Open Language Data Initiative: Advancing Low-Resource Machine Translation for Karakalpak

Mukhammadsaid Mamasaidov
Tahrirchi
m.mamasaidov@tahrirchi.uz

Abror Shopulatov
Tahrirchi
a.shopolatov@tahrirchi.uz

Abstract

This study presents several contributions for the Karakalpak language: a FLORES+ devtest dataset translated to Karakalpak, parallel corpora for Uzbek-Karakalpak, Russian-Karakalpak and English-Karakalpak of 100,000 pairs each and open-sourced fine-tuned neural models for translation across these languages. Our experiments compare different model variants and training approaches, demonstrating improvements over existing baselines. This work, conducted as part of the Open Language Data Initiative (OLDI) shared task, aims to advance machine translation capabilities for Karakalpak and contribute to expanding linguistic diversity in NLP technologies.

1 Introduction

The Karakalpak language, a member of the Turkic language family, is primarily spoken in the Republic of Karakalpakstan, an autonomous region within Uzbekistan, Central Asia. Current estimates suggest a native speaker population to be around 900,000 individuals (Ethnologue, 2024). Linguistically, Karakalpak is an agglutinative language which belongs to the Kipchak branch of the Turkic language family and shares close affinities with Kazakh and Nogai (Berdimuratov and Dáwletov, 1979).

As a low-resource language, Karakalpak presents significant challenges in the field of natural language processing, particularly in machine translation. Major translation platforms such as Google Translate (Google, 2024) currently do not offer support for this language as of the time of writing this paper, underscoring the need for dedicated research and development in this area.

This study, conducted as part of the Open Language Data Initiative (OLDI) shared task, presents fine-tuned neural models for Karakalpak translation, a fine-tuned version the No Language Left Behind (NLLB) model (NLLB Team et al., 2022). In line with OLDI's goals of expanding language resources, we release several key contributions:

1. A FLORES+ devtest dataset (NLLB Team et al., 2022) translated to Karakalpak
2. Parallel corpora for Uzbek-Karakalpak, Russian-Karakalpak and English-Karakalpak of 100,000 pairs each [1]
3. Open-sourced fine-tuned neural models for translation across Uzbek, Russian, English and Karakalpak languages [2]
4. Scripts for Latin-Cyrillic transliteration for Karakalpak

Our research aims to advance the state of machine translation for Karakalpak, contributing to the broader OLDI objective of improving natural language processing capabilities for low-resource languages. This work demonstrates how shared tasks like OLDI can drive progress in expanding linguistic diversity in NLP technologies.

2 Related work

The field of machine translation for low-resource languages has experienced significant advancement with the advent of the No Language Left Behind (NLLB) model families. These innovative models demonstrate the capability to facilitate translation across more than 200 languages, leveraging extensive collections of online corpora.

Another notable multilingual translation model is MADLAD-400 (Kudugunta et al., 2024), which extends the capabilities of large language models to cover 400 languages, including many low-resource languages and Karakalpak. This model represents a significant step forward in expanding the reach of

[1] https://huggingface.co/datasets/tahrirchi/dilmash

[2] https://huggingface.co/collections/tahrirchi

Proceedings of the Ninth Conference on Machine Translation (WMT), pages 606–613
November 15–16, 2024. ©2024 Association for Computational Linguistics

English	**Karakalpak**
According to Japan's nuclear agency, radioactive caesium and iodine has been identified at the plant.	Yaponiya yadro agentligi maǵlıwmatlarına kóre, stanciyada radioaktiv ceziy hám yod bar ekenligi anıqlanǵan.
The result of plotting analysis will be posted to a public website.	Syujet analiziniń nátiyjesi ǵalabalıq veb-saytqa jaylastırıladı.
The station's web site describes the show as "old school radio theater with a new and outrageous geeky spin!"	Stanciya veb-saytında show "jańa hám ádettegiden basqasha ájáyıp aylandıratuǵın eski mektep radio teatrı!" dep táriyiplenedi.

Table 1: Examples from the FLORES+ dataset for English-Karakalpak language pair

machine translation to a broader range of linguistic communities.

In the specific context of Karakalpak machine translation, several notable efforts have been made. A prominent example is the Apertium platform (Forcada et al., 2011), a rule-based machine translation system designed for low-resource languages. Utilizing finite-state algebra and rule-based methodologies, Apertium has developed morphological analyzers and spell-checking tools for Karakalpak[3]. Furthermore, it has produced machine translation systems for language pairs for Uzbek-Karakalpak, Kazakh-Karakalpak, and Tatar-Karakalpak.

A recent contribution to the Karakalpak translation comes from the Turkic Interlingua (TIL) team (Mirzakhalov, 2021), who introduced a model specifically trained on Turkic languages and corpora, with Karakalpak included in its linguistic scope. This initiative not only enhances the translation capabilities for Karakalpak but also contributes to the broader landscape of Turkic language processing. Additionally, the team has made significant strides in corpus development, introducing parallel corpora for numerous Turkic language pairs, including those involving Karakalpak.

Moreover, a proprietary online translation service for Karakalpak exists at `https://from-to.uz/`. To provide a comprehensive evaluation of Karakalpak machine translation capabilities, we will assess this tool's performance using its API, comparing it with our proposed models. This comparison will offer insights into both open-source and commercial solutions for low-resource language translation.

To our best knowledge, these developments collectively represent important steps towards improving machine translation capabilities for Karakalpak and other low-resource languages within the Turkic language family.

As an additional benchmark, we will include Claude-3.5-sonnet, a commercial large language model (LLM) with multilingual capabilities. While not specifically designed for Karakalpak translation, Claude-3.5-sonnet represents the current state of general-purpose language models and can provide valuable insights into how well such models perform on low-resource language tasks.

3 Datasets

3.1 FLORES+ Devtest Dataset

This study introduces the Karakalpak FLORES+ devtest dataset, which comprises 1012 sentences translated from English to Karakalpak. The FLORES+ datasets, derived from Wikimedia content, have been widely employed in the evaluation of foundational models within the NLLB family.

This dataset was developed under the auspices of the Open Language Data Initiative (OLDI). Two annotators were responsible for the translation of a devtest split, with subsequent cross-verification to ensure accuracy. The Karakalpak translations adhere to the most recent iteration of the Latin script orthography (see Table 1).

The Karakalpak orthography has experienced multiple changes recently. Both Latin and Cyrillic scripts are utilized, with the Latin script, introduced in 1995, undergoing several revisions. Notable modifications occurred in 2009 and 2016, with the latter replacing digraphs with diacritic letters to overcome previous limitations. Conversion scripts for Cyrillic and older Latin versions to the current system are available on GitHub[4].

3.2 Training data

The training dataset comprises diverse parallel corpora sourced from multiple domains, including news articles, literary works, lexicographic resources, and educational materials. Specifically, the corpus encompasses on average across three languages:

[3] `https://github.com/apertium/apertium-uzb-kaa`

[4] `https://github.com/tahrirchi/kaa-scripts`

- 23% sentences from news sources
- 34% sentences from books (novels, non-fiction)
- 24% sentences from bilingual dictionaries
- 19% sentences from school textbooks
- 4,000 English-Karakalpak pairs from Gatitos Project (Jones et al., 2023)

In total, the dataset consists of 100,000 sentence pairs for Uzbek-Karakalpak, Russian-Karakalpak, and English-Karakalpak each, making 300,000 pairs in total. Since there were too few bitexts with English, we decided to create English-Karakalpak dataset by translating Russian sentences from the Russian-Karakalpak dataset to English using Claude 3.5 Sonnet (See Appendix A). To promote further research and development in this field, we have made these corpora publicly available.

3.3 Data Mining Process

For mining parallel sentences, we apply only local mining, when we are sure that parallel sentences are to be mined from the translations of the same book, document or article. For alignment, we use LaBSE embeddings, although Karakalpak is not a supported language in LaBSE. We found that due to similarities of Karakalpak to already included Uzbek and Kazakh languages, LaBSE performed well for aligning sentences so we skipped this step.

The sentence alignment method we use is identical to the one applied for Erzya, as described by (Dale, 2022). We utilize LaBSE (Language-agnostic BERT Sentence Embedding) (Feng et al., 2020) to generate embeddings for each sentence pair. To calculate the alignment score, we first determine the cosine similarity between these embeddings. We then adjust this similarity by multiplying it with a length ratio - specifically, the length of the shorter sentence divided by the length of the longer sentence.

Using dynamic programming, we identify the sequence of sentence pairs that maximizes the total similarity score. Finally, we apply a threshold to filter out low-scoring alignments.

4 Translation Experiments

4.1 Model Training

For our experiments, we utilized the nllb-200-distilled-600M model, which is a transformer-based neural machine translation model with an encoder-decoder architecture. This model comprises 12 layers and employs the following approach: the source and target languages are indicated by the first tokens of the encoder and decoder inputs, respectively. This architecture allows the model to process and translate between numerous language pairs. The training process for our experiments consisted of several key steps:

4.1.1 Tokenizer Preparation

Initially, we trained a SentencePiece (Kudo and Richardson, 2018) tokenizer on an expanded mono-corpus of approximately 300,000 Karakalpak sentences with a total of 16,000 vocabulary length. We decided to train a separate tokenizer because we hypothesized that the intial vocabulary of the NLLB model was not suited for Karakalpak, as there were some non-ASCII characters in the writing system (see Table 2). We also provide an evaluation of a model without training a separate tokenizer and compare the model's performance with and without additional trained tokens.

Á á Ǵ ǵ Í ı Ń ń Ó ó Ú ú

Table 2: Non-ASCII letters from Karakalpak Latin alphabet.

4.1.2 Vocabulary Expansion

Following tokenizer training, we augmented the model's vocabulary. This expansion resulted in a total of 269,399 tokens, representing an increase of 13,195 tokens from the original model configuration. We then resized the model's token embeddings and initialized the new embeddings by averaging the embeddings of their constituent subtokens from the original vocabulary.

4.1.3 Model Variants

We developed three distinct model variants to evaluate the impact of additional tokens and training data composition:

1. **dilmash-raw**[5]: This model was trained exclusively on a our own parallel corpus comprising 300,000 sentence pairs in Uzbek, Russian, and English on the original nllb-200-600M.

2. **dilmash**: Same as **dilmash-raw**, but it is a fine-tuned model with additional tokens which

[5] dilmash `[dil-mash]` *n. (from Karakalpak)* an oral interpreter

Model	en-kaa	ru-kaa	uz-kaa	kaa-en	kaa-ru	kaa-uz
madlad-400	2.68 / 22.48	2.01 / 19.93	1.31 / 16.81	28.42 / 53.06	16.95 / 41.12	10.26 / 38.75
apertium-uzb-kaa	-	-	12.26 / 42.27	-	-	5.61 / 35.82
google-from-kaz	-	-	-	20.95 / 44.63	13.55 / 36.91	-
google-from-uzb	-	-	-	21.40 / 45.50	13.78 / 37.64	-
nllb-200-600M-from-kaz	-	-	-	4.32 / 23.35	3.12 / 16.86	3.91 / 25.26
nllb-200-600M-from-uzb	-	-	-	8.89 / 32.26	5.82 / 26.33	4.83 / 29.68
from-to.uz	-	-	**20.18 / 53.22**	-	-	11.18 / 41.37
claude-3.5-sonnet	11.17 / 33.37	9.02 / 34.02	12.74 / 35.17	**37.06 / 61.41**	**25.70 / 51.23**	**22.38 / 54.71**
dilmash-raw	14.37 / **45.65**	11.41 / **42.99**	16.16 / 48.88	30.01 / 54.81	16.34 / 42.01	19.19 / 51.92
dilmash	12.31 / 42.22	10.72 / 40.29	16.13 / 48.42	28.75 / 53.70	15.69 / 41.58	18.52 / 51.03
dilmash-TIL	**15.02** / 45.43	**12.00** / 42.07	17.59 / 49.90	32.07 / 56.45	17.53 / 43.52	19.83 / 52.58

Table 3: Evaluation of several models on sacreBLEU/chrF++ across various language pairs with Karakalpak on FLORES+ devtest set.

were trained on a bigger Karakalpak monocorpus.

3. **dilmash-TIL**: This variant was trained on the same dataset and tokenizer configuration as the **dilmash**, but supplemented with a strategically sampled subset from the TIL corpus. The sampling strategy was as follows:
   - 20% of parallel datasets containing Uzbek or Kazakh
   - 5% of all other datasets in the TIL corpus

To maintain balance with the Karakalpak dataset, we imposed an upper limit of 300,000 sentence pairs on the TIL corpus sample for the **dilmash-TIL**. This constraint ensured that the Karakalpak data was not overwhelmed by the additional multilingual data, while still allowing for potential improvements in cross-lingual transfer and overall model performance.

With a batch size of 1024 and using the AdaFactor (Shazeer and Stern, 2018) optimizer, we trained each model variant for 3 epochs. We employed a learning rate of 1e-4 with a linear warmup over the first 10% steps, followed by a constant learning rate schedule. Weight decay was set to 0.01 to help prevent overfitting.

To maximize computational efficiency and enable training on larger batch sizes, we utilized DeepSpeed ZeRO Stage 3 (Rasley et al., 2020) for model parallelism across 16 GPUs. This configuration allowed us to effectively distribute the model parameters and optimize memory usage, facilitating faster training times.

4.2 Evaluation Metrics

To evaluate the performance of our translation models, we employ two widely used metrics in machine translation:

- sacreBLEU (Post, 2018)
- chrF++ (Popović, 2017)

sacreBLEU, a standardized BLEU implementation, calculates the similarity between the machine-generated translation and one or more reference translations based on n-gram precision. It addresses inconsistencies in tokenization and BLEU computation across different implementations. chrF++, an extension of the character n-gram F-score, computes the F-score of character n-grams and word unigrams, which is particularly useful for morphologically rich languages, like Karakalpak or Uzbek.

5 Results and Discussion

Our evaluation on the FLORES+ Karakalpak devtest reveals several interesting insights into the performance of various translation models. The results, presented in Table 3, demonstrate the effectiveness of our proposed models, dilmash, dilmash-raw, and dilmash-TIL, in comparison to existing approaches.

Notably, the dilmash-raw model, which was trained on the original nllb-200-600M without additional tokens, outperforms the dilmash model with expanded vocabulary in most language pairs. This result suggests that the initialization of new token embeddings may have introduced some challenges. Our hypothesis is that the new token embeddings weren't initialized optimally, and before the model could learn good values for them, they may have affected other model parameters. The limited amount of Karakalpak data alone might not have been sufficient for the model to fully compensate for this initial distortion.

The dilmash-TIL model, which incorporates additional multilingual data from the TIL corpus, consistently outperforms both dilmash and dilmash-raw across all language pairs. This improvement is particularly notable in the ***-kaa** directions, with gains of up to 2.71 BLEU points (en-kaa) compared to dilmash. These results underscore two important points: first, the potential of using related Turkic language data to enhance translation quality for low-resource languages like Karakalpak; and second, that the additional training data and epochs may have allowed the model to better utilize the expanded vocabulary, overcoming the initial challenges faced by the dilmash model. To provide a more qualitative assessment of our models' performance, we have included translation examples in Appendix B.

While expanding the vocabulary can potentially improve model performance, careful consideration must be given to the initialization of new embeddings and the amount of training data available. The success of the dilmash-TIL model suggests that incorporating data from related languages and allowing for longer training periods can help overcome these challenges, ultimately leading to improved translation quality.

Interestingly, the Claude-3.5-sonnet model demonstrates superior performance in the **kaa-*** directions, surpassing our models by a significant margin. This suggests that large language models may have a particular advantage in understanding content in low-resource languages, possibly due to their extensive pretraining on diverse multilingual data.

The performance of other baseline models provides additional context. Google Translate when treating Karakalpak as Kazakh or Uzbek, achieves respectable results but falls short of our models and Claude-3.5-sonnet. The NLLB-200-600M model, despite not being originally trained on Karakalpak, shows some ability to transfer knowledge when treating Karakalpak as Uzbek rather than Kazakh. This aligns with linguistic expectations, given the closer relationship between Karakalpak and Uzbek (both in linguistic similarity and writing scripts).

6 Conclusion

Our key contributions in this work include:

1. Creation of a FLORES+ devtest dataset for Karakalpak.

2. Development of parallel corpora for Uzbek-Karakalpak, Russian-Karakalpak, and English-Karakalpak, each containing 100,000 sentence pairs.

3. Open-sourcing of fine-tuned neural models for translation across Uzbek, Russian, English, and Karakalpak languages.

4. Open-sourcing of scripts for Latin-Cyrillic transliteration for Karakalpak.

Looking ahead, we plan to explore data augmentation techniques to further enhance our models' performance. One promising approach is to leverage the capabilities of Claude-3.5-sonnet for back-translation, potentially expanding our training data with high-quality synthetic examples.

Additionally, we aim to expand our dataset by mining more data from a wider range of books and sources. This will not only increase the volume of our training data but also improve its diversity, potentially leading to more robust and versatile translation models.

7 Limitations

While our study presents some advancements in Karakalpak machine translation, several limitations should be noted. First, the relatively small size of our dataset, despite being substantial for a low-resource language, may limit the model's ability to generalize across diverse domains and linguistic contexts. Second, the reliance on machine translation for creating the English-Karakalpak dataset introduces potential biases and errors that could affect model performance. Additionally, our evaluation is primarily based on automatic metrics, which may not fully capture the nuances of translation quality, particularly for a morphologically rich language like Karakalpak. Future work should address these limitations through expanded data collection, human evaluation, and more diverse testing scenarios.

8 Acknowledgements

We thank David Dalé for his valuable insights and guidance throughout the project. Our heartfelt appreciation goes to Najimova Perizad and Nurlan Pirjanov for their expertise and assistance with the Karakalpak language. We are also grateful to Atabek Murtazaev and Nurniyazov Ajiniyaz for their support and contributions to this work. Their collective efforts and knowledge have been instrumental in improving the quality of this work.

References

E. Berdimuratov and A. Dáwletov. 1979. *Házirgi qaraqalpaq tili*. Qaraqalpqastan, Uzbekistan SSR. Textbook for philology students at higher education institutions (in Cyrillic).

David Dale. 2022. The first neural machine translation system for the Erzya language. In *Proceedings of the first workshop on NLP applications to field linguistics*, pages 45–53, Gyeongju, Republic of Korea. International Conference on Computational Linguistics.

Ethnologue. 2024. Karakalpak. Accessed: 2024-07-25.

Fangxiaoyu Feng, Yinfei Yang, Daniel Cer, Naveen Arivazhagan, and Wei Wang. 2020. Language-agnostic bert sentence embedding. *arXiv preprint arXiv:2007.01852*.

Mikel L Forcada, Mireia Ginestí-Rosell, Jacob Nordfalk, Jim O'Regan, Sergio Ortiz-Rojas, Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, and Francis M Tyers. 2011. Apertium: a free/open-source platform for rule-based machine translation. *Machine translation*, 25:127–144.

Google. 2024. Google translate. `https://translate.google.com`. Accessed: July 25, 2024.

Alexander Jones, Isaac Caswell, Orhan Firat, and Ishank Saxena. 2023. Gatitos: Using a new multilingual lexicon for low-resource machine translation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 371–405.

Taku Kudo and John Richardson. 2018. Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. *arXiv preprint arXiv:1808.06226*.

Sneha Kudugunta, Isaac Caswell, Biao Zhang, Xavier Garcia, Derrick Xin, Aditya Kusupati, Romi Stella, Ankur Bapna, and Orhan Firat. 2024. Madlad-400: A multilingual and document-level large audited dataset. *Advances in Neural Information Processing Systems*, 36.

Jamshidbek Mirzakhalov. 2021. Turkic interlingua: a case study of machine translation in low-resource languages. Master's thesis, University of South Florida.

NLLB Team, Marta R. Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, Anna Sun, Skyler Wang, Guillaume Wenzek, Al Youngblood, Bapi Akula, Loic Barrault, Gabriel Mejia-Gonzalez, Prangthip Hansanti, John Hoffman, Semarley Jarrett, Kaushik Ram Sadagopan, Dirk Rowe, Shannon Spruit, Chau Tran, Pierre Andrews, Necip Fazil Ayan, Shruti Bhosale, Sergey Edunov, Angela Fan, Cynthia Gao, Vedanuj Goswami, Francisco Guzmán, Philipp Koehn, Alexandre Mourachko, Christophe Ropers, Safiyyah Saleem, Holger Schwenk, and Jeff Wang. 2022. No language left behind: Scaling human-centered machine translation.

Maja Popović. 2017. chrf++: words helping character n-grams. In *Proceedings of the second conference on machine translation*, pages 612–618.

Matt Post. 2018. A call for clarity in reporting BLEU scores. In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 186–191, Belgium, Brussels. Association for Computational Linguistics.

Jeff Rasley, Samyam Rajbhandari, Olatunji Ruwase, and Yuxiong He. 2020. Deepspeed: System optimizations enable training deep learning models with over 100 billion parameters. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 3505–3506.

Noam Shazeer and Mitchell Stern. 2018. Adafactor: Adaptive learning rates with sublinear memory cost. In *International Conference on Machine Learning*, pages 4596–4604. PMLR.

A Prompt for translating from Russian to English using Claude-3.5-sonnet

```
You are a professional translator specializing in Russian to English translations.
Your task is to translate the given Russian text into English with the highest level
of accuracy, preserving the original meaning and context. Use proper grammar,
punctuation, and idiomatic expressions appropriate for English speakers.
Do not include any additional explanations or commentary; provide only the translated text.

Russian: {sent}
English:
```

B Translation examples from dilmash-TIL

eng	**kaa**	**dilmash-til (eng → kaa)**
The winter can be deceptively chilly: temperatures rarely go below freezing, but the wind and humidity combine to make it feel colder than what the thermometer says.	Qıs aldamshı dárejede suwıq bolıwı múmkin: temperatura kemnen-kem jaǵdaylarda muzlawdan tómenge túsedi, bıraq samal hám ıǵallıq termometr aytqanınan kóre suwıqlaw seziliw ushın birlesedi.	Qıs júdá suwıq bolıwı múmkin: temperaturanıń buzılıw dárejesinen az ushırasıwı múmkin, biraq samal hám ıallıqtıń birigiwinen termometrdiń aytqanına qaraganda suwiq boladı.
kaa	**eng**	**dilmash-til (kaa → eng)**
Transport joq ekenligi qáweteri oyındı topar tárepdarlarısız jabıq esikler artında ótkeriwge májbúr bolıw múmkinshiligın asırdı.	Fears of lack of transportation raised the possibility that the game would be forced to play behind closed doors without the team's supporters.	Concern over the lack of transportation increased the possibility of being forced to hold the game behind closed doors without group supporters.
rus	**kaa**	**dilmash-til (rus → kaa)**
Также полезно носить кольцо (только не слишком дорогое на вид).	Júzik taǵıw da paydalı bolıp tabıladı (júdá qımbat kórinetuǵın júzik emes).	Sonday-aq, dóńgelek kiyiwde paydalı boladı (tek júdá qımbat kórinetuın emes).
kaa	**rus**	**dilmash-til (kaa → rus)**
Ullı piramida Firaun Xufudı ullılaw ushın jaratılǵan hám kóplegen kishi piramidalar, qábirler hám ibadatcxanalar Xufunıń hayalları hám shańaraq aǵzaların ullılaw ushın qurılǵan.	Великая Пирамида была создана в честь фараона Хеопса, а много меньших пирамид, гробниц и храмов были построены в честь жён Хеопаса и его родственников.	Большая пирамида была создана для восхваления фараона Хуфа, а многие небольшие пирамиды, могилы и храмы были построены для восхваления жен Хуфа и членов его семьи.
uzb	**kaa**	**dilmash-til (uzb → kaa)**
Ayrim atomlar turg'un bo'lmagan yadroga ega, bu esa ularning kichkina turtki bilan yoki turtkisiz parchalanishga moyilligini anglatadi.	Ayırım atomlar turaqsız yadrolarǵa iye, yaǵnıy olar azǵantay yamasa hesh qanday túrtki bolmaǵan halda bóleklenedi.	Ayırım atomlar turaqsız yadroga iye, bul bolsa olardıń kishi túrtki menen yaki túrtkisiz bóliniwine beyim ekenligi anlatadı.
kaa	**uzb**	**dilmash-til (kaa → uzb)**
Keshesi 150 den 200 ge shekem nusqalar tayarlandı, házirde "Dunlap broadsides" dep ataladı.	Tun davomida 150 dan 200 ga qadar nusxalar tayyorlandi, ular hozirda "Danlep yon zambaraklari" deb ataladi.	Kechasi 150 dan 200 gacha nusxalar tayyorlandi, hozirda "Dunlap broadsides" deb ataladi.

Table 4: Some translation examples of dilmash-TIL model on FLORES+ sentences.