ISSN:2828-6235
Impact factor:13.77

International Journal of Literature and Language Studies(IJLLS)

CULTURAL APPROPRIATENESS OF MACHINE-GENERATED CONGRATULATIONS IN UZBEK AND RUSSIAN: A CORPUS-INFORMED EVALUATION FRAMEWORK

Yuldasheva Sabina Yusufovna
Lecturer, Department of Russian Language and Literature, Kokand State University, Kokand, Uzbekistan; ORCID: https://orcid.org/0009-0008-4685-4884; e-mail: yuldashevasabina18@gmail.com
doi:https://doi.org/10.5281/zenodo.23303736

Abstract. Large language models are increasingly used to write congratulatory messages, yet their output is usually judged by fluency rather than by conformity to the etiquette norms of a particular linguaculture. This paper proposes a corpus-informed framework for evaluating the cultural appropriateness of machine-generated congratulations in Uzbek and Russian. Norm profiles are derived from a reference corpus of authentic congratulatory texts [N texts] and describe, for each communicative situation, the expected form of address, level of formality, obligatory rhetorical moves, and typical set of wish concepts. Generated texts are compared with these profiles along six criteria, combining automatic measures with expert ratings. The framework is tested on [N] texts produced by [models] in response to prompts varying the occasion, the age and status of the addressee, and the channel. The study shows [main findings]. The framework contributes to the evaluation of culturally sensitive text generation for low-resource languages such as Uzbek.
Keywords: congratulation, speech etiquette, cultural appropriateness, large language models, text generation evaluation, politeness, Uzbek, Russian, low-resource languages.

1. Introduction

Writing a congratulation seems a trivial task, and it is precisely for such routine texts that people now turn to chatbots and language models. A user may ask a model to congratulate a grandmother on her birthday, a colleague on a promotion, or a teacher on Teachers' Day. The resulting text is usually grammatical and fluent. Whether it is appropriate is a different question: a congratulation that ignores the addressee's age, uses the wrong pronoun of address, or omits the blessings expected in a given culture may be perceived as cold or even disrespectful.

This problem is especially acute for Uzbek. Uzbek is a low-resource language in natural language processing: models are trained on far less Uzbek data than English or Russian data, and their knowledge of Uzbek communicative norms is often filtered through translation from dominant languages. As a result, a model may produce an Uzbek text that is a calque of an English or Russian template rather than a congratulation that an Uzbek speaker would actually write.

Existing metrics for text generation, such as similarity to reference texts or fluency scores, do not capture this dimension. What is needed is an evaluation grounded in empirical knowledge of how congratulations are actually written in a given linguaculture and situation. The present paper addresses this need.

The research questions are as follows. RQ1: Which parameters of the congratulation genre are critical for its cultural appropriateness in Uzbek and Russian? RQ2: How can norms for these parameters be derived from a corpus of authentic texts? RQ3: To what extent do machine-generated congratulations conform to these norms, and does conformity differ between Uzbek and Russian?

2. Related Work

Etiquette genres and politeness. Congratulation is a ritual genre whose core is an expressive speech act in J. Searle's sense, typically combined with wishes and compliments. Its realization

depends on the factors identified by P. Brown and S. Levinson: social distance, relative power, and the weight of the imposition. H. Spencer-Oatey extended this view with the concept of rapport management, emphasizing that politeness norms are culture-specific, and A. Wierzbicka showed that seemingly universal speech acts are shaped by culture-specific values. In Russian studies of speech etiquette, N. I. Formanovskaya described the system of etiquette formulas and their distribution across situations, while M. Hakimov examined pragmatic norms of Uzbek communication, including the central role of respect for elders.

Cultural awareness in NLP. Recent research has drawn attention to the cultural limitations of language technologies. P. Joshi and colleagues documented the sharp imbalance in resources between a few dominant languages and the majority of the world's languages. D. Hershcovich and colleagues argued that cross-cultural NLP must address not only language but also differences in values, common ground, and communicative norms. E. Bender and colleagues warned that large models reproduce the perspectives dominant in their training data. These works establish the problem but offer few genre-specific instruments for measuring cultural appropriateness.

Gap. No evaluation framework exists for etiquette genres in Uzbek, and comparative Uzbek–Russian studies of congratulation are lacking. This paper combines genre analysis, corpus linguistics, and generation evaluation to fill this gap.

3. Reference Corpus and Norm Profiles

The empirical basis of the framework is a reference corpus of authentic congratulations in Uzbek and Russian [N texts, N tokens], collected from greeting cards, official messages, public social media posts, and messenger texts shared with the authors' consent. All personal data were anonymized. Each text is described by situational metadata: occasion, channel, formality, the relationship between the participants, and the age and status of the addresser and the addressee.

From the corpus, a **norm profile** is derived for each type of situation. A profile is a set of empirically observed tendencies rather than prescriptive rules. For example, for the situation "birthday, adult child to elderly parent, Uzbek, informal channel", the profile specifies the expected form of address (a kinship term, often with a respect marker, as in Hurmatli onajon 'dear respected mother'), the polite pronoun Siz, the obligatory moves (address, congratulation, wish), and the most frequent wish concepts (health, longevity, often accompanied by a religious formula). For a Russian situation "birthday, friend to friend, messenger", the profile specifies a first name, the informal pronoun ty, a short text, and the concepts of happiness and success.

Each profile records not only the most frequent choice but also the range of acceptable variation and the choices that are absent or rare in the corpus. This distinction is important: a generated text should not be penalized for being creative, but it should be flagged if it uses a choice that the corpus suggests is inappropriate for the situation.

4. Evaluation Framework

The framework evaluates a generated congratulation along six criteria (Table 1). Criteria 1–5 are compared with the norm profile of the corresponding situation; criterion 6 is assessed by native-speaker experts.

Table 1. Criteria of cultural appropriateness

Criterion	Corpus-based norm	Typical violation
1. Form of address	Kinship term, name, or title with respect markers matching the addressee's age and status	First name or no address when writing to an elder
2. Pronoun and formality	Siz / vy or sen / ty according to the	Informal pronoun to an elder; official style in a message to a friend

	relationship; register matching the channel	
3. Obligatory moves	Congratulation and wish present; evaluation in official texts	Missing wish; missing appraisal of merits in an official message
4. Wish concepts	Concept set close to the situation profile	Absence of health and longevity wishes for an elderly Uzbek addressee
5. Religious and cultural formulas	Used where typical, absent where atypical	Religious formula in an official state message, or its absence where expected
6. Naturalness	Expert judgement: would a native speaker write this?	Calques of English or Russian templates in Uzbek texts

Criteria 1–5 are operationalized as features that can be annotated manually and, in later stages, detected automatically. For each criterion, a generated text receives a score of 2 (matches the profile), 1 (acceptable variant within the observed range), or 0 (choice absent or rare in the corpus for this situation). Criterion 6 is rated by three native-speaker experts on a five-point scale. Agreement between raters is measured with Krippendorff's α, which suits ordinal data and more than two raters.

An overall appropriateness index can be calculated as the mean of the normalized criterion scores. However, the profile of scores across criteria is more informative than the single index, since it shows exactly where a model departs from cultural norms.

5. Experimental Design

To test the framework, a set of prompts was constructed that systematically varies four factors: occasion (birthday, wedding, religious holiday, professional achievement), the addressee's age and status (elder relative, peer, superior, subordinate), channel (greeting card, messenger, official letter), and language (Uzbek, Russian). The prompts were formulated neutrally, without hints about the expected etiquette, for example: "Write a birthday congratulation to my grandmother".

Each prompt was submitted to [models and versions] [N] times, yielding [N] generated texts. The texts were annotated according to the criteria by two annotators following shared guidelines; disagreements were resolved through discussion. The scores of generated texts were then compared with those of authentic corpus texts written for the same situations, which serve as a baseline.

6. Results

[This section is to be completed with the experimental data. The planned analyses are outlined below.]

6.1. Overall appropriateness. Mean scores by model and language: [data]. The comparison with the authentic baseline shows how far each model is from human norms.

6.2. Profile of violations. Distribution of scores by criterion: [data, chart]. This analysis reveals which aspects of etiquette are most problematic for the models, for instance the form of address or the choice of wish concepts.

6.3. Uzbek versus Russian. Differences in scores between the two languages: [data, statistical test]. The hypothesis is that Uzbek texts deviate more from the norm profiles because of the limited Uzbek training data.

6.4. Situational factors. Dependence of scores on the addressee's age and status and on the channel: [data]. The hypothesis is that violations are most frequent in situations requiring elaborate expressions of respect.

7. Discussion

The framework shifts the evaluation of generated texts from “Is this text correct?” to “Is this text appropriate here?”. This shift is essential for etiquette genres, whose social function is to maintain relationships rather than to convey information. A grammatically flawless congratulation can fail precisely at this function.

The corpus basis is a key advantage of the approach. Instead of relying on the intuitions of individual experts, the norms are derived from authentic texts and can be updated as communicative practices change, for instance as messenger communication shortens and informalizes congratulations. At the same time, expert judgement remains necessary for naturalness, which cannot yet be reduced to formal features.

The results also have practical implications. Developers of chatbots and writing assistants can use norm profiles to adjust generation for different situations, for example by prompting the model about the addressee’s age and status. For language teaching, the profiles provide a clear description of how etiquette norms differ between Uzbek and Russian, which is useful for learners of Russian in Uzbekistan and for learners of Uzbek.

Limitations: [size and composition of the reference corpus, number of models tested, subjectivity of naturalness ratings — to be specified]. Future work includes automatic detection of criteria 1–5 and extension of the framework to other etiquette genres and other Turkic languages.

8. Conclusion

The paper has proposed a corpus-informed framework for evaluating the cultural appropriateness of machine-generated congratulations in Uzbek and Russian. Norm profiles derived from authentic texts make it possible to identify where generated texts depart from the etiquette expectations of a linguaculture. The framework offers a replicable instrument for culturally sensitive evaluation of text generation, which is particularly needed for low-resource languages such as Uzbek.

References

1. Bender E. M., Gebru T., McMillan-Major A., Shmitchell S. On the dangers of stochastic parrots: Can language models be too big? // Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. — 2021. — P. 610–623.
2. Brown P., Levinson S. C. Politeness: Some Universals in Language Usage. — Cambridge: Cambridge University Press, 1987.
3. Formanovskaya N. I. Rechevoy etiket i kul’tura obshcheniya [Speech etiquette and the culture of communication]. — Moscow: Vysshaya shkola, 1989.
4. Hakimov M. O‘zbek pragmalingvistikasi asoslari [Fundamentals of Uzbek pragmalinguistics]. — Tashkent: Akademnashr, 2013.
5. Hershcovich D. et al. Challenges and strategies in cross-cultural NLP // Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. — 2022. — P. 6997–7013.
6. Joshi P., Santy S., Budhiraja A., Bali K., Choudhury M. The state and fate of linguistic diversity and inclusion in the NLP world // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. — 2020. — P. 6282–6293.
7. Krippendorff K. Content Analysis: An Introduction to Its Methodology. — 2nd ed. — Thousand Oaks: Sage, 2004.
8. Searle J. R. A classification of illocutionary acts // Language in Society. — 1976. — Vol. 5, No. 1. — P. 1–23.
9. Spencer-Oatey H. (ed.) Culturally Speaking: Culture, Communication and Politeness Theory. — 2nd ed. — London: Continuum, 2008.

10. Wierzbicka A. Cross-Cultural Pragmatics: The Semantics of Human Interaction. — Berlin; New York: Mouton de Gruyter, 1991.