Political Discourse Analysis : A Case Study of 2014 Andhra Pradesh State Assembly Election of Interpersonal Speech Choices

by

Lalitha Kameswari VA, Radhika Mamidi

in

The Pacific Asia Conference on Language, Information and Computation
(*PACLIC-2018*)

Hong Kong Polytechnic University Jockey Club Auditorium, Hong Kong

Political Discourse Analysis : A Case Study of 2014 Andhra Pradesh State Assembly Election of Interpersonal Speech Choices

Lalitha Kameswari
Language Technologies Research Centre
KCIS, IIIT Hyderabad
Telangana, India
v.a.lalitha@research.iiit.ac.in

Radhika Mamidi
Language Technologies Research Centre
KCIS, IIIT Hyderabad
Telangana, India
radhika.mamidi@iiit.ac.in

Abstract

Since the beginning of the 20th century, people have started studying the correlation between language and culture. They observed how language is used in discourses to establish power relations in the society. Eggins (2004) found that the link between language and the choice made by the speaker in the exchange enable us to see speakers making meaning about interpersonal: the extent of their intimacy, their level of familiarity with each other and their attitudes and judgments. In political speeches, the speaker uses language to persuade the voters, influence their perceptions and build a positive interpersonal identity. Keeping in mind this result-oriented attempt by the speakers, Van Dijk and others (1997) describe discourse as political when it has a direct functional role as a form of political action in the political process. In this paper, we will look at four such speeches given by notable politicians from both winning and losing parties during the campaign of Andhra Pradesh State Assembly elections of 2014 and closely observe the linguistic choices made at the lexical and semantic levels. By a contrastive analysis of the speeches of winning and losing parties, we can identify the linguistic features which contribute to the outcome.

1 Introduction

Communication is one of the main pillars of politics. Language is used in various forms at different levels of communication such as speeches, arguments, press releases, pamphlets, advertisements, and manifestos. Charteris-Black (2011) states that "within all types of political systems, from autocratic, through oligarchic to democratic; leaders have relied on the spoken word to convince others of the benefits that arise from their leadership."

Language is a double edged sword in political discourses. It can effectively avoid conflicts and also stir controversies within no time. The way politicians frame their expressions, to a large extent determines who they are and whether or not they will succeed in their profession. So keeping in mind the end goal of appealing to the voters, political leaders construct their discourses strategically in the best way possible. Van Dijk's sociocognitive approach talks about one such strategy - the ideological polarisation between 'Us' and 'Them' where the speaker emphasises 'our good' and 'their bad', simultaneously de-emphasising 'our bad' and 'their good' (Van Dijk, 2006a). In addition to this, a speaker often has to establish multiple identities among various socio-economic groups using language as a medium to build these identities. Appropriate use of pronouns and kinship terms include or distance the speaker, audience, and other politicians. Therefore political discourses hold a great importance in the field of Critical Discourse Analysis.

When it comes to politics at the state level, parties often use the local/regional language for reaching out to a maximum number of voters, as well as creating a shared cultural and ethnic identity. This paper attempts to analyze the campaign speeches of one such state assembly elections of Andhra Pradesh in 2014. Telugu, being the official language of Andhra Pradesh and Telangana, is the language

used in these speeches. We have selected four campaign speeches given by different politicians belonging to winning and losing parties as our political discourse dataset. As there is no written record for the speeches from any reliable source, we manually transcribed them. After re-checking, the transcriptions were cleaned, classified and sorted to get the sentence count, word count, and frequency of each word.

2 Related Work

Critical Discourse Analysis (CDA) started emerging as a research field in the late 20th century. Different models by Fairclough, Van Dijk, Reisigl, and Wodak have contributed a lot to this field.

2.1 Fairclough's model of CDA

Fairclough (2001) first made a distinction between text and discourse. He claims that text is a form of product, while discourse is the larger social interaction process and text is a part of it. He developed a framework consisting of three dimensions of discourse notions, providing a three-dimensional method for discourse analysis. Fairclough (2013) explains that discourse can be understood as "(i) a language text, spoken or written, (ii) discourse practice (text production and text interpretation), (iii) sociocultural practice."

2.2 Van Dijk's contributions to CDA

Van Dijk (2001) sought to establish a relation between discourse and political ideologies in terms of the structures of political discourse. For instance, he observed the use of biased lexical items, syntactic structures such as actives and passives, pronouns such as 'us'and 'them', metaphors or topoi, arguments, implications, and many other properties of discourse. Van Dijk (2006b) looked at discourse as a medium through which political ideologies are acquired, expressed and propagated.

2.3 Discourse Historical Approach (DHA)

Reisigl and Wodak (2017) built the DHA based on the notions of critique, ideology, and power - which form the main focus of CDA. For DHA, language is a means to gain and maintain power by the use powerful people make of it. Hence DHA critically analyses the use of language by powerful people and the ways in which linguistic forms are used in various expressions and manipulations of power.

They aim to answer five basic questions about a given discourse - "(i) How are persons, objects, phenomena/events, processes and actions named and referred to linguistically? (ii) What characteristics, qualities, and features are attributed to social actors, objects, phenomena/events, and processes? (iii) What arguments are employed in the discourse in question? (iv) From what perspective are these nominations, attributions, and arguments expressed? (v) Are the respective utterances articulated overtly; are they intensified or mitigated?". We shall be answering similar questions while analyzing our data in this study.

2.4 Other recent studies

Drawing mainly from the existing methods and frameworks in CDA, there have been a lot of applications recently. Fatin et al. (2017) reviewed the theoretical and practical aspects of CDA in the analysis of language use in social context. Steffens and Haslam (2013) analyzed the official election campaign speeches of successful and unsuccessful Prime Ministerial candidates in Australian Federal elections and observed the differences in their choice of personal and collective pronouns. Jin and Lu (2013) conducted a contrastive study of speeches given by Mr. Obama and Mr. McCain using aspects of Systemic Functional Grammar to explain why Obama performed relatively better. Al-Sharoufi (2006) used methods of CDA to analyze the strategies used in Arabic newspaper editorials which convey ideological messages to the readers.

2.5 CDA for discourse in Telugu

Telugu is a Dravidian language native to India. There are about 75 million native Telugu speakers but there are very few resources which make it difficult for corpus creation and analysis. Very little work has been done so far in Telugu. Suryakanthi and Sharma (2015) worked on Discourse Translation from English to Telugu. To our knowledge, there hasn't been a previous study on political discourses in Telugu from a linguistic perspective.

3 Dataset

3.1 Background and data selection

The previous elections to the Andhra Pradesh State Legislative assembly were held in 2014. They were of special interest to the parties as well as to the people. They were the first elections after the Telangana bill for the formation of a separate state was assented by the President of India in March 2014. Due to this reason, parties and leaders put in vigorous efforts to campaign and reach out to the voters. Hence we have chosen this specific election campaign to gather political discourses for our data.

Out of the 294 total seats, Telugu Desam Party (TDP) and its ally Bharatiya Janata Party (BJP) emerged as the winning coalition securing 126 seats. Youth, Labour and Farmer Congress Party (YSRCP) lost to them, securing only 70 seats [1]. We have considered two speakers from the winning alliance - Mr. Nara Chandrababu Naidu, the leader of TDP and Mr. Pawan Kalyan, the founder of Janasena party who was a key supporter of the TDP-BJP alliance. Similarly, we selected two speakers from the losing party - Mr. Y.S. Jaganmohan Reddy, founder of the nascent party YSRCP and Ms. Y.S. Sharmila, sister of Mr. Jaganmohan Reddy. Both of them were the children of former Chief Minister Dr. Y.S. Rajasekhara Reddy. Of these four speakers, Mr. Chandrababu Naidu and Mr. Jaganmohan Reddy were the Chief Ministerial candidates.

Once we had finalized these four speakers, we collected one campaign speech of each speaker. We selected the speech such that - (i) It was given in an open area in front of a huge gathering of people from all sections of the society. (ii) The words of the speaker were clearly audible to minimize errors in transcription. (iii) The duration was sufficiently long to facilitate analysis. (iv) It had a good number of views online and was telecasted in all major regional channels.

3.2 Collection of data

We had the video recordings of the selected speeches available on YouTube[2]. The total duration of all the speeches was 70 minutes. The background noise in the audio was eliminated using Audacity[3], which is an open-source digital audio editor and recording application software. We manually transcribed each speech while listening to the speaker for best accuracy. The transcribed data was then checked and compared with the original speech by a team of two native Telugu speakers to eliminate errors, if any. This is the gold data we will be referring to in our analysis.

4 Methodology

The gold data was segmented into individual sentences. Parallelly, it was broken down into the word level and arranged lexicographically along with the frequency of occurrence of each word in the speech.

4.1 Classification of sentences

After segmentation, our dataset comprised of 613 sentences. To observe specific patterns of linguistic choices made by the speakers, sentences were classified into 8 categories as shown in Table 1. If a particular sentence glorifies the speaker, his/her party or any of its members and supporters, it was labelled as Self-Positive. When a speaker attempted to criticise other leaders and parties, that sentence was labelled as Other-negative. For example, *"nEnu abhivruddhi chEsAnu, AdAyam penchAnu"* which translates to *"I have worked for welfare and prosperity and have contributed to the increased income"* is a Self Positive sentence in Mr. Chandra Babu Naidu's speech whereas Mr. Jaganmohan Reddy said *"ivALa unna I rAjakIya nAyakulu vOTu lA kOsam seaT lA kOsam yE gaDDi ainA kUDA tinaDAniki venakADaTlEdu"* in a similar context which means *"Today's political leaders (referring to the ruling party) are not even hesitating to go to any extent just to win votes in the elections"* is classified as an Other-Negative sentence. When a speaker just gives some information, the sentence is Neutral. Questions were labelled as Interrogative. All promises made by the speakers were put into a separate category. When the speaker urges the audience to do something or commands some other political figure, the sentence is Imperative. Instead, if he/she politely makes a request or humbly shows gratitude towards the public or the

[1]Source:ElectionCommissionreport-http://eci.nic.in/eci_main/StatisticalReports/AE2014/Stat-Report-Andhra-Pradesh2014.pdf

[2]https://www.youtube.com/

[3]https://www.audacityteam.org/

	Self Positive	Other Negative	Inform-ative / Neutral	Interro-gative	Imper-ative	Promises	Local Sentiment	Request /Gratitude	Total
Chandrababu Naidu	29	20	33	13	18	25	8	1	147
Pawan Kalyan	37	38	58	20	20	0	25	3	203
Jaganmohan Reddy	15	51	74	7	8	13	0	6	174
Sharmila	22	19	15	5	10	9	4	10	94

Table 1: Sentence count of each speaker

Sentence %	Self Positive	Other Negative	Inform-ative / Neutral	Interro-gative	Imper-ative	Promises	Local Sentiment	Request /Gratitude
Chandrababu Naidu	19.73%	13.60%	22.45%	8.84%	12.25%	17.01%	5.44%	0.68%
Pawan Kalyan	18.23%	18.72%	28.57%	9.85%	9.85%	0.00%	12.32%	2.46%
Jaganmohan Reddy	8.62%	29.32%	42.53%	4.02%	4.59%	7.47%	0.00%	3.45%
Sharmila	23.40%	20.22%	15.96%	5.32%	10.64%	9.57%	4.25%	10.64%

Table 2: Percentage distribution of sentences

	Positive/ Appreci-ative	Negative/ Derogat-ory	Kinship Terms	First Person Pronouns	Second Person Pronouns	Third Person Pronous	Interr-ogative Pronouns	Nouns	Functi-onal Words	Past Verbs	Present Verbs	Future Verbs	Others	Total
Chandrababu Naidu	30	34	23	62	34	27	25	197	186	45	78	71	76	888
Pawan Kalyan	84	42	5	117	28	60	29	431	448	84	90	45	218	1681
Jaganmohan Reddy	29	53	79	102	25	54	48	674	722	61	109	119	330	2405
Shamrila	30	24	7	15	22	16	9	421	303	52	25	22	176	1122

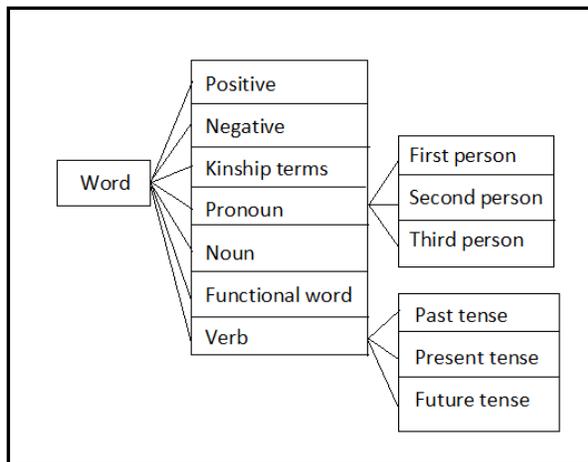Table 3: Word count of each speaker

Figure 1: Categorisation of words

previous leaders, the sentence comes under the Request / Gratitude category. Apart from these, a political discourse often aims to connect to the audience through a shared identity. One way of creating such identity is by evoking a sentiment of belongingness for the speaker among the target audience. Such sentences were labelled as Local sentiments.

Each speech was classified into these categories and the frequency of sentences in each category was recorded, as shown in Table 1. Table 2 shows the percentage of each category for each speech based on the total number of sentences. To get an average percentage of each category for comparison among the winning and losing party speeches, mean value of the corresponding percentage in the two respective speeches was calculated.

4.2 Classification of words

At the word level, our data set consisted of 6096 words. All the words in each speech were classified into 7 major lexical categories and were further subdivided as shown in Fig 1. Other words which didn't fall into any of the categories shown were not considered for further analysis. After omission of 800 such words, we were left with 5296 words in our dataset. Frequency of words in each category is shown in Table 3. From the word count and the sentence count of each speech, we obtained the average length of a sentence for each speech. Similar to the classification done for sentences, we calculated the average percentage of words in

each category for winning and losing speeches. Apart from that, we recorded the percentage of first person, second person, and third person pronouns with respect to the total number of pronouns used by the speaker. We also calculated the percentage of verbs used in each tense. From all this data, we tried to observe the similarities and differences among the speeches.

5 Observation of linguistic choices

5.1 Observations at the sentence level

All the speakers talked about the problems which the state was facing, the merits and demerits of previous governments and their plans for development. They included some rhetorical questions, touched upon the local sentiments and made some promises and requests. Table 4 shows the average percentage of each type of sentences in the winning and the losing speeches. We observed almost the same average percentage of promises in both the winning and the losing speeches. However, other differences are important to observe. The winning speeches have a slightly higher percentage of self-positive sentences whereas the losing ones have a significantly higher percentage of other-negative sentences. This leads to our inference that more focus on convincing the voters about one's own capabilities and strengths creates a good impression about the speaker, whereas constant criticism about other parties proves counterproductive.

Earlier work by Goffman (1976) states the importance of asking questions in social interactions. When it comes to politics, "the ability to question the actions and intentions of governments is a crucial part of democracy" (Pitkin, 1967). During campaign speeches, asking questions can help a speaker get the attention of the audience and make them reflect on the speech. A higher percentage of interrogative sentences in the winning speeches found in our analysis supports these studies.

While campaigning, it is important to create a shared social identity among the voters. Kroskrity (1999) defines social identity as "the linguistic construction of membership in one or more social groups or categories". For the creation of such identity, speakers tend to bring the "Us" factor by often

	Self Positive	Other Negative	Informative / Neutral	Interrogative	Imperative	Promises	Local Sentiment	Request / Gratitude
Winning	18.98%	16.16%	25.21%	9.35%	11.05%	8.50%	8.88%	1.57%
Losing	16.01%	24.77%	29.25%	4.67%	7.61%	8.52%	2.12%	7.05%

Table 4: Average percentage of sentences in each category

referring to the local area where they are speaking. For example, during his campaign speech at Pileru in Chittoor district of Andhra Pradesh, Mr. Chandrababu Naidu said *"pIlEru nAku kotta kAdu. 30 ELLu I jillA lO unnAnu 35 ELLu rAjakIyAla lO unnAnu"*, which translates to *"I'm not new to pIlEru. I've stayed here for 30 years and have been in politics for 35 years"*. This perhaps assures the voters that the speaker knows what their problems are, better than the others do. Our findings that the winning speeches have a higher percentage of local sentiments support this hypothesis.

Imperative statements and requests are two contrastive ways which speakers use to persuade the audience. *"ippuDu saraina nirNayam mIru tIskOvAli"* *(You should take the right decision now)* is an imperative statement by Mr. Chandrababu Naidu whereas *"dayachesi fAn gurtu kE mI amUlyamaina vOTu veyyAlani mA prArthana"(We request you to please vote for the 'fan' symbol)* is a request made by Ms. Sharmila. Both of them are asking the audience to vote for them but repeated requests give the audience a feeling that the speaker might be desperate for votes. The higher percentage of imperative statements in winning and requests in losing speeches is observed in our data.

Apart from this, we compared the average length of a sentence among the speeches. It was found that the number of words per sentence was more in the losing speeches whereas winning speakers kept their sentences short and simple (Table 5).

5.2 Observations at the word level

On careful observation, words carry a speaker's intention along with the literal meaning they convey. Table 6 shows the average percentage of words used by the winning and the losing speakers in each category. The difference between the positive and negative words is 1.02% in winning speeches and -0.17% in the losing speeches. This clearly shows that suc-

Speaker	Total Words	Total Sentences	Average words per Sentence
Chandrababu Naidu	888	147	6.04
Pawan Kalyan	1681	203	8.28
Jaganmohan Reddy	2405	174	13.82
Sharmila	1122	94	11.93

Table 5: Average sentence length of each speaker

cessful speakers create an overall positive feeling in the minds of voters and look at the brighter side of things whereas repeatedly highlighting some negative aspects might leave the voters frustrated and unhappy. A slightly higher percentage of kinship terms is seen in the losing speeches. This can be seen as the lack of confidence on the speaker's part where he/she tries to impress the voters just by assuming the superficial roles analogous to family members.

The winning speeches contain a relatively lesser number of nouns when compared to the losing ones. We observed the frequently used nouns in all the speeches, which are related to politics, elections, and campaigning. Mr. Pawan Kalyan used the noun *sImAndhra* 28 times in his speech, followed by most frequent nouns such as *prajalu (people)* and *congress*, which indicate that most of his speech was about how the people of Seemandhra would benefit in future and how the existing Congress party had failed in its governance. Ms. Sharmila's frequently used nouns were *Rajasekhara Reddy, YS-RCP* and *pArty*. Most of her speech was about Dr. Y.S.Rajasekhara Reddy, his achievements as a Chief Minister and how their party was a tribute to his work. This perhaps failed to offer the voters a strong promise for the future in contrast to the winning speeches.

	Positive	Negative	Kinship	Nouns	Functional Words
Winning	4.19%	3.17%	1.45%	23.91%	23.80%
Losing	2%	2.17%	1.96%	32.77%	28.51%

Table 6: Average percentage of words in each category

	First Person	Second Person	Third Person	Interrogative
Winning	46.38%	17.57%	22.38%	13.67%
Losing	38.12%	24.12%	26.67%	11.09%

Table 7: Average percentage of pronouns in each category

	Past	Present	Future
Winning	30.78%	40.65%	28.58%
Losing	36.83%	31.49%	31.70%

Table 8: Average percentage of verbs in each tense

Table 7 compares the usage of pronouns among winning and losing speeches. We find that successful speakers make more use of first person and interrogative pronouns when compared to the others. *nEnu*, which means I, is the most frequently used pronoun in both the winning speeches. This agrees with the findings of Thomas and Wareing (2004) that the use of first person singular pronoun "clearly declares who is responsible" and reflects the sharing of interests between audience and the speaker. The usage of the second person pronoun could be perceived as an increased distance between the speaker and the audience. The third person pronoun is generally used to refer to other leaders and parties, usually in the "us" and "them" comparisons.

On the basis of statistics of tenses used in verbs, the percentage of present tense verbs is significantly higher in the winning speeches (Table 8). Wang (2010) explains this - "present tense ranks with top priority since the addresses are to present the domestic and worldwide situations ranging from political, economic and cultural fields at present." The use of future tense is roughly the same in both the categories. The higher use of past tense verbs in the losing speeches indicates that they lay more emphasis either on glorifying achievements of their party in the past or criticising the other parties for their previous decisions, instead of discussing the present situations and future plans.

6 Conclusions and Future Work

We conducted a comparative study of the linguistic choices in the campaign speeches of the winning and the losing parties. The winning speeches had short and simple sentences, with more focus on their own capabilities rather than criticising others. They had more questions to get the attention of the audience and provide reasoning to their ideologies. Successful speakers included many local sentiments and made imperative statements in their speech which urged the voters to take the right decision whereas the unsuccessful speakers relied more on requesting the audience to vote for them. The difference between positive and negative vocabulary is positive in case of winning speeches and negative for losing speeches. A higher percentage of first person and interrogative pronouns was more effective than the usage of the second and the third person pronouns. The usage of present tense in verbs was observed more frequently in case of the winning speeches, past tense more in the losing speeches whereas future tense usage was nearly the same. We also observed trends among the frequently used nouns and kinship terms.

This detailed analysis will help us in identifying the key features while classifying and analyzing political speeches computationally. Our future work includes predicting the outcome of an election (or the percentage of votes) based on the kind of speeches given by a candidate.

References

Hussain Al-Sharoufi. 2006. Critical discourse analysis of political editorials in some arabic newspapers. *11th Proceedings of the Pan Pacific Association of Applied Linguistics*, pages 8–27.

Jonathan Charteris-Black. 2011. *Politicians and rhetoric: The persuasive power of metaphor*. Springer.

Suzanne Eggins. 2004. *Introduction to systemic functional linguistics*. A&C Black.

Norman Fairclough. 2001. *Language and power*. Pearson Education.

Norman Fairclough. 2013. *Critical discourse analysis: The critical study of language*. Routledge.

Nor Fatin, Abdul Jabar, and Kamariah Yunus. 2017. Critical discourse analysis (cda) as an approach to the analysis of language. 97, 01.

Erving Goffman. 1976. Replies and responses. *Language in society*, 5(3):257–313.

Pengsun Jin and Fengfeng Lu. 2013. A contrastive study of political speeches in presidential election of interpersonal meaning. *Studies in Literature and Language*, 6(3):79.

Paul V Kroskrity. 1999. Identify. *Journal of linguistic anthropology*, 9(1-2):111–114.

Hanna F Pitkin. 1967. *The concept of representation*. Univ of California Press.

Martin Reisigl and Ruth Wodak. 2017. The discourse-historical approach (dha). 01.

Niklas K Steffens and S Alexander Haslam. 2013. Power through us: Leaders use of we-referencing language predicts election victory. *PloS one*, 8(10):e77952.

T Suryakanthi and Kamlesh Sharma. 2015. Discourse translation from english to telugu. In *Proceedings of the Third International Symposium on Women in Computing and Informatics*, pages 222–227. ACM.

Linda Thomas and Shan Wareing. 2004. *Language, society and power: An introduction*. Routledge.

Teun A Van Dijk et al. 1997. What is political discourse analysis. *Belgian journal of linguistics*, 11(1):11–52.

Teun A Van Dijk. 2001. Discourse, ideology and context. *Folia linguistica*, 35(1-2):11–40.

Teun A Van Dijk. 2006a. Discourse, context and cognition. *Discourse studies*, 8(1):159–177.

Teun A Van Dijk. 2006b. Politics, ideology, and discourse.

Junling Wang. 2010. A critical discourse analysis of barack obamas speeches. *Journal of language teaching and research*, 1(3):254–261.