

Multimodal Reference by Means of the Pronoun *We* and Hand Gestures in a Novel Corpus of Parliamentary Opening Debates

Costanza Navarretta
University of Copenhagen
Njalsgade 76, bygning 4A,2
2300 Copenhagen S
costanza@hum.ku.dk

Abstract

Political discourse has persuasion as its main goal and the identification of the referents of pronouns in it is of great importance. This paper presents a novel multimodal corpus of Danish parliamentary opening debates. It also describes a study of multimodal reference by means of the first-person plural pronoun *vi* (we) and the co-occurring hand gestures in a subset of the corpus. The data in the study consists of 219 speeches of two prime ministers from the opening debates in 2013, 2014, and 2021-2024. In the speeches, the prime ministers answer questions of parliament members from the opposition. The uses of the first-person plural pronoun in political speeches are particularly interesting since the pronouns can refer to different groups, such as the government, the parliament, the country, or a specific party and can be used by politicians to achieve consensus or distinguish their politics from that of others. The main hypothesis we want to investigate in the study is whether the pointing gestures vary in their trajectory depending on the intended referents. The results of our study confirm this hypothesis for the most frequent referent types and show how pointing hand gestures are used by the two prime ministers to help their audience individuating the correct referents of *vi*, and emphasise them. Our data also indicates that co-speech hand gestures are in some cases used to show the attitude of the speakers toward what they are saying.

Keywords: Multimodal Parliament Corpus, Annotation, Pronominal reference

1. Introduction

Political discourse has persuasion as its main goal (Poggi and Vincze, 2009) and the identification of the referents of pronouns in it is of great importance, e.g. (Alavidze, 2016). Face-to-face communication is embodied and reference to objects is multimodal, that is, it is made not only by means of words but also by gestures. It is therefore important to analyse how politicians use both spoken language and gestures in communication, and more specifically in this study for referring to different groups.

In this paper, we present a multimodal corpus consisting of video- and audio-recorded opening debates from the Danish parliament *Folketinget*¹. The opening debates take place on the first Thursday after the opening of a parliament year and comprise discussions of the government programme and other political issues. The speeches during the opening debates are more spontaneous than those in other parliament meetings and are therefore particularly interesting to analyse multimodally. The multimodal corpus of opening debates from 2013 to 2024 has been collected and is being annotated.

In the paper, we also present a study of how speakers indicate various groups of referents not only by means of speech, in this case the first-person plural personal pronoun *vi* (we), but also through the co-occurring hand gestures, and par-

ticularly deictic (pointing) hand gestures.

The main hypothesis that we want to investigate is whether the deictic gestures co-occurring with the pronoun *vi* are different, as we expect, that is point towards different directions, depending on the type of referent, which comprise for example the government, the parliament, the whole country, or other entities such as the European Community, or the whole world.

The data in the study are a subpart of the opening debates' corpus and comprise the speeches of two Danish prime ministers in six opening debates from the years 2013, 2014, 2021, 2022, 2023, and 2024. The speeches consist of their answers to questions posed by politicians from the opposition, who address the prime minister's opening speech, held on the preceding Tuesday, and other political issues of their interest.

The uses of first-person plural pronouns in political speeches from different countries have been addressed in several studies (Zupnik, 1994; Proctor et al., 2011; Rehbein et al., 2021). The uses of *vi* (we) in Danish parliament debates is even more relevant because, according to political scientists, Danish governments follow various strategies to create consensus with as many parties as possible in order to maintain power, even when they have a minority government. Therefore, they often invite the parliament to join the government in supporting their policies. One of the consequences of this consensus-based behaviour is that many Dan-

¹<https://www.ft.dk/>

ish bills are passed with unanimous votes (Green-Pedersen and Thomsen, 2005).

The paper is organised as follows. In the next section, we present background studies, and in section 3 we describe the multimodal corpus we have collected and the annotations we are adding to it. In section 4, we account for the annotations coded for the present investigation of multimodal reference and, in sections 5 and 6 we present and discuss the findings in the annotated data. Finally, section 7 contains the conclusions, the limitations of the study, and possible future work.

2. Background Literature

Several researchers investigate the uses of first-person plural pronouns and distinguish the following types: *exclusive*, *restricted inclusive*, and *inclusive*. They suggest that these different uses in political discourse can help identify political groups, for example, coalition versus opposition, or different standpoints such as nationalism and populism. For example, Zupnik (1994) describes a study of the persuasive function of the pronoun *we* in the transcription of a televised political debate. She concludes that the pronoun marks the discourse space and indicates the role changes of the participants during the debates.

Allen (2007) investigate the occurrences of first- and third-person plural pronouns in six speeches of two Australian politicians. They find that the two politicians use the two pronominal types to construct and indicate their own identities, those of their party, and those of their opponents.

Proctor et al. (2011) study the occurrences of *we* in American political debates and interviews with several politicians. Their study shows that the politicians use *we* to induce nationalistic emotions and achieve their career goals. Proctor et al. (2011) also find differences in the way politicians use the pronoun in debates and in interviews.

The occurrences of first-person singular and plural pronouns in speeches by Romania's Prime Minister Adrian Năstase and the United Kingdom's Prime Minister Tony Blair have been investigated by Stănculete (2019). This study shows that the use of these pronouns reflects the different views on democracy of the two politicians and their understanding of their own role in the democratic process. According to the data, Adrian Năstase uses the pronoun *I* more frequently than Tony Blair, posing himself at the centre of the political scene. In contrast, Tony Blair often uses the pronoun *we*, and this is interpreted by the authors as an indication that Blair understands the importance of collaboration in a democracy.

The referents of the pronoun *we* in the transcriptions of 36 parliamentary debates of the German

Bundestag from 2017 to 2021 have been manually annotated by Rehbein et al. (2021). In the annotations, the authors distinguish the following referent groups: a) members of a committee, b) the country (Germany), c) generic use (one), d) the government, e) the parliament, e) members of a specific party, f) groups of people defined by social variables such as age, profession, and religion, g) groups of individuals, and h) geo-political groups on supranational level. The annotations are then used to train and test a classifier to automatically identify the referents of the pronoun.

Reference to persons and objects in face-to-face communication is achieved not only by means of words but also via pointing gestures. The shape and function of pointing gestures in different cultures have been extensively studied, e.g. (Kita, 2003; Kendon, 2004; Enfield et al., 2007).

Pizzuto and Capobianco (2008) analyse the uses of first- and second-person pronouns and co-occurring hand gestures as devices to identify interlocutors in both vocal and sign languages.

Adetunji (2006) investigate the use of deictic gestures accompanying first- and second-person pronouns in two political speeches and show that politicians use multimodal reference to associate with positive actions or dissociate from controversial actions, which they or their collaborators have performed.

Poggi and Vincze (2009) analyse the persuasive power of hand gestures and gaze in fragments of video- and audio-recorded electoral debates. The behaviours of a French and two Italian politicians were investigated. According to the authors, persuasion is achieved through Aristotle's three strategies, *logos*, *ethos* and *pathos* and the three politicians use gestures and gaze in agreement with either the persuasive structure of the specific discourse fragment or their general political strategy.

Finally, the deictic gestures co-occurring with second-person pronouns in televised British and American political debates have been analysed in (Navarretta and Paggio, 2013). The investigation shows that the politicians who are highest rated by the audience are also those who succeed in indicating more clearly their different interlocutors multimodally.

In the present study, we combine research aimed at identifying the referents of first-person plural pronouns in political discourse (Rehbein et al., 2021) with multimodal studies of the use of deictic gestures, which co-occur with personal pronouns (Pizzuto and Capobianco, 2008; Adetunji, 2006).

3. A Multimodal Corpus of Opening Debates

The audio- and video-recordings of the Danish political speeches from the past 25 years are available on the Danish Parliament's television website ².

We have collected the speeches from the opening debates of the parliament in the period from 2013 to 2024³.

The Hansards of the parliament debates are available from the Danish Parliament's website⁴, and some of them are also in the Danish part of the ParlaMint corpora (Erjavec, 2025b,a; Erjavec et al., 2024), ParlaMint-DK, which covers the parliament speeches from October 2014 to June 2022. Moreover, the ParlaMint corpora contain useful metadata and linguistic annotations that could be added to the multimodal version of the speeches.

The Danish hansards are produced by the *Office of the Parliament Hansard*. They have corrected the transcriptions of the speeches for factual and minor speech errors. They have also changed them so that they resemble written language. Therefore, punctuation marks have been added to the transcripts, while pauses, speech marks, hesitations, and self-corrections have been excluded. These changes facilitate reading the resulting hansards, but are problematic for studies of spoken and multimodal communication. Therefore, in our corpus, we want to use the verbatim transcriptions of the videos. We are using Whisper OpenAI ⁵ to produce these transcriptions. We are then manually correcting them also using the transcriptions in ParlaMint-DK. The WER rate is around 34%. At the moment, we have only corrected the transcriptions used in this study. Another strategy to align the Danish Hansards and the corresponding audio-files has been proposed in (Kirkedal et al., 2020) excluding the transcriptions of the debates, which could not be automatically aligned with speech. This is the case for all the speeches we are interested in, since they deviate mostly from the hansards and contain many dialogue markers, laughs, slips of the tongue, and self-corrections⁶.

The data used in this study are the answers of two first ministers to questions from opposition par-

ties and comprise 219 speeches⁷, which we have multimodally annotated (see section 4). The transcriptions and annotations presented in this study will also be available through the CLARIN-DK repository.

The two prime ministers whose speeches we analyse are Helle Thorning-Schmidt and Mette Frederiksen, both social democrats. Helle Thorning-Schmidt was the first Danish female prime minister (2011-2015) and she was the head of two governments. The Helle Thorning-Schmidt I government (2011-2013) was a coalition led by the Social Democratic party (Social Demokratiet), the largest Danish party, and comprised the Danish Social Liberal party (Radikale Venstre), a centre-right party, and the Socialist People's party (Socialistik Folkeparti), a left wing party. The Helle Thorning-Schmidt II government (2013-14) consisted only of the Social Democratic party and the Danish Social Liberal party and was formed after the Socialist People's party decided to leave the government.

Mette Frederiksen followed Helle Thorning-Schmidt as the leader of the Social Democratic party in 2015, and she has been the prime minister since 2019. The Mette Frederiksen I government (2019-22) consisted only of social democrats, while the Mette Frederiksen II government (2022-) is led by the social democrats and comprises the Liberal party (Venstre), a party that earlier was leading right-wing governments, and the Moderates (Moderaterne), a liberal novel party started by the former leader of the Liberal Party before the 2022 elections. The elections to the Parliament have been just announced and will take place on March 24, 2026.

The duration of the annotated video files (the 219 answers to questions of the two prime ministers) in this study is approximately 3 and half hours.

The video and audio files, as well as the verbatim transcriptions, have been imported into the multimodal annotation tool, ELAN (Wittenburg et al., 2006)⁸.

Table 1 shows the year of the opening debates, the speaker, the number of tokens (both words and speech markers), the occurrences of the pronoun *vi*(we) in the speeches and the relative frequency of the pronoun in them. Table 1 shows that the relative frequency of the pronoun *vi* varies from speech to speech, and its frequency is higher in the speeches of Helle Thorning-Schmidt than in the speeches of Mette Frederiksen.

²The website is at <https://www.ft.dk/da/aktuelt/tv-fra-folketinget>.

³The links to the videos can be shared and the videos can be downloaded for private use. We will distribute the transcriptions of the debates and the multimodal transcriptions in the CLARIN-DK repository <https://dkclarin.dk/clarin.dk/>, the Danish part of the CLARIN infrastructure <https://clarin.e>.

⁴<ftp://oda.ft.dk>

⁵<https://openai.com/index/whisper/>

⁶The transcriptions described in (Kirkedal et al., 2020) are not more freely available. We got the relevant information from the first author of the paper.

⁷Each answer to a question by the opposition is a speech.

⁸<https://archive.mpi.nl/tla/elan>

Year	Speaker	tokens	vi	%
2013	Thorning-S	9,601	236	0.025
2014	Thorning-S	5,402	153	0.03
2021	Frederiksen	11,406	79	0.007
2022	Frederiksen	23,509	88	0.004
2023	Frederiksen	25,890	107	0.004
2024	Frederiksen	26,411	162	0.006
Total	Both	102,219	825	0.008

Table 1: The speeches' year, the speaker, the number of the tokens in the speeches, the occurrences of *vi*, and the relative frequency of the pronoun

4. Multimodal Annotations

In this section, we describe the annotations of the referents of *vi* and of the co-occurring hand gestures in the speeches.

4.1. Annotation of the Referents of *Vi*

We have manually annotated the referents of the pronoun *vi* in one ELAN track. The choice of referent categories was inspired by the classes proposed by Rehbein et al. (2021). We identified the following referent types: a) GOV: the government, b) PARTY: a specific party, c) PARL: the parliament, d) DENM: Denmark, e) EU_NATO: EU, Europe or NATO, f) WORLD: all the world, g) GEN: generic (one), f) PERS: the speaker or a specific person.

Two coders independently annotated the referents of *vi* in the speeches of the prime ministers in 2014 and 2022 in order to test inter-coder agreement. The rest of the speeches were annotated by one coder. Both annotators have worked with political debates for some years, and the guidelines they received consist in the list of possible referent groups.

The inter-coder agreement in terms of Cohen's kappa (Cohen, 1960) is 0.91. The few disagreement cases regard whether the referents are the parliament (PARL) or the whole country (DENM), and, in two cases, whether the pronoun refers to the government (GOV) or the prime minister's party (PARTY). In these cases, we have decided to use the widest referent: the country (DENM) and the government (GOV), respectively.

In table 2, we show the referents of the pronoun *vi* in the opening debates from the various years. The most frequent referent types of *vi* in the prime ministers' speeches are the parliament (PARL), the government (GOV), and the country (DENM).

4.2. Annotation of Hand Gestures Cooccurring with *Vi*

Hand gestures co-occurring with *vi* in the transcribed debates have been annotated with prede-

defined features as proposed in the MUMIN framework (Allwood et al., 2007). The ELAN scheme for the framework is available in the CLARIN-DK repository⁹. In table 3, we show the features that we used to describe hand gestures. The first seven features in the table indicate whether the gesture is made with one or two hands, whether the gesture is single or repeated, and characteristics of the gesture's shape and trajectory. These features are a simplified version of the features proposed in (McNeill, 1992). The last feature in the table describes the semiotic function of gestures. Its types are based on (Peirce, 1931) and consist of the following classes:

indexical deictic gestures are also known as *pointing* gestures. They indicate objects, locations, or persons.

indexical non-deictics gestures are always, in the case of hand gestures, *batonic* gestures, also called beats. They follow the rhythm of speech and often emphasise the words with which they co-occur.

iconic gestures show some semantic features of the objects they represent by similarity or homomorphism. For example, iconic gestures can show the direction of a movement or the shape of an object. Iconic gestures also include metaphoric gestures, such as moving the hands around one's own head to indicate thinking.

symbolic gestures, also known as *emblems*, have conventionalized form and meaning. An example is the victory gesture made extending the second and middle fingers to mimic the letter V. Symbolic gestures are culturally dependent.

Semiotic other indicate gestures that have more functions. In our data, they are either pointing, or iconic gestures, which at the same time are beats.

Two coders manually annotated the semiotic types of hand gestures in the speeches from 2024. The two coders were given the descriptions of the semiotic types as well as two examples of each type. The inter-coder agreement in terms of Cohen's kappa is 0.87. All disagreement cases regard whether a gesture is a deictic or a deictic and beat at the same time (SemioticOther). In these cases, we chose the most general class, (*SemioticOther*)¹⁰. In the videos of the debates, the speaker is not shown all the time. In some cases, the reactions of

⁹<http://hdl.handle.net/20.500.12115/43>

¹⁰The shape of the hand gestures were only annotated by a coder.

Year	GOV	PARTY	PARL	DENM	EU_NATO	WORLD	GEN	PERS
2013	109	1	64	39	1	0	19	0
2014	65	0	56	17	13	0	1	0
2021	7	0	45	18	0	0	4	5
2022	14	0	56	15	3	0	0	0
2023	31	2	42	30	0	0	0	0
2024	43	19	50	33	4	3	8	0
Total	269	22	288	152	21	3	32	5

Table 2: The referents of *vi* in the analysed speeches

Attribute	Value
Handedness	BothHands, Symmetric, BothHands, Asymmetric, RightSingleHand, LeftSingleHand
HandRepetition	Single, Repeated
Palm	Open, Closed, Other
Palm Direction	Up, Down, Side, Backward, Forward, Change
Trajectory RightHand	Forward, Backward, Side, Up, Down, Complex
Trajectory LeftHand	Forward, Backward, Side, Up, Down, Complex
SemioticType	IndexicalDeictic, IndexicalNonDeictic, Iconic, Symbolic, SemioticOther

Table 3: Features describing the shape and function of hand gestures

some of the members of the parliament are instead recorded. Moreover, not all hand gestures made by the speaker are clearly visible in the videos. In these cases, the gestures have not been annotated. In table 4, we give an overview of the semiotic types of the hand gestures identified in the opening debates. We have not included symbolic gestures in the table, since they did not occur in the data. The last three columns of table 4 show how many gestures are single and how many are repeated¹¹, and the percentage of occurrences of *vi*, which are accompanied by gestures. Not surprisingly, the most frequent gestures co-occurring with *vi* are beats, which follow the speech rhythm and have an emphatic function. However, we also identified many deictic gestures and multifunctional batonic and deictic gestures. Table 4 also shows that the frequency of gestures co-occurring with the pronoun *vi* varies from speech to speech. Moreover, it seems that Mette Frederiksen is using multimodal refer-

¹¹We have not annotated the times each gesture was repeated.

ence more frequently than Helle Thorning-Schmidt. However, since the videos are not always showing the speaker, and we have annotated fewer speeches of Helle Thorning-Schmidt than of Mette Frederiksen, this observation should be verified by annotating and comparing more data.

In figure 1, a screenshot of the ELAN tool with the opening debate from 2013 is shown.

5. Multimodal Reference of *vi* in the Opening Debates

In this section, we analyse the gestures co-occurring with the pronoun *vi* and provide some examples of their occurrences, focusing especially on deictic gestures in order to test our hypothesis that deictics are pointing in different directions depending on the referent type of the pronoun *vi*. In cases of a sufficient number of annotations of a class, we also control that the direction of the hand gestures related to a certain type of referent occurs significantly often, applying Pearson's chi-square

Year	Deictic	Beat	Iconic	Other	Total	Single	Repeated	%
2013	15	19	2	5	41	33	8	0.174
2014	3	12	1	1	17	9	8	0.105
2021	7	22	6	0	35	22	13	0.443
2022	8	23	4	4	39	33	6	0.443
2023	5	17	1	4	27	19	8	0.252
2024	15	33	8	17	73	50	23	0.45
Total	53	126	22	31	232	165	66	0.28

Table 4: The types of hand gesture co-occurring with *vi* in the speeches from each year

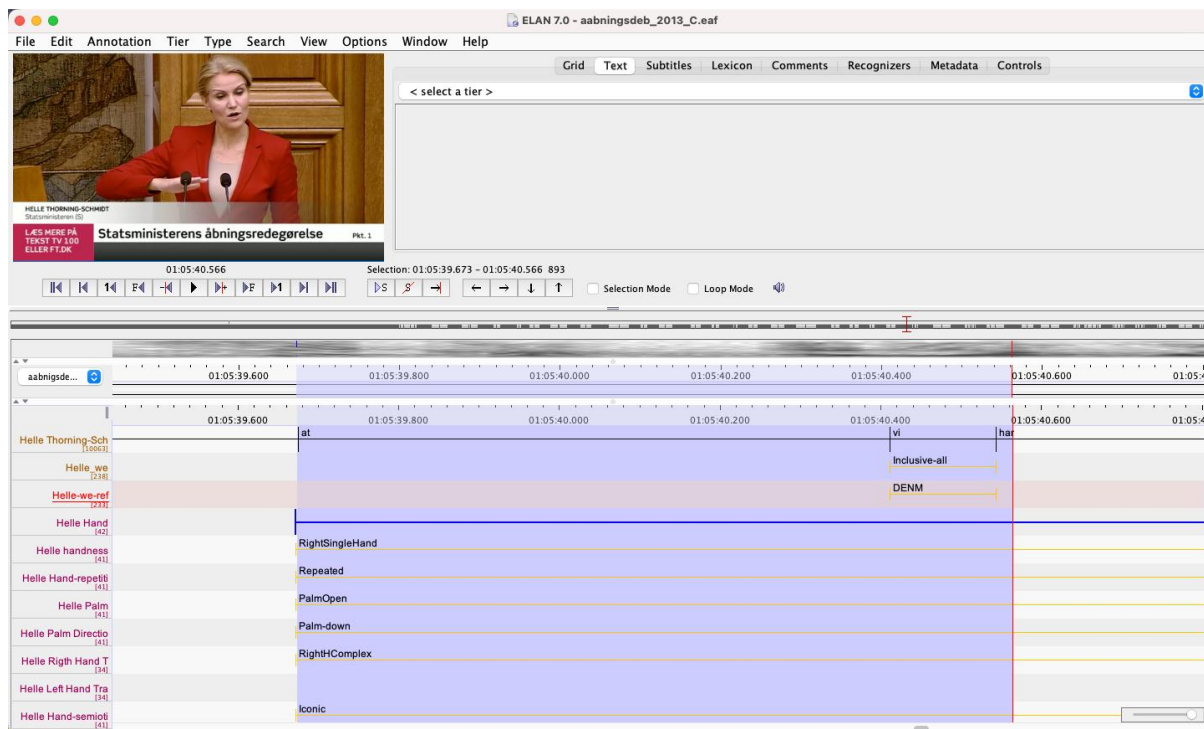


Figure 1: Screenshot from the ELAN tool with the annotation of the opening debate in 2013.

test.

Table 5 shows the distribution of the semantic types with respect to the various referent types. Not surprisingly, most deictic gestures co-occur with the pronoun *vi* when it refers to the most frequent referent types, the parliament, the government, and the country (Denmark). When *vi* referred to some groups, e.g. PERS and WORLD, it is not accompanied by deictic gestures in the data.

When the two prime ministers refer to the government or their own party by means of speech and gestures, they point to themselves while saying *vi* in 20 out of 21 cases¹² ($\chi^2 - \text{square} = 10.8068$ and $p < 0.01$).

An example of this multimodal reference is in figure 2, which shows Helle Thorning-Schmidt pointing to herself with two hands while saying *vi* and referring to the government. A second example



Figure 2: Reference to the government by Helle Thorning-Schmidt in a screenshot from the opening debate in 2013.

of multimodal reference is in figure 3 that shows Mette Frederiksen while she points to herself with her right hand and says *vi* referring to her party. In three cases, in which the referred party is not that

¹²We have included deictics which are at the same time batonic.

Semantic Type	GOV	PARL	DENM	EU	PARTY	GEN	WORLD	Total
Deictic	15	20	10	1	4	3	0	53
Beat	30	55	25	3	5	8	0	126
Iconic	5	12	4	1	0	0	0	22
Other	7	10	11	2	1	0	0	31
Total	57	97	50	7	10	11	0	232

Table 5: The semiotic types co-occurring with the various referent types of *vi*



Figure 3: Mette Frederiksen refers to her own party saying *vi* and touching herself. Screenshot from the opening debate in 2023.

of the prime ministers, they point with one hand and with their body towards a specific location in the room, probably indicating the area where the members of that party are sitting.

When the two politicians refer multimodally to the parliament, they point to the audience while saying *vi* in 24 out of 26 cases ($chi - square = 11.3369$ and $p < 0.001$).

Examples of the multimodal reference to the parliament by Helle Thorning-Schmidt and Mette Frederiksen are shown in figure 4 and 5, respectively. In the first figure, a screenshot from the 2014 opening debate, Helle Thorning-Schmidt refers to the whole parliament saying *vi* and pointing to the audience (the parliament members) with her left hand. In



Figure 4: Helle Thorning-Schmidt refers to the whole parliament saying *vi* and pointing to the parliament members in front of her with her left hand. Screenshot from the opening debate in 2014.

figure 5, a screenshot from the opening debate in 2022, Mette Frederiksen refers to the parliament saying *vi* and pointing to the parliament members in front of her with both hands. When the two politi-



Figure 5: Mette Frederiksen refers to the parliament saying *vi* and pointing to the audience (the parliament members) with both hands. Screenshot from the opening debate in 2022.

cians refer to the country, they point to the audience with two hands in 7 out of 10 cases. An example of this multimodal reference is shown in figure 6, a screenshot from the opening debate in 2021. In



Figure 6: Mette Frederiksen refers multimodally to the country saying *vi* and pointing forward with both open arms. Screenshot from the opening debate in 2021.

three cases, however, Mette Frederiksen points to herself when she refers to Denmark, indicating that she is particularly involved in what she is addressing and thus including herself as representative of the country. These occurrences, as well

as other cases of hand gestures that co-occur with the first-person plural pronouns, mainly indicate the attitude of the speaker toward what she is saying. For example, Mette Frederiksen touches her chest several times with both hands when she speaks about other issues that are particularly important to her. This "touching-the-heart" gesture often occur when she speaks about what her party or government has accomplished or is planning to do, but in a couple of cases it also co-occurs when she refers to the parliament, or to the country, expressing hope that the parliament or the country together with her will do what she is proposing. An example of the "touching-the-heart-gesture" is in figure 7. Mette Frederiksen touches her chest with one hand over the other while saying *vi ønsker at sætte kommunerne fri* (we (the S party) wish to set the municipalities free). The iconic gestures in the



Figure 7: Mette Frederiksen touches her chest while saying "we (the S party) wish to set the municipalities free". Screenshot from the opening debate in 2024.

data co-occur with a larger part of the utterances, which contain the pronoun *vi*. These gestures illustrate in different ways what the speaker is talking about. For example, Mette Frederiksen extends her right arm while saying (*de udfordringer*) *vi ser ind i på den lange bane* ((the challenges) we will face in the long term) (see figure 8). Both single-hand and both-hands gestures were present in the recorded videos of the two speakers. We did not find any difference in the two prime ministers' use of both-hand, left- and right-hand gestures.

6. Discussion

The analysis of multimodal reference in our annotated data confirm our hypothesis that different pointing gestures co-occur with the first-person plural pronoun *vi* depending on whether the speaker refers to her government or party and when she refers to the parliament or the country. In the former cases, the prime ministers point or touch themselves, while in the latter ones, they often point toward the audience. In a few cases, however, the de-



Figure 8: Mette Frederiksen illustrates "on the long term" in the utterance "the challenges we will face on the long term" by extending her right arm. Screenshot from the opening debate in 2024.

ictic hand gestures show the attitude of the speaker toward what she is saying, and more specifically that the subject they are addressing is particularly important to them.

The pointing gestures co-occurring with *vi* can help to identify the referents of the pronoun when they are present. However, the most common hand gestures co-occurring with the pronoun are beats, which emphasise the words they co-occur with. We have also found a number of iconic gestures, which co-occur with larger fragments of speech in which the pronoun *vi* appears. These gestures illustrate the whole fragment and not specifically the pronoun.

We have also found that some of the deictic gestures, which co-occurred with the first-person plural pronouns, show the attitude of the speaker towards what she is saying. This is especially the case for Mette Frederiksen.

When we analysed the gestures related to the referent annotated as *Party*, we found that this category is too vague since it covers both the party of the speaker and other parties. We had to disambiguate the two types of referent when we analysed the hand gestures co-occurring with *vi*. In the future, we will split the group *Party* into two categories: *Own-Party* and *Other-Party*.

In our study, we have only investigated multimodal reference by speech and hand gestures. However, other types of gestures such as body posture, head movements, and eye gaze can also be deictic. Furthermore, all types of gestures and especially facial expressions are important when studying the attitudes of people toward what they say or what is said by other people, and therefore they should be included when studying attitudes. In fact, the body posture and facial expression of Mette Frederiksen in figure 7 also contribute to show her attitude towards what she is saying. We have started annotating these other gesture types

in the data.

7. Conclusion, Limitations, and Future Work

In this article, we have presented a multimodal corpus of Danish video- and audio-recorded parliamentary opening debates, which we have collected and are multimodally annotating.

We have also described a study of the referents of the first-person plural pronoun *vi* (we) in a subset of the corpus, consisting of the speeches of two prime ministers during the opening debates. These data have been multimodally annotated and inter-coder agreement experiments have been conducted to evaluate the consistency of the annotations. Our investigation of the uses of the pronoun *vi* in the speeches shows that the most frequent referents of the pronoun *vi* are the parliament, the government, and the country. The frequent reference to the parliament (an inclusive use of the first-person plural pronoun) confirms that Danish prime ministers try to obtain consensus with as many parties as possible, as proposed in (Green-Pedersen and Thomsen, 2005).

In our analysis of how the two prime ministers refer to the different referents of *vi* by means of co-speech hand gestures, we found that pointing gestures and combined pointing and batonic gestures contribute to the identification of the referents. Thus, the results of our study confirm, for the most frequent referent groups, our hypothesis that the pointing gestures co-occurring with the first-person plural pronoun have different directions depending on the referent type. In fact, when the prime ministers refer to the government or their own party (and in some cases, the two coincided), they point to or touch their chest. When they refer to the parliament or the country, they point toward the parliament members who sit in front of them in most cases, but there are always exceptions. Some of the hand gestures co-occurring with the pronoun *vi* also indicate the attitude of the speaker towards what she is saying. In particular, the "touching-the-heart" gesture shows that the subject addressed is important to her.

Moreover, only part of the occurrences of the pronoun are accompanied by co-speech gestures, and only part of these co-occurring gestures are deictic. Therefore, multimodal reference can only contribute to the identification of the pronominal referents in these cases. However, multimodal reference is important because it is used when the two prime ministers want to identify clearly and/or emphasise the referents of the pronouns. In some cases, they focus on the achievements or plans of their government or party, in other cases, they look for consensus in the parliament or speak on behalf

of the whole nation. These uses are also in line with the proposal in (Poggi and Vincze, 2009) that hand gestures are one of the persuasive means utilized by politicians.

Concluding, the present investigation of multimodal reference can only in few cases contribute to pronominal reference resolution, but it contributes to understanding multimodal communication, and not only in political speeches, and its findings can also be useful to generation systems of multimodal behaviours.

There are many limitations to this study. First of all, only the speeches of two prime ministers in six opening debates have been analysed. Secondly, only hand gestures have been considered. However, humans use both eye gaze, body posture, and head movements to point to objects. These gestures will also be considered in future annotations of the corpus.

Another limitation is that we only addressed the multimodal reference in relation to the first-person pronoun *vi*, since its use is particularly interesting in political discourse. However, multimodal reference regards all referring expressions, and also the use of other pronouns is interesting in political speeches as described in section 2.

In the future, we will finish the time-aligning work on all the speeches in the opening debates, and we will annotate all types of visible co-speech gestures in the videos pursuing to produce so much of the work as possible semi-automatically.

8. Acknowledgements

We thank our former colleague Dorte Haltrup Hansen who was the second coder in the inter-coder agreement experiments.

9. Bibliographical References

- Akinbiyi Adetunji. 2006. Inclusion and exclusion in political discourse: Deixis in Olusegun Obasanjo's speeches. *Journal of Language and Linguistics*, 5(2):177–191.
- Maia Alavidze. 2016. The use of pronouns in political discourse. *International Journal of Arts & Sciences*, 9(4):349.
- Wendy Allen. 2007. Australian political discourse: Pronominal choice in campaign speeches. *Selected Papers from the 2006 Conference of the Australian Linguistic Society*.
- Jens Allwood, Loredana Cerrato, Kristiina Jokinen, Costanza Navarretta, and Patrizia Paggio. 2007. The mumIn coding scheme for the annotation

- of feedback, turn management and sequencing. *Multimodal Corpora for Modelling Human Multimodal Behaviour. Special Issue of the International Journal of Language Resources and Evaluation*, 41(3–4):273–287.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46.
- Nick J Enfield, Sotaro Kita, and Jan Peter De Ruiter. 2007. Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, 39(10):1722–1741.
- Tomas Erjavec, Matyáš Kopp, Nikola Ljubešić, Taja Kuzman, Paul Rayson, Petya Osenova, Maciej Ogrodniczuk, Çağrı Çöltekin, Danijel Koržinek, Katja Meden, Jure Skubic, Peter Rupnik, and et al. 2024. [Parlamint ii: advancing comparable parliamentary corpora across europe](#). *Language Resources and Evaluation*.
- Tomaž et al. Erjavec. 2025a. [Linguistically annotated multilingual comparable corpora of parliamentary debates ParlaMint.ana 5.0](#). Slovenian language resource repository CLARIN.SI.
- Tomaž et al. Erjavec. 2025b. [Multilingual comparable corpora of parliamentary debates ParlaMint 5.0](#). Slovenian language resource repository CLARIN.SI.
- Christoffer Green-Pedersen and Lisbeth Hoffmann Thomsen. 2005. Bloc politics vs. broad cooperation? the functioning of danish minority parliamentarism. *The Journal of Legislative Studies*, 11(2):153–169.
- Adam Kendon. 2004. *Gesture: Visible action as utterance*. Cambridge University Press.
- Andreas Søeborg Kirkedal, Marija Stepanovic, and Barbara Plank. 2020. Ft speech: Danish parliament speech corpus. In *INTERSPEECH 2020*, pages 442–446. International Speech Communication Association (ISCA).
- Sotaro Kita. 2003. *Pointing: Where language, culture, and cognition meet*. Psychology Press.
- David McNeill. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.
- Costanza Navarretta and Patrizia Paggio. 2013. Multimodal behaviour and interlocutor identification in political debates. In *International Workshop on Political Speech*, pages 99–113. Springer.
- Charles Sanders Peirce. 1931. *Elements of Logic. Collected Papers of Charles Sanders Peirce. Volume Two*, edited by Hartshorne, C. & Weiss, P. Harvard University Press.
- Elena Antinoro Pizzuto and Micaela Capobianco. 2008. Is pointing “just” pointing?: Unraveling the complexity of indexes in spoken and signed discourse. *Gesture*, 8(1):82–103.
- Isabella Poggi and Laura Vincze. 2009. *Gesture, Gaze and Persuasive Strategies in Political Discourse*, pages 73–92. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Katarzyna Proctor, I Lily, and Wen Su. 2011. The 1st person plural in political discourse—american politicians in interviews and in a debate. *Journal of pragmatics*, 43(13):3251–3266.
- Ines Rehbein, Josef Ruppenhofer, and Julian Bernauer. 2021. Who is we? disambiguating the referents of first person plural pronouns in parliamentary debates. In *Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021)*, pages 147–158.
- Iliana-Mihaela Stănculete. 2019. I vs. we. first person personal pronouns in political speeches. *East-West Cultural Passage*, 19(1):121–134.
- Peter Wittenburg, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. [ELAN: a professional framework for multimodality research](#). In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC’06)*, Genoa, Italy. European Language Resources Association (ELRA).
- Yael-Janette Zupnik. 1994. A pragmatic analysis of the use of person deixis in political discourse. *Journal of pragmatics*, 21(4):339–383.