensionality reduction We want to know whether the two-dimensional landscape resulting from factor analysis keeps the relative distance between parties intact. Hence, we compare the distance between the parties in the two-dimensional landscape of Figure 3 to the Euclidean distances between the party centroids for the full 768-dimensional semantic space, as is displayed in the distance matrix in Figure 7.

The distance matrix shows that the coalition parties are located close to each other, and further away from any of the opposition parties. This is also visible in our political landscape, where there is a big gap over the first latent factor between the coalition and opposition parties. The parties that are located furthest from the coalition on the political landscape (DENK, PvdD, SP, Volt) also have the largest distances to the coalition in high-dimensional space. Moreover, we see a cluster of opposition parties CDA, CU, D66 and GL-PvdA. In the mean distance matrix, we also see that these parties are close together.

5. Discussion

Our results showed that we were able to create a computational political landscape on the basis of the parliamentary debate speeches for the Schoof government period.

5.1. RQ1 Creation of the Political Landscape

We made a computational political landscape by fine-tuning BERT and reducing dimensionality using factor analysis. The resulting landscape is stable over an unseen test dataset, and the dimensionality reduction keeps relative distances between parties intact.

While analysing the first latent factor in Figure 3, a clear divide between opposition parties and the four coalition parties is visible. When looking at the order of the opposition parties from left to right, we see it closely resembles the left-wing to right-wing axis of Kieskompas's political landscape in Figure 4. It also is logical that the right-wing and centrist parties are closer to the coalition than the left-wing parties, as the coalition is also right-winged (Van Holsteyn and Irwin, 2025).

NSC was closer to the middle in the Kieskompas landscape while in our landscape it was located on the right of the first dimension. This might be explained by the difference in source material and timeline on which the two political landscapes were created. The Kieskompas used the party programs written before the elections in 2023 while we base our landscape on speeches from the period after the coalition formation where NSC clearly chose the right-oriented direction (van den Berg, 29 November 2024).

We would have expected that JA21 and FVD (the temple icon right under in Figure 4) were placed much further to the right in our political landscape than they actually were, as these are known as strongly right-oriented parties. A possible explanation could be that these parties have the least amount of speech samples as was shown in Figure 1, which limited the speakers' overall contribution to the semantic space.

Analysis of the second latent factor was more difficult. Manual examination of the individual speeches revealed a clear pattern of less grounded speeches that sometimes attack other politicians, to grounded speeches based on arguments. After running LIWC over the speeches, we found that the 'analytical thinking' score corresponds to our findings. Hence, we give the second axis an analytical thinking scale, where a lower second latent value means a higher degree of analytical thinking. Many of the parties that are on the less analytical side (or more intuitive/personal side) of the landscape are populist parties (FVD, JA21, PVV and SP (Rooduijn, 2021). As Zanotto et al. (2024) claimed, populists tend to use rhetorical strategies in their speeches. These rhetorical claims are usually not supported by arguments, and hence these parties end up at the side of the analytical thinking axis representing lower levels of analytical thinking. This is in line with the findings of the study of Jordan et al. (2019) who showed that populist US president Trump is on the lower end of the analytic thinking scale of the LIWC tool.

5.2. RQ2 Schoof in the Political Landscape

Our results showed that both in the high dimensional representation and in the reduced semantic space, Schoof was located most closely to NSC. The Euclidean distance matrix (Figure 7) also indicated that Schoof was positioned close to all coalition parties, while those parties were farthest from DENK, PvdD, SP and Volt, which are all left-oriented. When we soleley look at the first latent factor, the left–right opposition–coalition dimension, we can safely assume that the statements of Schoof were well in line with the statements from the right-oriented coalition parties. However, when we examine the full distance matrix in detail, we find that Schoof is closest to NSC among the coalition parties. The landscape also shows a large distance on the second analytical-thinking dimension between Schoof and PVV. From this perspective, we can conclude that although Schoof was ideologically neutral within the coalition, his analytical-thinking levels aligned closely only with NSC and VVD.

6. Conclusion

We made the following contributions: we created a stable computational model of the Dutch political landscape during the Schoof government period based on debate speeches from the House of Representatives. We validated this model against a model created by an independent research institute and confirmed that both models show similar trends on the socio-economic left–right axis. Finally, our analysis indicates that, based on his parliamentary statements, Prime Minister Schoof can be described as ideologically neutral within the coalition, yet his analytical-thinking style does not align with both BBB and PVV.

Our study has several limitations. The data distribution for the different parties in our model was skewed and not all parties were represented equally. When reducing the multi-dimensional space to two dimensions, many aspects are lost and were not taken into account in our analysis. In the comparison against the Kieskompas landscape, we use both a different data source — political party programs and debate speeches — and a different time frame as the Kieskompas was constructed before the elections and the debates cover the period after the formation process. As an additional validation we could use the political party programs to create a BERT model and compare that against the manually constructed Kieskompas and our current landscape.

We envision several paths for future work. We would like to study party dynamics by zooming in on prominent party members and study how close or distant they are from their average party representation: do we see closely clustered members or are they perhaps closer to other parties? Furthermore, we would be very interested to expand our analysis over a longer timeline and investigate whether we could predict some of the dynamics of changes in the political landscape such as the fusion between GL and PvdA (Van Holsteyn and Irwin, 2025) or the separation by Pieter Omzigt from CDA to start the new party NSC (NOS Nieuws, 20 August 2023).

This study focused on the Dutch parliamentary debates of the Schoof government but the proposed method is certainly easily applicable to datasets from other countries and parliaments as the approach is language independent.

7. Acknowledgments

We would like to thank Prof. James Pennebaker for his advice and suggestions for our study. This work contributed to the HAICu project with file number NWA.1518.22.105 which is financed by the Dutch Research Council (NWO).

8. Bibliographical References

Ryan L Boyd, Ashwini Ashokkumar, Sarah Seraj, and James W Pennebaker. 2022. The development and psychometric properties of LIWC-22. Technical report, University of Texas at Austin, Austin, TX. https://www.liwc.app.

Wietse De Vries, Andreas van Cranenburgh, Arianna Bisazza, Tommaso Caselli, Gertjan van Noord, and Malvina Nissim. 2019. BERTje: A Dutch BERT Model. *arXiv preprint arXiv:1912.09582*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv:1810.04805*.

Daniel Diermeier, Jean-François Godbout, Bei Yu, and Stefan Kaufmann. 2012. Language and ideology in congress. *British Journal of Political Science*, 42(1):31–55.

Kayla N. Jordan, Joanna Sterling, James W. Pennebaker, and Ryan L. Boyd. 2019. Examining long-term trends in politics and culture through language of political leaders and cultural institutions. *Proceedings of the National Academy of Sciences*, 116(9):3476–3481.

KiesKompas BV. 2023. Kieskompas Tweede Kamer verkiezingen in 2023. https://tweedekamer2023.kieskompas.nl/.

Anand Konjengbam, Subrata Ghosh, Nagendra Kumar, and Manish Singh. 2018. Debate Stance Classification Using Word Embeddings. In *Big Data Analytics and Knowledge Discovery*, pages 382–395, Cham. Springer International Publishing.

Duncan McDonnell and Stefano Ondelli. 2022. The Language of Right-Wing Populist Leaders: Not So Simple. *Perspectives on Politics*, 20(3):828–841.

Jayme L. Neiman, Frank J. Gonzalez, Kevin Wilkinson, Kevin B. Smith, and John R. Hibbing. 2016. Speaking different languages or reading from the same script? word usage of democratic and republican politicians. *Political Communication*, 33(2):212–240.

NOS Nieuws. 20 August 2023. Omtzigt doet mee aan verkiezingen met eigen partij: Nieuw Sociaal Contract.

NU.nl. 2 June 2024. Welke partijen zitten voor welke fractie in het Europees Parlement?

Keith T. Poole. 2005. *Spatial Models of Parliamentary Voting*. Analytical Methods for Social Research. Cambridge University Press.

Ludovic Rheault and Christopher Cochrane. 2020. Word Embeddings for the Analysis of Ideological Placement in Parliamentary Corpora. *Political Analysis*, 28(1):112–133.

Matthijs Rooduijn. 2021. Populisme, Nederland en verkiezingen. https://www.uva.nl/shared-content/faculteiten/nl/faculteit-der-maatschappij-en-gedragswetenschappen/nieuws/2021/02/verkiezingen-populisme-nederland-en-verkiezingen.html.

Martijn Schoonvelde, Anna Brosius, Gijs Schumacher, and Bert N Bakker. 2019. Liberals lecture, conservatives communicate: Analyzing complexity and ideology in 381,609 political speeches. *PloS one*, 14(2):e0208450.

Take Sipma, Marcel Lubbers, Tom Van der Meer, Niels Spierings, Kristof Jacobs, et al. 2021. Versplinterde vertegenwoordiging. Nationaal Kiezersonderzoek 2021.

Karen Sparck Jones. 1972. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1):11–21.

Charles Spearman. 1904. "General Intelligence" Objectively Determined and Measured.

Yla R. Tausczik and James W. Pennebaker. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, 29(1):24–54.

J.Th.J. van den Berg. 29 November 2024. Het treurig lot van NSC. https://www.parlement.com/column/202411/het-treurig-lot-van-nsc.

Joop JM Van Holsteyn and Galen A Irwin. 2025. The Dutch parliamentary elections of November 2023. *West European Politics*, 48(2):464–477.

Matthew Wall, André Krouwel, and Thomas Vitiello. 2014. Do voters follow the recommendations of voter advice application websites? A study of the effects of kieskompas.nl on its users' vote choices in the 2010 Dutch legislative elections. *Party Politics*, 20(3):416–428.

Sergio E. Zanotto, Diego Frassinelli, and Miriam Butt. 2024. Language Complexity in Populist Rhetoric. In *Proceedings of the 4th Workshop on Computational Linguistics for the Political and Social Sciences: Long and short papers*, pages 61–80, Vienna, Austria. Association for Computational Linguistics.

9. Language Resource References

Tweede Kamer der Staten-Generaal. 2024. *Plenaire verslagen*. Retrieved December 22, 2024 from `https://www.tweedekamer.nl/kamerstukken/plenaire_verslagen`.

A. Appendix A. Acronyms of political parties

Acronym	Party name	English translation	EP group
BBB	BoerBurgerBeweging	Farmer-Citizen Movement	EPP
CDA	Christen-Democratisch Appèl	Christian Democratic Appeal	EPP
CU	ChristenUnie	Christian Union	EPP
D66	Democraten 66	Democrats 66	RE
DENK	DENK	DENK	-
FVD	Forum voor Democratie	Forum for Democracy	NI
GL-PvdA	GroenLinks - Partij van de Arbeid	GreenLeft - Labour Party	G/EFA - S&D
JA21	Het Juiste Antwoord ’21	The Correct Answer ’21	ECR
NSC	Nieuw Sociaal Contract	New Social Contract	EPP
PvdD	Partij voor de Dieren	Party for the Animals	Left
PVV	Partij voor de Vrijheid	Party for Freedom	PfE
SGP	Staatkundig Gereformeerde Partij	Reformed Political Party	ECR
SP	Socialistische Partij	Socialist Party	Left
Volt	Volt	Volt	G/EFA
VVD	Volkspartij voor Vrijheid en Democratie	People’s Party for Freedom and Democracy	RE

Table 2: The full version of the acronym for each party name, the translation of the party name to English, and their European Parliament (EP) group as of the start of the Schoof government in 2024 (NU.nl, 2 June 2024).