# Towards an Analyzer of Emotions for Texts in Russian in Bilingual Perspective

**Anastasia Kolmogorova, Alina Malikova, and Alexander Kalinin**

## 1 Introduction

The emotional text analysis of textual data has seen a great success these last 5 years. Mainly it is due to two facts. First, the technologies in AI have achieved the level where it becomes necessary to ensure not only the exchange of verbal messages between humans and machines, but also their mutual empathetic understanding [1]. Second, the last decade, the way of feeling emotions and speaking about them has drastically changed – people do not hide them anymore, in contrast – to be extremely sensitive to minor affective fluids becomes a new norm of emotionality in social networks [2].

Within a new research path known as Affective Computing paradigm, a particular attention is paid to how computer could become aware about emotional state of a person by processing different (and textual too) data from her. Our project is especially focused on textual data processing for analyzing a palette of emotions expressed in it. The main advantage of the project is that its aim is to elaborate a valid instrument to predict the emotion in texts in Russian.

However, while doing the project, we faced several cases, which showed us some limitations of the emotional analysis technics. First, it concerns the selection of assessors for emotional texts dataset annotation. Their emotional sensitivity, very subjective and depending on many heterogeneous factors, is put as a basis for machine learning algorithms. As Russia is a multilingual territory with many ethne speaking Russian as well as their autochthonous idiom, we became interested in

A. Kolmogorova (✉)
National Research University Higher School of Economics, Saint Petersburg, Russia

Siberian Federal University, Krasnoyarsk, Russian Federation

A. Malikova · A. Kalinin
National Research University Higher School of Economics, Saint Petersburg, Russia



R. Bolgov et al. (eds.), *Digital Geography*, Springer Geography,
https://doi.org/10.1007/978-3-031-50609-3_13

how differently the monolingual Russians, on one hand, and the bilinguals speaking Russian and one more national language, on the other hand, act as assessors while doing the emotional annotation of texts of social networks in Russian.

This chapter aims to present the results of the experiment with two non-discrete annotations of the same sample of texts in Russian done: first, by Russian monolinguals and second, by Tuvan-Russian bilinguals.

The chapter is structured as follows: Section 2 describes the ideology of emotion analysis and main models of emotions implemented in emotion analysis tasks and outlines the problem of differences of detecting emotions in text by mono- and bilinguals; Section 3 features the methodology of the experiment including data and annotation tools; Section 4 presents the obtained results and opens a discussion; in Sect. 5, we draw the perspectives of further work and, finally, we make conclusions.

## 2 The Emotion Analysis of Text Data in Bilinguals: Preliminary Remarks

Emotions are our permanent companions long all our life: Whatever we do – when we are thinking, dreaming, or speaking – their subtle fluids always affect our behavior. The first attempts to catch them in text data by using computational methods gave birth to the approach known nowadays as Sentiment Analysis. Since the beginning of twenty-first century, the term is employed to describe both the specific task of determining a document's sentiment polarity and the more general problem area of automatically identifying a range of emotional and attitudinal "private states" [3]. However, with time and due to the development of new technologies, the researchers felt the insufficiency of binary or ternary classifications (i.e., positive or negative, or sometimes neutral) proposed within the sentiment approach. Thus, from sentiment analysis emerges emotion analysis of text data. The scholars from the field presume that the goal of emotion analysis is to recognize an emotion, rather than a sentiment in texts, and to find the best input representations and algorithms to classify the data (sentences, paragraphs, entire documents) into predefined classes of emotions [4].

Each technic in emotion analysis is obligatory based on a model of emotions. Thus, to describe emotions, the scholars operate either dimensional or categorical models [5].

## 3 Models of Emotions Exploited in Emotion Analysis

The categorical models treat emotions as distinct entities which number differs from 6 categories in Ekman's classification [6], via 9 – in Tomkins affect typology [7] and, finally, towards famous Plutchik's wheel of emotions [8]. Many projects in Affective Computing exploit the aforementioned emotional systems or their parts to perform the task in emotion detection, emotion recognition, and emotional text analysis (e.g., [9–11]).

However, in the last years, research practices turn towards dimensional models of emotions, because they allow obtaining much more nuanced emotion representation of text. Dimensional models consider particular emotions as points in emotional continuum that is mostly visualized as 3-D space because of three parameters taken into account. This is the case of VAD model [12] (Valence (polarity), Arousal (a calm-excited scale), and Dominance (perceived degree of control in social situation)) and Osgood's multi-dimensional scaling (MDS) model [13] using Evaluation, Potency, and Activity scales to draw the continuity of emotions.

One of the most recent achievements in this domain is the three-dimension model of H. Lövheim well-known as Lövheim Cube. It considers three neurotransmitters that are supposed to trigger emotions – serotonin, noradrenaline, and dopamine [14]. Our project in Emotional Text Analysis relies on this continual emotional representation.

Hugo Lövheim stated that the different levels of the three monoamines – dopamine, noradrenaline, and serotonin – are linked with the emotional states experienced by an individual. He represented the correlations existing between emotions and monoamines by orthogonal monoamines axes. Altogether, the axes form a cube. He placed eight emotions as eight vertices of the cube.

Thus, on the first vertical axis, depending on the level of noradrenaline, he puts the emotions of Shame/Humiliation, Fear/Terror, Enjoyment/Joy, Contempt/Disgust (each emotion has a double name because the first nomination indicates the weakest degree of its manifestation, while the second the strongest) on the axis bottom (minimal level of the noradrenaline) and the emotion of Distress/Anguish, Anger/Rage, Interest/Excitement and Surprise – on the axis top. According to the level of serotonin, the same emotions were placed differently on another "left to right" axis: Shame/Humiliation, Fear/Terror, Distress/Anguish, Anger/Rage – at the beginning because of the minimal monoamine's level, and Contempt/Disgust, Enjoyment/Joy, Surprise and Interest/Excitement on the end of the axis.

Finally, the third "front to back" axis showing the level of the dopamine starts from the low-dopaminergic emotions of Shame/Humiliation, Distress/Anguish, Contempt/Disgust, Surprise and goes to the high-dopaminergic emotions of Fear/Terror, Anger/Rage, Enjoyment/Joy, and Interest/Excitement.

Localized in this way, the emotions form four oppositions, which correspond with the supporting diagonals inside the cube:

1. Distress – Enjoyment ("–" serotonin; "+" noradrenaline; "–" dopamine/"+" serotonin; "–" noradrenaline; "+" dopamine).
2. Anger – Disgust ("–" serotonin; "+" noradrenaline; "+" dopamine/"+" serotonin; "–" noradrenaline; "–" dopamine).
3. Shame – Excitement ("–" serotonin; "–" noradrenaline; "–" dopamine/"+" serotonin; "+" noradrenaline; "+" dopamine).
4. Fear – Surprise ("–" serotonin; "–" noradrenaline; "+" dopamine/"+" seroto-nin; "+" noradrenaline; "–" dopamine).

Thus, such dichotomies allow us to conceptualize paired emotions as antipodes: from the point of the combination of three monoamines and their level, for example, fear is a negative form of surprise; anger is aggressive antonym of disgust, etc. We used these four dichotomies to design the interface for the emotion assessment in experiments with mono- and bilinguals detailed in Sect. 3.

While organizing the annotation procedure and then interpreting its results, we discovered that the assessment magnitude values – a magnitude of a pooled vector that can be thought of as the distance from the center of the cube to the point of a particular text estimation by a certain user [15] – depend on the individual's empathy score: More empathetical the person is, more radical assessments she gives. In such case, the solution we found was to eliminate outliners in assessments. However, the case was insightful and symptomatical: It demonstrates that emotion assessment procedure is very volatile and sensitive to numerous psycholinguistic factors. In other words, each model is a priori biased because it depends on the dataset it is trained on, but the data itself depends on the annotation specificity.

To prove the idea, we formulate a research question:

Does emotion annotation done by monolingual differ from the annotation conducted by bilinguals? If yes, it means that the analyzer trained on the monolingual emotional introspective judgments won't run correctly for bilinguals.

## *3.1 Emotions in Bilingual Brain and Linguistic Consciousness*

Bilinguals' cognitive system excites researchers in cognitive sciences, linguistics, and psychology who try to see this "wonderful mechanics" that allows bilinguals to process the input in two or more languages. The research focus falls on language lateralization in bilinguals [16], mechanisms of heritage bilingualism [17], cognitive specificities of child bilingualism [18], and emotion activation by first and second language words.

There exist two main theoretical conceptions explaining how words in bilingual mental lexicon are linked with emotions.

According to Revised Hierarchical Model [19], bilingual mental lexicon has both lexical and conceptual levels for each language. Even if in first language (L1) the link between word and concept is stronger than in second language (L2), the second language word has a compensatory ability to activate simultaneously conceptual level in first and second language lexicons. In this way, when perceiving emotion bearing words, bilinguals process conceptual information about emotions stocked in both mental lexicons. In the frame of another conception, the Emotion Context-of-Learning Theory [20], the researchers postulate that L1 words are better activators of emotions than L2 words because they have been leant in more affect saturated ambience – in communication with parents and caregivers.

The thesis about a correlation existing between language status (L1 or L2) and effectiveness of cognitive processing of emotion words or emotion bearing words was confirmed in a few experiments with lexical decision task tests [20], procedures for measuring brain activity using EEG [21], and fMRT [22].

However, in the field of emotion analysis, the specificity of emotion in text assessment by bilinguals has not yet been studied. It is worth saying that nowadays, the well-known annotated datasets in Russian and other languages represent text data collections assessed uniquely by monolinguals [9, 10, 23–25] – we did not find in Web of Science and Scopus databases any paper detailing a dataset where text data was evaluated by bilinguals.

However, given the multinational territory of Russia, the problem seems waiting for its solution. Moreover, a particular characteristic of Russian-national bilingualism is that Russian functions as a communicatively strong language mediating interactions between Russian majority and national minorities. The necessity to ensure a mutual understanding between them on the level of emotion perception conditions the relevance of such study – we must be sure that texts in Russian addressing Tuvan or Khakass people make them feel the emotions we wanted to express.

The bilingual community that interests us within our project in emotion analysis is Republic of Tuva.

### *3.2 Bilingualism in Tuva*

The region of Krasnoyarsk borders with another territory – the Republic of Tuva where the autochthonous people are, mostly, Tuvan-Russian bilingual. The growth of Tuvan-Russian bilingualism as social phenomenon begins since the moment when Tuva had integrated the Soviet Union as an autonomous region in 1944.

During the last decades, the status of Russian language has changed many times: The Constitution of 1996 postulated that Russian in Tuva was the official and federal language, according to the Constitution of 2001 г. – only official language, the Low of languages in Tuva (2003) fixed it as official and international language [26]. Due to some social changes, in the 1980s–1990s, Russian lost some of its functions in Tuvan community and it led to the reduction of number of those who speak Russian on the territory [27]. Nowadays, the situation seems to be very unbalanced: The urban population prefers using Russian at work, in institutional context, and Tuvan – to communicate with family members and friends; in villages, people speak, mostly, Tuvan language, even if in schools, since 2018, all subjects are taught in Russian [28, 29]. Answering to our questionnaire, 65 Tuvan informants (mean age 21.3, students) pointed out that their native language is Tuvan; 70% of them speak Tuvan only at home, 6% at work; only 9% speak Russian at home, 62% at work or in the university. As for subjectively assessed difficulties, 56% of informants do not have any difficulties in using Russian, 36% of respondents declare having some problems from time to time, and 8% complain to have regular problems in speaking Russian.

In sum, the researchers classify the general linguistic situation on the territory of Tuva as unbalanced and rapidly changing in the last decade. The conclusion reflects entire Tuvan population and is applicable to our pull of informants too.

## 4 Methodology

Section 3 includes two kinds of methodological remarks: first, description of our methodology when working with monolinguals and second, design of our experiment with bilinguals. Even if the base of these two experiments is similar, there exists some specificity to be mentioned.

## 5 Data

For our whole project, we selected 80–100 words emotional posts from the most popular Russian social network VKontakte. Using emotive hashtags as cues, we retrieved about 4000 texts from three public groups: Overheard, Room № 6, and Caramel. Let us name it as "Basic data."

For two experiments in monolinguals and bilinguals, we used a selection of 48 texts randomly retrieved from the whole collection: six texts for each of eight emotions according to Lövheim model. We will call it "Experimental data."

## 6 Annotation Procedure in Monolinguals

When we start our work, we faced the problem of absence of datasets annotated in an appropriate non-discrete way, which corresponds better to Lövheim dimensional model. Thus, we had to obtain data by ourselves by using Yandex Toloka crowdsourcing engine, which provides pools of Russian speaker assessment, tools for organizing user-interface for annotation tasks, and orchestrating acquisition process [31, 33].

Since it was quite difficult to make the assessor specify the point in 3-d space, we went another way – the use internal diagonals of the Cube as scales along which the estimation was performed (see Fig. 1). The default position was a center of the Cube, which was treated as neutral configuration. Assessors were adjusting the positions of the point on the diagonals to specify the "presence" or intensity of a given emotion. The scales proposed to informants were four (see Fig. 2): Shame–Excitement, Disgust–Anger, Fear–Surprise, and Enjoyment–Distress. The main limitation of the proposed assessment tool is that our informants could not assign any emotional value to both of opposed emotions forming a scale – they could do it

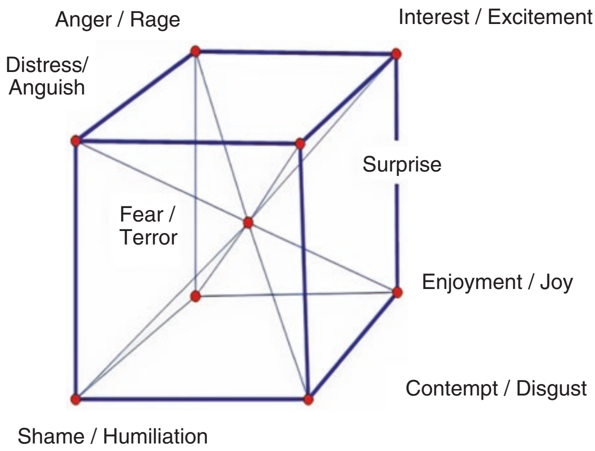



**Fig. 1** Lövheim Cube diagonals

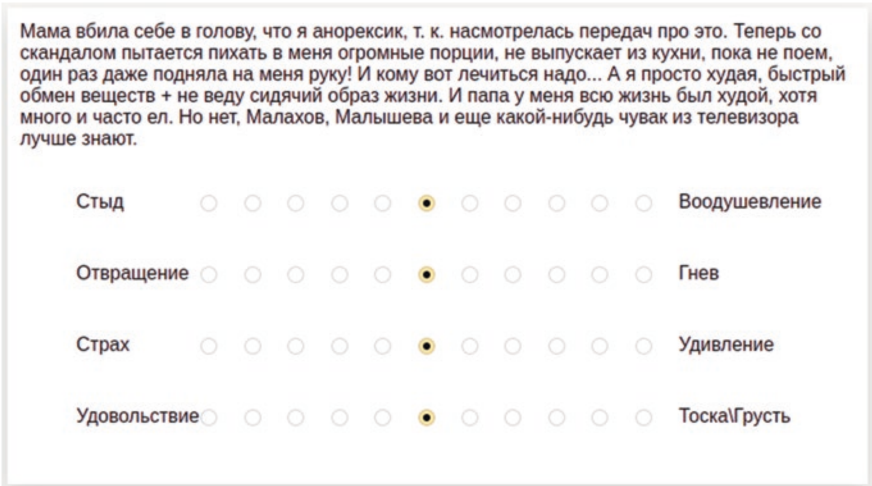



**Fig. 2** Example of annotation UI

for only one of them. In this way, we lose some of the nuances of annotation; instead, we gain in number of annotated texts and speed of annotation.

In sum, 2000 informants from Toloka took part in annotation. It means that our "Basic data" was annotated by two or more informants from Toloka – it will be our "Basic dataset."

In this chapter, when we speak about the comparison of assessment in bilinguals and monolinguals, we use as monolingual dataset the sample of 48 texts, which were annotated by 174 respondents selected from the mentioned above pull from Toloka according to the criteria of age and gender to fit more our pull of bilinguals in second experiment. Each of 48 texts was annotated by each of 174 annotators ("Experimental monolingual dataset").

A sample of user interface (UI) that we designed by ourselves for the task assessment is presented in Fig. 2. The outliners were eliminated from the body of estimations.

As a result, we had, first, raw estimations and, second, a mapping of texts and four corresponding vectors with angle derived from diagonal vector and magnitude derived from user estimation scalar. To transform users' answers into Lövheim Cube entities (point described by neurotransmitters' values), we used a pooling technique of averaging of all four vectors.

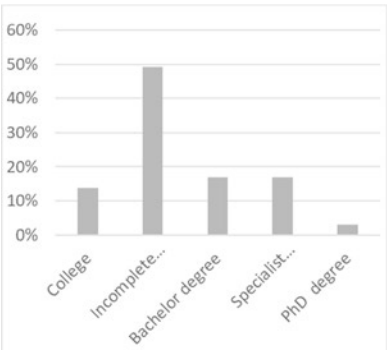
In this way, we could transform terms of emotions and their intensity into coordinates resembling three neurotransmitters. This part is more thoroughly explained in our previous work [30].

Regarding target variables, we faced a dilemma – whether to use pooled Cube coordinates derived from diagonals in terms of neurotransmitter values or to choose diagonals assessment values "per se." From the point of view of following Lövheim Cube ideology, it is better to operate in neurotransmitter space to provide "neurobiological coordinates," but raw assessment values are more interpretable from human point of view. Thus, we decided to experiment both with 3-d coordinates and raw values.

## 7 Annotation Procedure in Bilinguals

To verify our hypothesis about differences in perceiving emotions in texts in monolinguals and bilinguals, we formed a group of 65 Tuvan-Russian bilinguals. All informants were recruited by posting an advertisement on social network VKontakte in groups visited by users from Tuva. Most of them were 18–25 years old students (see Fig. 3). 70% of informants were women and 30% men.

For each informant, the annotation task was preceded by a questionnaire Bilingual Language Profile (BLP) [32] consisting of 4 modules: (1) Language History (including questions about age of acquisition and age of comfortable use of



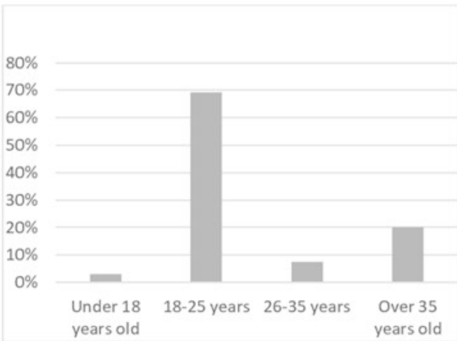




**Fig. 3** Bilingual informants' education and age

each language); (2) Language Use (including questions about percentage of use in an average week with friends/with family/at school or work, etc.); (3) Language Proficiency (including questions about informant's self-assessment of language competences: speaking, understanding, writing, and reading); and (4) Language Attitudes (including questions about the informants' cultural identification made via languages).

Each question response in the BLP is scalar and associated with a certain point value. By calculating the final informant's score, we can define his dominant language. Using this criterion, we divided all bilingual informants into three groups: those with Tuvan as dominant language (T), those with dominant Russian (R), and those who show characteristics of balanced bilingualism (B).

After having answered to the questionnaire, informants started to assess 48 texts (8 for each emotion according to Lövheim model) in Russian using the same interface with four scales as monolinguals had already used (Fig. 2). Each of 48 texts was assessed by each of 65 annotators ("Experimental bilingual dataset"). The outliners were eliminated from the body of estimations.

# 8 Results and Discussion

## *8.1 Comparison of Raw Data*

When speaking about raw data, we mean estimations obtained for each text under annotation directly on each scale.

In Table 1, we show Student's *t*-test values obtained in comparison of 48 texts estimations given by groups of monolinguals and bilinguals on four scales. Results are regrouped according to the dominant emotion assigned to texts by informants in "Basic dataset": We applied statistical measures to groups of "sad," "anger," and other texts evaluated as such by 2000 assessors in our "Basic dataset."

**Table 1** Student's *t*-test values obtained in comparison of 48 texts in Russian estimations given by group of monolinguals and bilinguals on four scales

| Category of texts | Shame–excitement | Disgust–anger | Fear–surprise | Enjoyment–distress |
|---|---|---|---|---|
| All texts | 0.168035836 | **0.000017095** | **0.000003959** | 0.650112520 |
| Anger | **0.024581956** | 0.143814899 | **0.000036465** | 0.721777102 |
| Distress | **0.018088926** | 0.834818531 | 0.155642855 | **0.000014443** |
| Disgust | **0.044788024** | **0.000000001*10$^{-6}$** | 0.072912310 | 0.965666585 |
| Enjoyment | **0.023280081** | 0.083019252 | 0.158893921 | 0.293755615 |
| Excitement | **0.000000340** | 0.307239340 | 0.179120165 | 0.117777343 |
| Fear | 0.089459445 | 0.665581657 | 0.074643384 | **0.034271318** |
| Shame | 0.835649519 | 0.351717329 | 0.918886773 | 0.834782760 |
| Surprise | **0.000078409** | 0.697143996 | **0.034617159** | 0.066264496 |

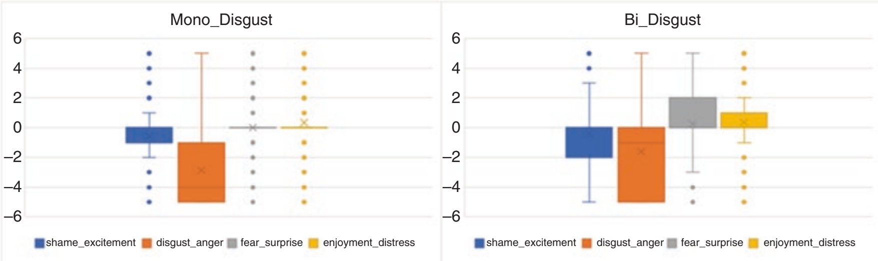



**Fig. 4** Range chats for the group of texts with dominant emotion of Disgust in monolinguals and bilinguals

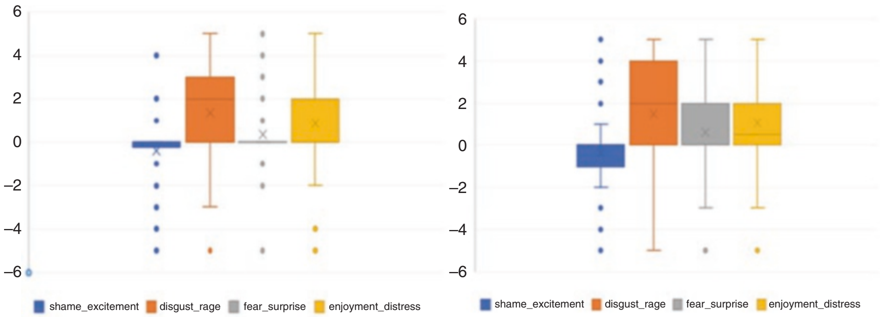



**Fig. 5** Range chats for the estimations of example 1 in monolinguals and bilinguals

As we can see, the significant *p*-values (in bold) appear not rarely. For all the integrity of texts, the difference is significant in estimations given by mono- and bilinguals on Disgust–Anger and Fear–Surprise scales. The scale strongly affected by discrepancies in estimations in two groups is Shame–Excitement scale where only two of eight text categories have no relevant differences in estimations.

In fact, a more detailed analysis demonstrated that the most of differences concern not primary emotions but secondary. Range chats (see Fig. 4) show that in the group of texts expressing Disgust (according to the Basic dataset), both monolinguals and bilinguals agree with primary emotion, but as for the secondary, bilinguals see in texts more of Surprise and Distress than monolinguals.

To focus on the discrepancies, we can analyze one example of text and its estimations:

Муж работает в крупной продуктовой компании. Достали родственники, которые постоянно удивляются тому, что он не тащит домой продукты. Да, блин, мы зарабатываем прилично и можем себе позволить купить все, что надо. Какой смысл из-за пакета молока лишиться достойной работы и испортить репутацию (*The husband works for a large grocery company. Got relatives who are constantly surprised that he does not drag home groceries. Yes, damn, we earn decently and can afford to buy everything we need. What’s the* *point of losing a decent job and ruining your reputation because of a milk carton*).

In general, the main evidence is that bilinguals' estimations (see Fig. 5) show larger assessment magnitude than in monolinguals: On "Shame–Excitement" scale, most estimations in bilinguals oscillate between 0 and −1, in monolinguals – about 0; on "Fear–Surprise" scale in bilinguals – between 0 and 2, in monolinguals – 0; on "Enjoyment–Distress" scale, the values do not show any discrepancy.

While comparing raw estimations in monolinguals and in (a) bilinguals with dominant Russian language, (b) bilinguals with dominant Tuvan language, and (c) balanced bilinguals, we see that the evaluations differ drastically when we compare monolinguals with bilinguals with dominant Tuvan language. In case b, we observe the significance of discrepancies on all four scales (see Table 2).

## *8.2 Comparison of Data Submitted to Pooling Technics*

To confirm the tendencies that we saw in raw data, we applied to both experimental datasets some pooling technics (see Sect. 3.2) allowing us to transform raw text estimations into its "emotional coordinates" in the Cube space.

The visualized data (see Fig. 6) echoes our findings demonstrated in the previous subsection: We see the clouds of red points slightly move towards the vertices of Distress and Surprise. If we have a look at histograms displaying in comparison the distribution of pooled Cube coordinates of texts derived from "Fear–Surprise" and "Enjoyment–Distress" diagonals in two experimental datasets (see Figs. 7 and 8), we can state that bilinguals give more of radical assessments to the pole of Surprise than monolinguals and prefer moderate but redundant estimations of the emotion of Distress.

It is to notice that histograms show normalized (not absolute) values because the number of informants in two groups was not the same. The presented scales (see Figs. 7 and 8) are logarithmic – the segment below $10^{-1}$ has an absolute value two times less than the same segment in the middle.

Finally, Tuvan-Russian bilinguals appear, according to the results of conducted research, more sensitive to emotions of Surprise and Distress than Russian monolinguals.

Searching to explain the observed inconsistency between monolinguals' and bilinguals' emotions in texts assessment, we see many probabilities. The differences may be conditioned by:

**Table 2** Student's *t*-test values for estimations in groups of Tuvan-Russian bilinguals with dominant Tuvan and monolinguals

| Shame–excitement | Disgust–anger | Fear–surprise | Enjoyment–distress |
|---|---|---|---|
| 0.018744842 | 0.000107612 | 0.00000002 | 0.040572310 |

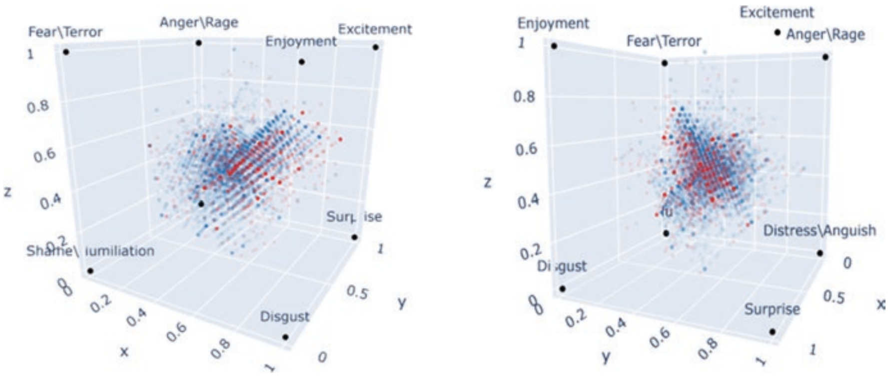



**Fig. 6** Emotional coordinates of texts annotated by monolinguals (in blue) and by bilinguals (in red) in two projections

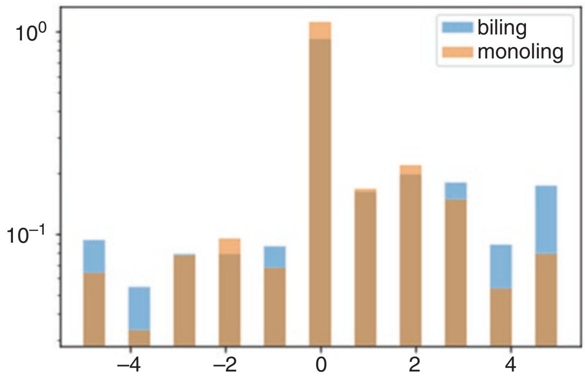



**Fig. 7** Histogram displaying in comparison the distribution of pooled Cube coordinates of texts derived from "Fear (−5) – Surprise (+5)"

- Psycholinguistic factor of language competence (native language/second language).
- Factor of national mentality deeply rooted in traditions and rituals of showing and managing emotional states in Russian linguo-cultural community and in Tuvan society.
- Ethnic factor of different environmental triggers for emotions (for Russians a sedentary lifestyle is proper, and Tuvans are nomads).

However, the main contribution of such experiment is that it proves: When we are speaking about technics and models of ETA, there could not be any universal annotated dataset nor dictionary with emotive word scores. The statement pretends to be true even when we work with the same language if we are always in the circle of the same linguistic forms.

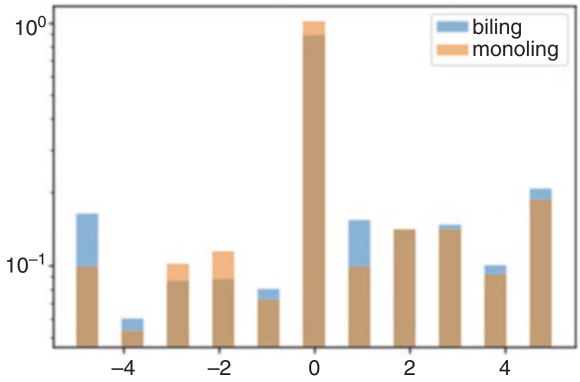



**Fig. 8** Histogram displaying in comparison the distribution of pooled Cube coordinates of texts derived from “Enjoyment (−5) – Distress (+5)”

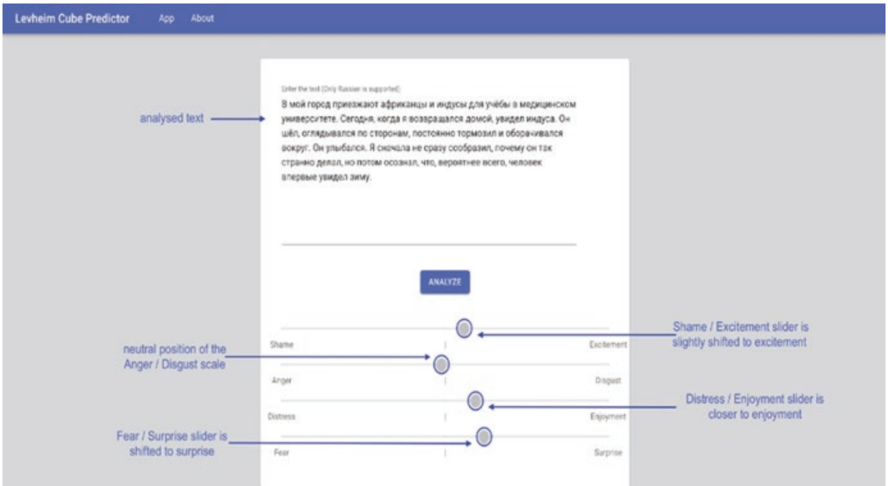



**Fig. 9** Emotion analyzer interface trained on the “Basic dataset” obtained from 2000 monolinguals

## 9 Further Research

As we have already run an analyzer of emotions in Internet-texts in Russian trained on our “Basic dataset” by using regression model of ML and it performs rather well (see its interface in Fig. 9), we consider the perspective of elaborating another analyzer – for predicting emotions in text in Russian for Tuvan-Russian bilinguals.

Even if we have not yet tested this annotated by bilinguals dataset in any ML model, we anticipate strong inconsistency of predictions done by already existing analyzer built up on dataset annotated by monolinguals and an analyzer trained on dataset annotated by Tuvan-Russian bilinguals.

## 10 Conclusion

The problem of annotation design and its quality is very important for further development of ETA paradigm. Our experiment pays a particular attention to the fact that annotators are not an abstracta – they are individuals belonging to a cultural, linguistic, and ethnic community and have different lifestyles and life experiences. That is why the concept of target-group of any emotion in text analyzer/classifier enters on the stage of projects in ETA. Not only high values of F1 score justify the project success but also the adequateness of its outputs to the estimations of target-clients.

**Acknowledgments** We want to thank the Krasnoyarsk Region Foundation for Research for supporting our research (project № 629 "Computer modeling of emotional perception of Internet texts in Russian by Tuvan-Russian bilinguals").

## References

1. Picard, R.W.: Affective Computing, 1st edn. MIT Press, Cambridge, MA (1997)
2. Koshkarova, N.N., Yakovleva, E.M.: Discourse of new emotionality: communicative practices of digital reality. Politicheskaya lingvistika. **5**, 147–152 (2019). https://doi.org/10.26170/pl19-05-15. (in Russian)
3. Abercrombie, G., Batista-Navarro, R.: Sentiment and position-taking analysis of parliamentary debates: a systematic literature review. J. Comput. Soc. Sci. **3**, 245–270 (2020). https://doi.org/10.1007/s42001-019-00060-w
4. Kim, E., Klinger, R.: A survey on sentiment and emotion analysis for computational literary studies. arXiv, 1808.03137. https://doi.org/10.17175/2019_008
5. Calvo, R.A., Kim, S.M.: Emotions in text: dimensional and categorical models. Comput. Intell. **29**, 527–543 (2013). https://doi.org/10.1111/j.1467-8640.2012.00456.x
6. Ekman, P.: An argument for basic emotions. Cognit. Emot. **6**(3–4), 169–200 (1992). https://doi.org/10.1080/02699939208411068
7. Tomkins, S.S.: Affect Imagery Consciousness. Volume 1: The Positive Affects. Springer, New York (1962)
8. Plutchik, R.: Emotions: a general psychoevolutionary theory. In: Scherer, K., Ekman, P. (eds.) Approaches to Emotion, pp. 197–219. Lawrence Erlbaum Associates, Hillsdale (1984). https://doi.org/10.4324/9781315798806
9. Alm, C.O., Rot, D., Sproat, R.: Emotions from text: machine learning for text-based emotion prediction. In: Raymond, J. (ed.) Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, pp. 579–586. Association for Computational Linguistics, Vancouver (2005)
10. Volkova, E., Mehler, B., Meurers, W.D., Gerdemann, D., Bülthoff, H.: Emotional perception of fairy tales: achieving agreement in emotion annotation of text. In: Inkpen, D., Strapparava, C. (eds.) Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pp. 98–106. Association for Computational Linguistics, Los Angeles (2010)
11. Bhowmick, P.K., Basu, A., Mitra, P.: Reader perspective emotion analysis in text through ensemble based multi-label classification framework. Comput. Inf. Sci. **2**(4), 64–74 (2009). https://doi.org/10.5539/cis.v2n4p64

12. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**(6), 1161–1178 (1980). https://doi.org/10.1037/h0077714
13. Osgood, C.E., May, W.H., Miron, M.S.: Cross-Cultural Universals of Affective Meaning. University of Illinois Press, Urbana (1975)
14. Lövheim, H.: A new three-dimensional model for emotions and monoamine neurotransmitters. Med. Hypotheses. **78**, 341–348 (2012). https://doi.org/10.1016/j.mehy.2011.11.016
15. Kolmogorova, A., Kalinin, A., Malikova, A.: Semiotic function of empathy in text emotion assessment. Biosemiotics. **14**(3) Cham: Springer (2021). https://doi.org/10.1007/s12304-021-09434-y
16. Chernigovskaya, T.V.: Cheshire Smile of Schrödinger's Cat: Language and Mind. Jazyki slavjanskoj kul'tury, Moscow (2017) (in Russian)
17. Polinsky, M.: Bilingual children and adult heritage speakers: the range of comparison. Int. J. Biling. **22**, 547–563 (2018). https://doi.org/10.1177/1367006916656048
18. Bialystok, E., Martin, M.M.: Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. Dev. Sci. **7**, 325–339 (2004)
19. Kroll, J.F., Stewart, E.: Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. J. Mem. Lang. **33**(2), 149–174 (1994). https://doi.org/10.1006/jmla.1994.1008
20. Harris, C.L., Gleason, J.B., Ayçiçeği, A.: Chapter 10 – When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In: Pavlenko, A. (ed.) Bilingual Minds: Emotional Experience, Expression, and Representation, pp. 257–283. Multilingual Matters, Clevedon (2006)
21. Kazanas, S.A., Altarriba, J.: Emotion word processing: effects of word type and valence in Spanish-English bilinguals. J. Psycholinguist. Res. **45**(2), 395–406 (2016). https://doi.org/10.1007/s10936-015-9357-3
22. Vélez-Uribe, I., Rosselli, M.: Electrophysiological correlates of emotion word processing in Spanish-English bilinguals. Biling. Lang. Congn. **24**, 31–55 (2021). https://doi.org/10.1017/S136672892000036X
23. Chen, P., Lin, J., Chen, B., Lu, C., Guo, T.: Processing emotional words in two languages with one brain: ERP and fMRI evidence from Chinese-English bilinguals. Cortex. **71**, 34–48 (2015). https://doi.org/10.1016/j.cortex.2015.06.002
24. Heras-Pedrosa, D.L.C., Sánchez-Núñez, P., Peláez, J.: Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. Int. J. Environ. Res. Public Health. **17**(15), 5542 (2020)
25. Datasets and trained models for Russian from Sberbank AI. https://huggingface.co/sberbank-ai
26. Borgoiakova, T.G.: Sociolinguistic Processes in the Republics of Southern Siberia. Khakas. gos. un-t im. N. F. Katanova, Abakan (2022) (in Russian)
27. Katunin, D.A.: Linguistic legislation in the republics of Siberia. Lang. Cult. **4**, 13–24 (2009) (in Russian)
28. Anayban, Z.V.: Interethnic situation and political processes in post-Soviet Tuva. Issledovaniya po prikladnoy i neotlozhnoy etnologii, p. 227. IEA RAN, Moscow (2011) (in Russian)
29. Arefiev, A.L., Bakhtikireeva, U.M., Sinyachkin, V.P.: Problems of bilingualism in the system of school language education in the Republic of Tuva. Novye issledovaniya Tuvy. **1**, 255–272 (2021) (in Russian)
30. Kalinin, A., Kolmogorova, A., Malikova, A.: Non-discrete sentiment dataset annotation: case study for Lövheim Cube emotional model. In: Alexandrov, D.A., Boukhanovsky, A.V., Chugunov, A.V., Kabanov, Y., Koltsova, O., Musabirov, I. (eds.) Digital Transformation and Global Society, DTGS 2020, Communications in Computer and Information Science, vol. 1242, pp. 154–164. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-65218-0_12
31. Kalinin, A., Kolmogorova, A., Malikova, A.: Non-discrete sentiment annotation for Lövheim Cube. Google Colab Notebook (2021) https://colab.research.google.com/drive/15_0q1ff7_2fuldHoC1qzh4U7av16qM1V

32. Birdsong, D., Gertken, L.M., Amengual, M.: Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. COERLL, University of Texas at Austin (2012) https://sites.la.utexas.edu/bilingual/
33. Kolmogorova, A., Kalinin, A., Malikova, A.: Lövheim cube-backed emotion analysis: from classification to regression. In: Alexandrov, D.A., et al. (eds.) Digital Transformation and Global Society. DTGS 2021. Communications in Computer and Information Science, vol. 1503, pp. 97–107. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93715-7_7