Deception Detection in Russian Texts

Olga Litvinova and **Tatiana Litvinova**
Voronezh State Pedagogical University
Lenina St, 86, Voronez, Voronezhskaya oblast', 394024
`olga_litvinova_teacher@mail.ru`
`centr_rus_yaz@mail.ru`

Pavel Seredin
Voronezh State University
Universitetskaya pl., 1, Voronez,
Voronezhskaya oblast', 394036
`paul@phys.vsu.ru`

John Lyell
Higher School of Economics
Myasnitskaya ul., 20, Moskva, 101000
`jjlyell@gmail.com`

Abstract

Psychology studies show that people detect deception no more accurately than by chance, and it is therefore important to develop tools to enable the detection of deception. The problem of deception detection has been studied for a significant amount of time, however in the last 10-15 years we have seen methods of computational linguistics being employed with greater frequency. Texts are processed using different NLP tools and then classified as deceptive/truthful using modern machine learning methods. While most of this research has been performed for the English language, Slavic languages have never been the focus of detection deception studies. This paper deals with deception detection in Russian narratives related to the theme "How I Spent Yesterday". It employs a specially designed corpus of truthful and deceptive texts on the same topic from each respondent, such that N = 113. The texts were processed using Linguistic Inquiry and Word Count software that is used in most studies of text-based deception detection. The average amount of parameters, a majority of which were related to Part-of-Speech, lexical-semantic group, and other frequencies. Using standard statistical analysis, statistically significant differences between false and truthful Russian texts was uncovered. On the basis of the chosen parameters our classifier reached an accuracy of 68.3%. The accuracy of the model was found to depend on the author's gender.

1 Introduction

Deception is defined as the intentional falsification of truth made to cause a false impression or lead to a false conclusion (Burgoon and Buller, 1994). Psychology studies show that all types of people students, psychologists, judges, law enforcement personnel detect deception no more accurate than chance (Bond and DePaulo, 2006). Vrij (2010) pointed out that machines are far outperform humans at detecting deception. Therefore, creation of new automatic techniques to detect deception are vital.

Scientists have been studying deception for a long time, attempting to design text analysis techniques to identify deceptive information. However, it is only very recently that methods of modern computational linguistics and data analysis have been employed in addressing this issue (Newman et al., 2003). With the growing number of Internet communications it is increasingly important to identify deceptive information in short written texts. This poses a great deal of challenge as there are no non-verbal cues in textual information, unlike in face-to-face communication.

Obviously there is no single linguistic feature which can with high accuracy partition deceptive from truthful texts. It is thus important to utilize a combination of certain frequency-based text parameters, making up what can be called a linguistic deception profile. The use of a selection of various parameters is vital in analyzing texts for deceptive information and Vrij was right to say that, a verbal cue uniquely related to deception, akin to Pinocchios growing nose, does not exist. However, some verbal cues can be viewed as weak di-

agnostic indicators of deceit (2010). In this way, it seems clear that a combination of features is more effective than isolated categories.

To discern deceptive patterns in communication in the field of Natural Language Processing (NLP), over the last 10-15 years, new approaches to deception detection have arisen, relying essentially on the analysis of stylistic features, mostly automatically collected, as with a vast majority of similarly related NLP tasks, for example, in native language identification (NLI), the task of detecting an authors native language form their second language writing (Shervin and Dras, 2015).

Many recent studies involving automated linguistic cue analysis, including studies concerning deception detection, have leveraged a general-purpose, psycho-social dictionary such as Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2007).

Most papers dealing with automated deception detection were performed using English texts with the evaluation of reliability/truthfulness of the narrative being addressed as a text classification task employing machine learning methods. However, more recently in NLP tasks, methods and models are tested across different languages (see Shervin & Dras (2015) an example of such work in the field of NLI).

To the best of our knowledge, the problem of deception detection as an NLP task has not to date been addressed for Russian, which is connected in large part to the lack of applicable data sets. The lack of standard data sets for this task motivated us to construct our own data set a corpus of truthful and deceptive narratives, written in Russian, on an identical topic from the same author. The corpus contains detailed information about each author (gender, age, psychological testing results etc.), and represents an additional contribution to this work. The corpus is now available on request, but in the near future it will be available only on a specially created site.

Using the previously mentioned corpus, a statistically significant difference between truthful and deceptive texts from the same author, written using an identical theme, was discovered. Utilizing these parameters we offer a new approach to the evaluation of the reliability/truthfulness of the Russian written narrative. The classifier was test separately for both men and women.

2   Related Work

Deception detection (in the framework of computational linguistics) is usually conceived of as a text classification problem where a system should classify an unseen document as either truthful or deceptive. Such a system is first trained on known instances of deception. One of the first studies to employ this approach was the one by Newman et al. (2003), who showed that by using supervised machine learning methods and quantitative text parameters as features one can automatically classify texts as deceptive or truthful. The authors obtained a correct classification of liars and truth-tellers at a rate of 67% when the topic was constant and a rate of 61% overall.

Frequently used features have been token unigrams and LIWC lexicon words starting originally with the above paper by Newman. LIWC (Pennebaker et al., 2007) is a text analysis program that counts words in psychologically meaningful categories. LIWC processes text based on 4 main dimensions: standard linguistic dimensions (1), psychosocial processes (2), relativity (3) and personal concerns (4). Within each dimension, a number of variables are presented, for example, the psychosocial processes dimension contains variable sets representing affective and emotional processes, cognitive processes and so forth. Using the LIWC 2015, up to 88 output variables can be computed for each text, including 19 standard linguistic dimensions (e.g., word count, percentage of pronouns, articles), 25 word categories tapping psychological constructs (e.g., affect, cognition), 10 dimensions related to relativity (time, space, motion), 19 personal concern categories (e.g., work, home, leisure activities), 3 miscellaneous dimensions (e.g., swear words, nonfluencies, fillers), and 12 dimensions concerning punctuation information. The default dictionary contains 2300 words, which are used as the basis for output categories. With a few exceptions, the output variables represent only a percentage of the total words found in LIWC dictionary (Pennebaker et al., 2007).

Several studies have relied on the LIWC lexicon to build deception models using machine learning approaches and showed that the use of semantic information is helpful for the automatic identification of deceit. For example, Mihalcea & Strapparava (2009) used LIWC, measuring several language dimensions on a corpus of 100 false and true

opinions on three controversial topics similar to Newman et al. (2003). They achieved an average classification performance of 70%, which is significantly higher than the 50% baseline. It is worth noting that they also tested the portability of the classifiers across topics, using two topics as training sets and the third topic as a test set. The fact that the average accuracy was significantly higher than the 50% baseline indicates that the learning process relies on clues specific to truth/deception, and it is not bound to a particular topic.

In a similar study of Spanish texts (Almela et al., 2013), the discriminatory power of almost all LIWC variables under the first two dimensions (linguistic and psychological processes), the most relevant ones, have been checked (73.6%).

Until now, very little attention has been paid to address the identification of deception based on demographic data using computational approaches because of the scarcity of resources for this task (Prez-Rosas and Mihalcea, 2014). We are aware of only two other resources for deception detection where demographic data is available (Prez-Rosas and Mihalcea, 2014; Verhoeven and Daelemans, 2014).

In the study (Fornaciari et al., 2013) the authors combined deception detection and personality recognition techniques, in order to get some insight regarding the possible relation between deception and personality traits from the point of view of their linguistic expression. They found that the machine learning models perform better with subjects showing certain kinds of personality traits (when taking into account the author's communication style, deceptive statements are more easily distinguished). However, as the authors themselves suggest, the relatively small amount of respondents allowed them to obtain only a few types of personalities.

In the study by Levitan et al. (2016) for oral speech it was shown that for deception detection, when they included binned NEO-scores, as well as gender and language, in addition to the prosodic and LIWC feature sets, accuracy of the classifier went up to 65%, i.e. there is a 25% relative increase over the majority class baseline and a 13% absolute increase.

For this particular study, we have made use of Linguistic Inquiry and Word Count. We used a LIWC Russian dictionary and also designed our own dictionaries (see explanations below).

The analysis was performed along 104 parameters used to distinguish truthful and deceptive texts.

3 Data and Settings

Firstly, in order to address the subject at hand, we must study a corpus containing truthful and deceptive texts. Collecting this type of text corpus constitutes a scientific task in itself (Fitzpatrick and Bachenko, 2012). Most text corpora being studied presently have a volume of limitations caused by too few respondents, as well as a paucity of deceptive and truthful texts written by the same individual and for the data set as a whole, due in large part to the difficulty of obtaining a control sample of texts in which the same author tells the truth for the sake of comparison. What is important in developing methods of lie detection in texts is to identify changes in the idiolect of the same individual when they produce both deceptive and truthful texts on the same topic. Additionally, as was noted, most corpora contain only English texts.

Another downside of the existing corpora is the shortage of detailed metadata providing the authors personal information (gender, age, education level, psychological testing data, etc.) to establish the effects of personality traits on how deceptive texts are produced.

In our paper we have used a text corpus Russian Deception Bank. It was launched in 2014 as part of a text corpus called RusPersonality (Litvinova et al., 2016). Deception Bank currently contains truthful and deceptive narratives (average text length is 221 words, SD = 15.2) of the same individuals on the same topic (How I spent yesterday) (see example in Table 1).

Since it was not a spontaneously produced language, it was deemed necessary to minimize the effect of the observers paradox by not explaining the ultimate aim of the research to the participants. In addition, to motivate them, the respondents were told that their texts (without information of which of them were truthful and which were not) would be evaluated by a trained psychologist who would attempt to tell a truthful text from a deceptive one. Each respondent whose texts would not be correctly evaluated would be awarded with a cinema ticket voucher.

The number of the authors is N = 113 as of now (46 males, 67 females, university students, all native speakers of Russian) and there are plans

Truthful Text	Deceptive Text
So here we were in Piter and went to the apartment that we had booked, it was not far from the city centre. Having dropped off our stuff, we went on a walk around the city centre and grabbed something to eat. Well, actually every afternoon we spent here was pretty much the same. In the evening we would go to any Pub or Bar and killed time there. Yes, killed time because it was not much fun. Maybe its because the people around werent much fun. Of course it was interesting to visit the museums and other sights of the city but I cant say that really left an impression that it was supposed to and all in all, I didnt feel too happy throughout that trip.	Having come to Piter, first thing we went to the apartment that we had booked, it was in the city centre, straight in Nevskiy, our window overlooked the beautiful views of Piter, especially in the evening when the sun went down, it was very beautiful. Of course you can spend ages walking the streets of the city and never get tired, while you are walking, you cant help being happy about everything you see around you. Every evening we would drive around different places in the city and sure thing, we dont have any clubs or pubs like that back home and I dont think we ever will. The way this city makes you feel is just special.

Table 1: Sample statements from the same author

to extend it. Apart from truthful and deceptive texts by each individual, Russian Deception Bank (as well as all the texts in RusPersonality) comes with metadata which provides detailed information about their authors (gender, age, psychological testing results). Hence, the annotated Russian Deception Bank will enable authors personal features (psychological and physical) to be considered as a factor contributing to the production of their deceptive texts.

We argue that these data are critical in designing an objective method of identifying intentionally deceptive information (Levitan et al., 2016). Each text was entered into a separate text file, and misspellings were corrected. Each of the 226 text files was analyzed using LIWC 2015 and a Rus-

sian language dictionary based on LIWC2007.

We have employed a basic Russian language dictionary that comes with the LIWC software and additionally developed our own users dictionaries (see explanations below). It's worth noting that the program's Russian dictionary is a simple translation of the corresponding English LIWC dictionary. For our study we selected categories that were the least dependent on the content of the texts. Hence the following parameters were selected:

- I STANDARD LINGUISTIC DIMENSIONS (19),

- II PSYCHOLOGICAL PROCESS DIMENSIONS (Affective Processes - 5, Cognitive Processes - 8, Perceptual Processes - 3, Relativity - 3),

- All Punctuation parameters (11).

Users dictionaries were also compiled according to the user manual:

- a dictionary of 20 most frequent function words in Russian Freq FW (20 parameters account for the uses of each word in a text and 1 parameter represents the proportion of the total uses of all such words in a text)

- a dictionary of demonstrative pronouns and adverbs Deictic (1 parameter accounts for the proportion of these words per to the total word length of a text)

- discourse markers DM (10)

- a dictionary of intensifiers and downtowners Intens (2 parameters)

- a dictionary of pronouns as parts of speech Pron (10)

- a dictionary of perception vocabulary PerceptLex (1 parameter)

- a dictionary of pronouns and adverbs describing the speaker Ego (I, my, in my opinion) (1 parameter)

- a dictionary of emotional words Emo (negative and positive, 2 parameters).

All in all, there are 104 parameters. The users dictionaries were compiled using the available dictionaries and Russian thesauri.

It was necessary to compile these particular dictionaries owing to the fact that the Russian dictionary that came with the software was a translation of a corresponding English dictionary and did not stand independent testing, i.e. if all the variables from the first group are identified unambiguously, there are doubts as to the semantic category of the second group and thus they have to be evaluated independently and objectively. The results were processed using SPSS 23.0 software.

4 Experiments

Originally we excluded the parameters that had a frequency of less than 50%. Here frequency of a parameter is defined as a ratio of non-zero values of a parameter to the number of all of the analyzed texts (both truthful and deceptive ones). The selected parameters are identified in the Table.

Further on we calculated and evaluated the variation coefficient of the text parameters that indicates the range of a linguistic parameter in the texts by the same author (Viktor V. Levitsky , 2004). This can be done using the following ratio:

$$V = \frac{\sum_i^n * \frac{|x_{Ti} - x_{Di}|}{x_{Ti}}}{n} * 100\% \qquad (1)$$

where x_{Ti} is the value of the i-th parameter in a truthful text, x_{Di} is the value of the i-th parameter in a deceptive text, and n is a selection size. The computed variation coefficients are shown in Table 2.

A statistical analysis (see Table 2) showed that the computed variation coefficient for the selected parameters ranges significantly. The parameters with correlation coefficient over 50 % were excluded at the next stage (see Levitsky (2004); Litvinova, (2015)).

In order to understand how the parameters of truthful and deceptive texts by the same author change in relation to the absolute value, we calculated the averaged values of each parameter. Table 2 presents a relative change in each parameter in deceptive texts in relation to truthful ones (in percentages).

In order to determine which of the originally selected text parameters could be used in further calculations, we tried to establish a connection between the variation coefficients of the text parameters, frequencies of the parameters in the texts

as well as the difference between the average values of the text parameters in a selection of truthful and deceptive texts. Using the methods of correlation analysis, we found that for a statistical significance level p¡0.05 there is no connection between the frequency of a parameter and a difference between the average values of truthful and deceptive texts. At the same time the calculation of the Pearson correlation coefficient for frequencies of the text parameters and their variation coefficient showed that there is a considerably strong connection r¿0.9 at p¡0.05 (a linear dependence between the two values). Therefore there is one important conclusion to be made: the use of only average values of text parameters in a selection is not always the best option as it does not allow for the distribution of a certain parameter in deceptive and truthful texts by the same author. In order to consider a type of the distribution of text parameters in the corpora of truthful and deceptive texts we used one of the most effective criteria for checking the normality is the use of the Shapiro-Wilk test for normality as it is stronger compared to the alternative criteria for small samples. However, some of the text parameters in deceptive texts (Sixltr, AllPunc, PersPronUser) change their distribution differently. Only the parameters with the following characteristics were chosen for the model to evaluate truthful and deceptive texts:

they are frequent (i.e. occur in no less than half of the texts); vary reasonably in the texts by the same author (on average in a selection); have normal distribution (since we have Student's statistics as a basis of our classifier).

It should be noted that in order to design the models, the parameters that are normally distributed in the corpus of truthful texts were employed. According to the calculation, only 10 parameters are normally distributed in truthful texts (see Table 3).

Hence in deceptive texts in Russian compared to truthful ones on the same topic there are more verbs, conjunctions overall, specifically the conjunctions and, as well as words for cognitive processes overall and inclusive words in particular, additional discursive markers, pronominal nouns, and more personal pronouns (even though it was only revealed at the 10% significance level). In truthful texts there are more prepositions and punctuation marks. Consequently, characteristic features of deceitful texts from a morphological

Parameter	Frequencies in truthful texts, %	Frequencies in deceptive texts, %	Difference in the averaged values of the parameters in deceptive texts in relation to truthful ones, %	Varia-tion Coeff.
Words per sentence (WPS)	100	100	5.53	22.47
Words¿6 letters (Sixltr)	100	100	-0.6	15.09
Function words (FW)	100	100	1.59	11.32
Total pronouns	97	99	2.59	29.14
Total pers pronouns	97	98	6.11	29.78
1st pers singular	94	93	5.71	38.54
1st pers plural	68	67	4.94	86.06
3rd pers singular	88	86	5.39	70.40
3rd pers plural	87	83	-7.57	70.72
Verbs	100	100	3.02	27.37
Adverbs	88	89	-3	59.92
Preposi-tions	100	100	-1.03	19.70
Conjunc-tions	100	100	3.94	27.93
Negations	83	76	2.91	87.17
Quantifiers	95	89	1.75	64.971
Numbers	56	52	13.46	121.19
Cognitive Processes	100	100	3.14	21.711
Insight	90	91	7.78	58.24
Causation	87	84	4.54	73.46
Discrep-ancy	65	60	-7.95	99.84
Tentative	78	76	2.73	83.32
Certainty	86	85	0.5	60.72
Inhibition	52	50	-3.92	117.41
Inclusive	100	98	5.24	33.96
Exclusive	83	76	2.73	72.3
Perceptual Processes	92	88	-11.67	58.07
Seeing	60	58	-13.75	95.76
Hearing	66	70	12.90	92.06
Feeling	54	47	-25.75	117.58
Space	100	99	1.18	25.5

Parameter	Frequencies in truthful texts, %	Frequencies in deceptive texts, %	Difference in the averaged values of the parameters in deceptive texts in relation to truthful ones, %	Varia-tion Coeff.
Time	100	98	-4.45	32.91
All Punc-tuation (AllPun)	100	100	-6.78	16.21
Period	100	100	-4.83	25.19
Comma	100	99	-1.52	32.66
Dash	79	61	-31.62	86.14
Freq FW	100	100	-0.05	10.83
(and)	98	98	7.31	35.39
(in)	97	96	-1.72	45.3
(not)	82	76	-2.36	85.88
(on)	89	88	4.97	62.78
(with)	80	86	10.71	74.6
(that)	77	71	5.88	75.77
(over)	50	58	16.36	114.15
(but)	68	58	-6.66	101.99
(like)	65	65	1.17	94.18
Deictic	95	90	-5.55	94.18
DM Additions	97	98	7.67	36.16
DM Sub-stitutions	74	76	27.96	77.84
Intensi-fiers	63	65	-17.77	91.19
Noun-like Pron	98	96	5.65	40.68
Adverb-like Pron	85	85	2.18	67.77
Adjective-like Pron	77	75	11.01	83.49
Number-like Pron	53	52	-21.31	110.55
Per-sPronUser	96	93	6.91	45.34
Per-ceptLex	58	50	-17.64	106.98
Ego	92	91	2.8	43.24
Positive Emo	80	75	-1.85	86.88

Table 2: Text analysis data

Parameter	Mean	t	p
Total pers pronouns		1.655	0.1
D	10.932		
T	10.298		
Verbs		3.979	0.0001
D	15.460		
T	15.305		
Prepositions		-3.352	0.001
D	13.366		
T	13.484		
Conjunctions		5.848	0.0001
D	8.496		
T	8.163		
Cognitive Processes		11.916	0.0001
D	18.050		
T	17.488		
Inclusive		9.236	0.0001
D	9.202		
T	8.735		
AllPun		-3.382	0.001
D	20.320		
T	21.801		
(and)		11.726	0.0001
D	3.948		
T	3.685		
DM Additions		12.915	0.0001
D	3.912		
T	3.658		
Noun-like Pron		5.798	0.0001
D	7.610		
T	7.212		

Table 3: Statistical differences between deceptive (D) and truthful (T) texts

standpoint are a greater amount of verbs, personal pronouns, pronominal nouns, conjunctive relationship markers, and a lesser amount of prepositions and punctuation marks. As it seems, this is connected to the fact that texts which contain such characteristics demand less cognitive effort in their creation, however this is merely a proposition and, of course, would need to be verified.

The basis of this model is Rocchio classification. For text classification we first created two centroids $[ST1...ST10]$ and $[SD1...SD10]$, based on the previously attained values of STi and SDi. These serve as averages of all the 10 various chosen parameters in both the truthful and deceitful texts..

For each text, in order to find whether they are or are not truthful, we then need to find the vector S, which consists of all elements Si, i.e. the 10 aforementioned parameters. Our classifier then determines the truthfulness of a text based on the similarity between the vectors of the test documents and the centroids specific to truthful and deceitful texts.

To measure the similarity of the the test set text vectors and the centroids we utilized the cosign similarity of the vector and the centroid. However, our experiment shows that in this case purely measuring cosine similarity actually has a very weak ability to classify texts. Thus with out experiment we decided it was better to use the function listed below (2), which represents a hybrid of both the Euclidean distance formula and the cosine simi-

larity between two vectors.

The similarity of vector S and centroid S_T is measured thus:

$$\chi_T^2 = \frac{1}{n} \sum_i^n \frac{(S_{Ti} - S_i)^2}{S_{Ti}} \qquad (2)$$

Analogously, the similarity of vector S and centroid S_D is measured as such:

$$\chi_D^2 = \frac{1}{n} \sum_i^n \frac{(S_{Di} - S_i)^2}{S_{Di}} \qquad (3)$$

We will assume, that in order to determine the type of text, it is sufficient to compare the values of χ_T^2 and χ_D^2. The text in question will be classified as deceitful if $\chi_T^2 > \chi_D^2$ and classified as truthful if $\chi_T^2 < \chi_D^2$

In order to test this approach, before designing the model, the texts were divided into the learning and test sets (70 %, i.e. 158 texts for train and 30 %, i.e. 68 texts for tests). In order to evaluate the suggested model the overall accuracy, which is the percentage of texts that are classified correctly, was computed. The accuracy of the suggested approach evaluated on the total test corpus was 68.3 %. Since data set has an equal distribution between truthful and deceptive texts, the baseline is 50 %.

In our study the accuracy of the model was tested individually for males and females and so was the overall one. The classification accuracy for males was 73.3 % and 63.3 % for females. Hence the analysis indicates that models for detecting deception in written texts could be further improved by considering the characteristics of their authors.

5 Conclusion

The average classification accuracy of 68.3 % although higher than the 50% baseline indicates that classification task is difficult and more research is needed to discover what methodology could be appropriate to improve the results. The analysis revealed that models for detecting deception in written texts could be significantly improved by considering the characteristics of their authors. Males and females lie in different ways. Thus models should be further designed for deceptive/truthful texts by males/females, for peoples of different ages, different psychological profiles in order for

them to be more accurate. This is a promising research field, however it has not been properly addressed as part of text-based deception detection because of the scarcity of resources for this task. We assume that the corpus of deceptive and truthful Russian texts with metadata providing various personal information about their authors (gender, age, education, results of psychological and neuropsychological testing and so forth) would contribute to further improvements in this field. Currently we are extending our corpus using real texts - recordings of job candidates in one of Russia's largest industrial companies. All of the candidates took a series of psychological tests. Parts of the interviews were classed as truthful/deceptive using polygraph readings, collection of extra information about the candidates as well as follow-up interviews. To the best of our knowledge, the corpus being designed has no equivalents.

The corpus is to be further expanded by increasing the number of texts as well as respondents. The features of the production of deceptive texts depending on the gender, age and psychological characteristics of their authors are also to be identified. There are plans to design a corpus of deceptive and truthful texts in the first language (Russian) as well as the second language (English) by the same author in order to identify possible structural and lexical differences between the linguistic expression of deceit in both languages.

Further we plan to expand upon our list of parameters and utilize various machine learning algorithms for classifying truthful and deceitful texts, and then compare these results to the method mentioned in this paper.

Acknowledgments

References

Angela Almela, Rafael Valencia-Garca, and Pascual Cantos. 2013. Seeing through deception: A computational approach to deceit detection in written communication. *Linguistic Evidence in Security, Law and Intelligence (LESLI)*, 1(1).

Charles F. Bond and Bella M. DePaulo. 2006. Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3):214–234.

Judee K. Burgoon and David B. Buller. 1994. Interpersonal deception: Ill effects of deceit on perceived communication and non-verbal behavior dynamics. *Journal of Nonverbal Behavior*, 18(2):155–184.

Eileen Fitzpatrick and Joan Bachenko. 2012. Building a data collection for deception research. In Eileen Fitzpatrick, Joan Bachenko, and Tommaso Fornaciari, editors, *Proceedings of the EACL Workshop on Computational Approaches to Deception Detection*, pages 31–38.

Tommaso Fornaciari, Fabio Celli, and Massimo Poesio. 2013. The effect of personality type on deceptive communication style. In *Intelligence and Security Informatics Conference (EISIC)*. IEEE.

Sarah Ita Levitan, Yocheved Levitan, Guozhen An, Michelle Levine, Andrew Rosenberg, Rivka Levitan, and Julia Hirschberg. 2016. Identifying individual differences in gender, ethnicity, and personality from dialogue for deception detection. In *NAACL Workshop on Computational Approaches to Deception Detection*, San Diego.

Tatyana Litvinova, Olga Litvinova, Olga Zagorovskaya, Pavek Seredin, Alexander Sboev, and Olga Romanchenko. 2016. "ruspersonality": A russian corpus for authorship profiling and deception detection. In *International FRUCT Conference on Intelligence, Social Media and Web (ISMW FRUCT 2016)*, pages 1–7, St. Petersburg, RU.

Tatiana Litvinova. 2015. On the problem of stability of parameters of idiostyle. In *Proceedings of Southern Federal University*, pages 98–106.

Rada Mihalcea and Carlo Strapparava. 2009. The lie detector: Explorations in the automatic recognition of deceptive language. In *Proceedings of the Association for Computational Linguistics (ACL-IJCNLP 2009)*, pages 1–7, Singapore, SG.

Matthew L. Newman, James W. Pennebaker, Diane S.Berry, and Jane Richards. 2003. Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29(5):665–675.

James W. Pennebaker, Cindy K. Chung, Molly Ireland, Amy L. Gonzales, and Roger J. Booth, 2007. *The Development and Psychometric Properties of LIWC2007*. Linguistic Inquiry and Word Count (LIWC), Austin, TX.

Vernica Prez-Rosas and Rada Mihalcea. 2014. Gender differences in deceivers writing style. *Journal Lecture Notes in Computer Science*, 8856:163–174.

Malmasi Shervin and Mark Dras. 2015. Multilingual native language identification. *Natural Language Engineering*, 1:1–53.

Ben Verhoeven and Walter Daelemans. 2014. Clips stylometry investigation (csi) corpus: A dutch corpus for the detection of age, gender, personality, sentiment and deception in text. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC14)*, Reykjavik, IS. European Language Resources Association (ELRA).

Viktor V. Levitsky . 2004. *Quantitative methods in linguistics*. Nova Kniga, Vinnytsia, UKR.

Aldert Vrij. 2010. *Detecting lies and deceit: Pitfalls and opportunities*. John Wiley and Sons, Chischester, UK, 2nd edition.