

WUT 2020
10th International Conference “Word, Utterance, Text: Cognitive, Pragmatic and Cultural Aspects”

**THE METHOD OF MATHEMATICAL STATISTICS IN A
PRAGMALINGUISTIC EXPERIMENT**

Maxim Ivchenko (a), Marina Laskova (b)*, Vladimir Lazarev (c), Elena Pishkova (d),
Irina Samarina (e)
*Corresponding author

(a) Southern Federal University, Rostov-on-Don, Russia, mvivchenko@sfnu.ru

(b) Southern Federal University, Rostov-on-Don, Russia, mvlaskova@sfnu.ru

(c) Southern Federal University, Rostov-on-Don, Russia, vladivlazz@rambler.ru

(d) Southern Federal University, Rostov-on-Don, Russia, pishkova@yandex.ru

(e) Southern Federal University, Rostov-on-Don, Russia, samarina@yandex.ru

Abstract

A lot of research has been carried out to identify a person's individual qualities but there have been insufficient studies into the peculiarities of the speech behaviour of people representing certain professional groups including politicians running for the presidency who may largely affect the life and well-being of their fellow citizens. Aims: Because most often politicians use scripts of their speeches carefully prepared for them by speechwriters, we decided to study spontaneous pre-election speeches of a group of US presidency candidates. Thus by means of an implicit pragmalinguistic method we aimed at diagnosing their individual personal qualities and those common for the whole professional group. Sample and Methods: We analysed the pre-election speeches of ten American politicians running for the post of US President and counted the speech signals they used which allowed us to identify their certain individual characteristics. This required the use of objective pragmalinguistic analysis, the modified content analysis and the method of mathematical statistics and the theory of probability. Results: We found that stereotyped features of a socio-professional group of politicians include objectiveness, self-assurance, rationality. The individual qualities of US presidency candidates diagnosed in our research are confirmed by the facts of their social behaviour described by different political analysts, political psychologists and biographers. Conclusion: We can use the received data for further research of the speech behaviour of the representatives of other social and professional groups. Besides, the results of the survey can be helpful for specialists in the sphere of psychological and pragmalinguistic diagnosis.

2357-1330 © 2020 Published by European Publisher.

Keywords: Pragmalinguistics, experiment, mathematical statistics, speech strategy, politicians.



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The modern linguistics largely employs statistical research methods which is caused by a number of reasons. Among them are the rapid development of linguistic subjects, such as sociolinguistics, psycholinguistics, pragmalinguistics which require statistical methods; the applied tasks connected with the automatic processing of the text; and above all, the search for new ways of enhancing the objectiveness and accuracy of the research. This article is aimed at describing the method of mathematical statistics used for achieving more accurate results in an objective pragmalinguistic experiment.

The experiment is called objective due to the absence of any kind of subjective assumptions either on the part of the examinee or the researcher. The opinion of the text sender about their own individual qualities cannot influence conducting the experiment and gaining the ultimate results, which excludes the subjectiveness on their part. The objectiveness of the researcher carrying out the experiment is gained through the use of methods of mathematical statistics and quantitative data processing. The experiment is called pragmalinguistic as we study the linguistic manifestation of the speaker's verbal activity (Oberemchenko & Degtyaryova, 2018). Moreover, the research has been conducted in the framework of implicit pragmalinguistics which studies a person's speech with an individual set of social and psychological characteristics as well as with an individual life experience and stereotypes of behavior (including verbal behavior). Implicit pragmalinguistics allows to diagnose a speaker's individual features by their speech. It deals with shades of meanings presented as a part of unconscious verbal behavior and in the form of speech habits (Pishkova, 2016).

The specific character and choice of the objective pragmalinguistic experiment as a method of research is determined by its object, purpose and material. The object of the current study is the verbal behavior of the communicants, namely politicians running for the US Presidency. The ultimate purpose of the research is diagnosing individual characteristics of the communicants in question. The material of the research is their discourse in the form of the texts. The texts are based on the scripts of pre-election debates of ten candidates running for the US Presidency (G. Bush Sr., J. Carter, W. Clinton, M. Dukakis, J. Kennedy, W. Mondale, R. Nickson, R. Perot, R. Reagan, G. Ford) in the speech genre of pre-election discourse. The choice of the pre-election discourse is determined by the verbal behavior of political leaders who talk on the spot and spontaneously, without due preparation. In such situations they automatically choose the linguistic means that are more habitual for them.

The objective pragmalinguistic experiment was carried out according to a certain scheme. To begin with, we specified the list of speech signals. Then we found texts of pre-election speeches of ten US presidential candidates. It was of paramount importance for us to follow the speech behavior of the examinees referring to the same gender (men), being of the same origin (Americans), being of the same mature age (46-69), connected with the same activity (politicians), pursuing the same objective (to become President of the country).

Next, the texts were divided into small syntactical groups (SSG) which are the unit of the current pragmalinguistic research. Their number for each candidate was 500. Then the matrix tables were made to be filled in with speech signals used by the candidates in their rhetoric

After that, we processed the data received, namely, we calculated the frequency of occurrence of implicit intentions in percentage terms using the following formula:

$$\mathbf{F\ format} = (\mathbf{E\ format\ SSG} \div \mathbf{E\ text\ SSG}) \times 100\%$$

where **F format** is the frequency of the format occurrence in percentage terms; **E** is the sum of SSG.

Let us consider the results of using the method of mathematical statistics for diagnosing the speech behavior of the examinees under research according to the strategy “Involvement/non-involvement of the communicants in the speech event”. This strategy is implemented in the text in the form of three speech formats:

- a personal format with only one communicant (either a sender or a receiver) being involved in the speech act;
- a social format with both communicants (both a sender and a receiver) being involved in the speech act simultaneously;
- objective format with neither of the communicants being involved in the speech act (neither a sender nor a receiver).

With the personal format being realized, the receiver becomes convinced of the author’s authority, righteousness and success. The list of speech signals of the personal format includes personal pronouns **I, you, we** (which are referred to either a sender or a receiver of the text) and their corresponding possessive, objective and reflexive pronouns; parenthesis with modal and attitudinal meaning (e.g., **Fortunately, ...**); Complex Subject phrases (**to be sure/to be likely + Infinitive**); verbs in the Imperative Mood; vocatives (**Ladies and gentlemen**); metatextual expressions and discourse connectors (**as if, as though, rather, however, firstly, secondly**, etc.) which make the text logically structured and show the line of thinking of the speaker or listener; interjections showing the speaker’s subjective attitude to an interlocutor and to the message itself; words **yes** and **know** which reflect the speaker’s standpoint to the utterance; empty lexemes (**well, you know, so to say, well now, so**) used for filling in hesitation pauses which often occur in public speaking, with the former proving the speaker’s involvement in the speech event.

When choosing a social format the author makes the listener psychologically involved in the speech event. The markers of the social format include inclusive pronouns “**we, you**” and their corresponding possessive and objective forms which refer to both the addresser and the addressee at the same time; indefinite personal pronouns “**one, you**”, indefinite pronouns “**any, anyone, every, everyone, each, all, nobody**”; infinitive of purpose; inclusive imperative (“Let’s think”) which encourages the listener for a joint activity.

While realizing an objective format of the strategy in question the speaker intuitively focuses on the content of the message, i.e. communication minus “a human factor”. From the point of view of psychology, such people are characterized as introverts being focused on things, while extraverts are focused on people. The set of speech signals of the objective format includes nouns or pronouns in the 3rd person singular and plural as a subject which are referred to a thing or a person who is not involved into the communication; Perfect Passive; Participle II.

As an example let us look at some utterances and detect speech signals of the formats according to the strategy “Involvement/non-involvement of the communicants in the speech event”. It should be noted that we considered exclusively utterances but not sentences, as the latter refer to the lower level of the

language system, while utterances refer to the discourse as a *speech* phenomenon. An utterance can contain one or more small syntactical groups (SSG) being a unit of analysis of the speech behavior.

- [1] *I hope*, [2] *you won't take my five seconds away from me* (M. Dukakis).

In this utterance the author uses a personal format in both SSGs: in the first one he uses the personal pronoun “I” and in the second one – a personal pronoun “you”, a possessive pronoun “my” and an objective pronoun “me”

- [1] *Our relationships with China are important and* [2] *we have to preserve them* [3] *if we want reliable partners* (W. Clinton).

This utterance consists of three SSGs each of which actualizes the social format: in the first SSG the author uses the possessive pronoun “our”, in the second and third SSGs he chooses the inclusive pronoun “we”.

- [1] *Defense industries are going to convert to civilian industries* (R. Perot).

There are no any speech signals of the personal or social formats in this example, consequently, the author exercises the objective format.

Each of the three formats (personal, social and objective) has a random percentage-based speech genre index (RSGI) which identifies the format. The arithmetic mean of the percentage rate for every format provides an average speech genre index (ASGI) which is meant to be mandatory.

Those RSGIs which are beyond the admissible scope of deviation from RSGI (ΔS) and which are beyond the confidence interval are considered to be diagnostic. By the confidence interval we understand the index gap which includes admissible deviation from ASGI towards increasing or decreasing indices. Such a confidence interval cannot be considered to be applicable for diagnosis. The admissible deviation (ΔS) and the confidence interval are calculated according to formulas of the theory of probability and mathematical statistics (Shulenin, 2012).

Let us review the data processed in Table 1 considering the speech behavior of US Presidency candidates in their pre-election discourse within the implicit impacting strategy “Involvement/non-involvement of the communicants in the speech event” (See Table 1).

From the point of mathematical statistics, RSGIs are viewed as discrete random variables which unlike continuous variables change in a stepwise and discontinuous way. To explain the distribution of those variables we had to stick to the basic numerical characters of a random variable. Such characters include

- Arithmetic mean (which is ASGI in our research). It is calculated according to a formula:

$$\mathbf{S\ format} = (\mathbf{F\ format\ A1} + \dots + \mathbf{F\ format\ An}) \div \mathbf{n}$$

where **S format** stands for ASGI; **F format A1**- the frequency of this format occurrence with the first author; **F format An** - the frequency of this format occurrence with the n-author; **n** - the number of the examined authors (from 1 to 10).

- Dispersion, which is a squared deviation of a random variable (RSGI) from the arithmetic mean (ASGI). Dispersion is calculated according to a formula:

$$D = (RSGI1 - ASGI)^2 + (RSGI2 - ASGI)^2 + \dots (RSGIn - ASGI)^2 \div n$$

where **D** stands for dispersion; **RSGI** – Random Speech Genre Index; **ASGI** – Average Speech Genre Index; **n** - the number of the examined authors.

Table 01. The speech behavior of US Presidency candidates in their pre-election discourse within the implicit impacting strategy “Involvement/non-involvement of the communicants in the speech event”

Formats Authors	Personal Format		Social Format		Objective Format		Total number of SSGs
	SSG	%	SSG	%	SSG	%	
Bush G. Sr.	172	34.4	114	22.8	214	42.8	500
Carter J.	187	37.4	141	28.2	172	34.4	500
Clinton W.	134	26.8	219	43.8	147	29.4	500
Dukakis M.	139	27.8	145	29	216	43.2	500
Kennedy J.	141	28.2	95	19	264	52.8	500
Mondale W.	136	27.2	162	32.4	202	40.4	500
Nickson R.	178	35.6	104	20.8	218	43.6	500
Perot R.	158	31.6	150	30	192	38.4	500
Reagan R.	166	33.2	227	45.4	107	21.4	500
Ford G.	109	21.8	204	40.8	187	37.4	500
Total format SSGs	1520		1561		1919		5000
ASGI		30.4		31.2		38.4	
Deviation (ΔS)		3.2		6.2		5.7	
Confidence interval		33.6-27.2		37.4-25		44.1-32.7	

- Mean square deviation, which is an admissible deviation from ASGI. It is calculated according to a formula:

$$\Delta S = t \times \sqrt{D} \div \sqrt{n}$$

where ΔS stands for an admissible deviation from ASGI; t – confidence probability coefficient which equals to 2.26 for ten respondents; D – dispersion, i.e. a squared deviation of a random variable (RSGI) from the arithmetic mean (ASGI); n - the number of the examined authors.

The confidence interval for the average variable (ASGI) is $ASGI \pm \Delta S$.

2. Problem Statement

The pragmalinguistic experiment is carried out to demonstrate the fact that the way the addresser uses his or her speech and which linguistic units they choose for implementing their impacting tasks, characterizes them themselves in the first place. The pragmalinguistic experiment underway is based on the methodology designed by Matveeva (2004).

3. Research Questions

- What is the unit of research?
- How can the method of mathematical statistics be applicable to carry out the pragamlinguistic experiment?
- What are the stereotyped features of a social and professional group of politicians?

4. Purpose of the Study

The specific character and choice of the objective pragmalinguistic experiment as a method of research is determined by its object, purpose and material. The object of the current study is the verbal behavior of the communicants, namely politicians running for the US Presidency. The ultimate purpose of the research is diagnosing individual characteristics of the communicants in question. The material of the research is their discourse in the form of the texts. The texts are based on the scripts of pre-election debates of ten candidates running for the US Presidency (G. Bush Sr., J. Carter, W. Clinton, M. Dukakis, J. Kennedy, W. Mondale, R. Nickson, R. Perot, R. Reagan, G. Ford) in the speech genre of pre-election discourse.

5. Research Methods

With the ultimate goal of the study taken into consideration, we have employed the method of content-analysis to process a big amount of experimental material. The content-analysis is defined as a research method designed to detect particular units and concepts in a text or a text corpus (Matveeva, 2004). The content-analysis can be of two types: qualitative (semantic) and quantitative (structural). The qualitative content-analysis allows us to conclude about the presence or absence of particular characteristics of the content, for instance, it sets up links between the text under analysis and the information about the author. The quantitative analysis is based on the calculation of certain formal characteristics in texts. Moreover, we are interested not in the formal characteristics themselves but in the extralinguistic reality beyond them, such as the individual characteristics of the author of the text.

The content-analysis carried out in the framework of implicit pragmalinguistics is referred to the second type. More than that, it is modified as it allows the researcher to conclude not about the semantic peculiarities of the text but rather about the individual personal qualities of the author of the text.

6. Findings

The research conducted on the strategy “Involvement/non-involvement of the communicants in the speech event” showed that out of ten candidates running for the presidential post three candidates went beyond the upper confidence interval boundary which accounts to 33.6%. Thus, J. Carter had 37.4%, R. Nickson had 35.6% and G. Bush Sr. had 34.4%. This can be interpreted as the fact that those politicians tended to be authoritarian, self-aggrandized and ego-centric.

The low indices of the personal format which go beyond the low boundary of the confidence interval (27.2%) are observed with G. Ford (21.8%) and W. Clinton (26.8%). This proves their modesty, lack of inclination to overemphasize their self or attract attention to them. They might have a low self-esteem (Laskova & Zueva, 2016, p.74; Rydchenko & Gerasimenko, 2019, p.64). According to an American political analyst M. Hermann, a politician’s low self-esteem can be caused by a heavy need for power as well as approval and support of his decisions on the part of the others (Hermann, 1983, p.5).

The other candidates are within the boundaries of the confidence interval which is 33.6%-27.2% (W. Mondale – 27.2%; M. Dukakis – 27.8%; J. Kennedy – 28.2%; R. Perot – 31.6%; R. Reagan – 33.2%), which makes the diagnosis of their individual qualities impossible.

As far as the social format is concerned, the average speech genre indices going beyond the upper boundary of the confidence interval (37.4%) are observed with R. Reagan (45.4%), W. Clinton (43.8%) and G. Ford (40.8%). This proves their perception of the listeners as eager interlocutors who are ready for cooperation and common responsibility. The manner of communicating with the public through demonstrating common problems has become habitual and subconscious. Choosing speech signals of the social format these candidates have an implicit impact on the public without realizing that themselves. This makes the people feel that they play a significant role in this common activity and support the candidate who appeals to them which goes in line with the genre of the pre-election discourse.

The authors whose average speech genre indices of the social format go beyond the lower boundary of the confidence interval (25%) include J. Kennedy (19%), R. Nickson (20.8%) and G. Bush Sr. (22.8%). It allows us to interpret them as people with a low focus on public.

The social format data of the rest of the authors is within the confidence interval (37.4%-25%) which makes their diagnosing difficult (J. Carter – 28.2%; M. Dukakis – 29%; R. Perot – 30%; W. Mondale – 32.4%).

The consideration of the objective format allows to conclude that the average speech genre index going beyond the upper boundary of the confidence interval (44.1%) is observed solely with one candidate – J. Kennedy (52.8%). Subconsciously sticking to the objective format this politician largely rests on facts while appealing to the public. The informative part of communication is more important for him than emotions or assessments. J. Kennedy tends to be introspective. On the one hand, he is characterized with being concentrated on an object, being logical, rational, realistic, organized, reserved and emotionally stable. On the other hand, such kind of people tend to be shut-in, estranged, unsociable, passive, motivationally challenged, indifferent to the surrounding people. Predominance in the objective format proves that the latent impact on the public results in the latter being under the impression of the speaker’s objectiveness and impartial attitude towards the situation around him.

The authors whose indices of the objective format go beyond the low boundary of the confidence interval are R. Reagan (21.4%) and W. Clinton (2.4%). The informative part of communication might seem less important to them than the personal communication factor. Speaking before the public these two candidates trigger the psychological mechanism of mass contagion through their involvement into the event described.

After interpreting the speech portraits of ten US presidential candidates we managed to diagnose their most common individual qualities including objectiveness, self-righteousness, speech activity, rationality, ability to stand their ground, oratorical skills and eloquence. The fact that the candidates under research belong to one social and professional group connected with the sphere of politics allows us to assert that the above mentioned individual characteristics are stereotyped for the representatives of this social and professional group. These findings are confirmed by other different research (Avedova, Korobka, & Frolova, 2019; Bartashova & Polyakova, 2018; Chanturidze, 2018; Lazarev & Khusane, 2019; Oberemchenko & Basenko, 2017; Ponomarenko, Vasilkova, Volskaya, Kasperova, & Nikolaeva, 2018; Raluca, 2017; Samarina, 2017).

7. Conclusion

Thus, the use of the method of mathematical statistics in a pragmalinguistic experiment allows a researcher to make the mathematically sound verbal portraits of the text senders as accurately and in as much detail as possible. Moreover, it makes it possible to give their interpretation and make a judgement about their communicative and social behaviour.

Acknowledgments

Funding support from the Southern Federal University Programme on supporting publication activities is gratefully acknowledged, as is the collegial contributions of colleagues in the Faculty of Translation and Information Technologies in Linguistics of the Southern Federal University.

References

- Avedova, R. P., Korobka, D. V., & Frolova, M. S. (2019). The means of creating a linguistic image of politicians in a political discourse (based on a pre-election campaign of Donald Trump). Informational and communicational culture: science and education.
- Bartashova, O. A., & Polyakova, S. E. (2018). Manipulating the mechanism of epistemic vigilance in political and legal discourses. *Journal of Siberian Federal University - Humanities and Social Sciences*, 11(5), 707-715.
- Chanturidze, Y. (2018). Functional and linguistic characteristics of Donald Trump's victory and inaugural speeches. *Journal of Language and Education*, 4(4), 31-41.
- Hermann, M. (1983). Assessing Personality at a Distance: A Profile of Ronald Reagan. *Quart. Rep.*, 7(6), 1-8.
- Laskova, M. V., & Zueva, R. S. (2016). Political caricature as a social and cultural universal in modern political linguistics. *Humanities and social sciences*, 1, 71-75. Retrieved from <http://hses-online.ru/2016/01/08.pdf>
- Lazarev, V. A., & Khusane, A. M. A. (2019). The analysis of politician's verbal behavior through the example of Barack Obama's public speaking: linguostylistic aspect. *Humanities and social sciences*, 6. Retrieved from http://hses-online.ru/2019_06.html

- Matveeva, G. G. (2004). Peculiarities of a content-analysis in processing the data of a pragmalinguistic experiment in latent pragmalinguistics. *Language. Discourse. Text. Rostov-on-Don, Part 2*. 125-129.
- Oberemchenko, E. Y., & Basenko, G. V. (2017). A pragmatic approach to studying peculiarities of verbal behavior of German-speaking diplomats. *Philological sciences. Theory and practice issues*, 6-2(72), 135-137.
- Oberemchenko, E. Y., & Degtyaryova, L. M. (2018). The identification of extralinguistic informational components in the structure of a personality of a message sender: pragmalinguistic aspect (Russian and German languages). *Philological sciences. Theory and practice issues*, 9-1 (87), 134-138.
- Pishkova, E. Y. (2016). *Pre-election discourse: pragmatics and translation: monograph*. Southern Federal University. Rostov-on-Don.
- Ponomarenko, E. B., Vasilkova, N. N., Volskaya, N. N., Kasperova, L. T., & Nikolaeva, A. V. (2018). Female political media discourse in American and English linguocultures: Speech strategies and tactics. *Xlinguae*, 11(2), 625-638.
- Raluca, B. (2017). Managing Interpersonal Relationships. *The European Proceedings of Social and Behavioural Sciences*, 6. 39-47.
- Rydchenko, A. S., & Gerasimenko, I. V. (2019). The structure of assessment in a political discourse and its nature (based on speeches of American politicians). *Topical issues of science*, 50, 62-66.
- Samarina, I. V. (2017). *Political discourse as an interpreter's challenge (pragmalinguistic approach): monograph*. Taganrog.
- Shulenin, V. P. (2012). *Mathematical statistics. Part 1. Parametrical statistics: students' book*. Tomsk: NTL. Retrieved from <http://www.fpmk.tsu.ru/sites/default/files/%D0%A5%D1%80%D0%B0%D0%BD%D0%B8%D0%BB%D0%B8%D1%89%D0%B5/%D1%84%D0%BF%D0%BC%D0%BA/%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%D0%B8/Shulenin-MS/MS-part1.pdf>