

EROL: Transformer-Based Multilingual Translation for Northeastern Indian Language Pairs

Laishram Thoibisana Devi[✉] and Grace Lee

Nalanda University, India

thoiclare@gmail.com

Abstract

We present EROL, a multilingual neural machine translation system for four English–Indic language pairs: English↔Assamese, English↔Bodo, English↔Manipuri (Bengali script), and English↔Manipuri (Meitei Mayek). Our systems are built on the Transformer-Base architecture with 6 encoder and 6 decoder layers, a 512-dimensional model size, 2048-dimensional feed-forward networks, and 8 attention heads. We employ shared source and target embeddings, GELU activation, layer normalization, dropout regularization, and byte-pair encoding (BPE) tokenization. The models are trained using the standard sequence-to-sequence cross-entropy objective, providing a simple yet effective baseline for multilingual translation in low-resource settings. Experimental results demonstrate competitive performance across all translation directions, achieving BLEU scores of 25.10 for Assamese→English, 23.50 for Bodo→English, 22.47 for Manipuri (Bengali)→English, 24.71 for Manipuri (Meitei Mayek)→English, 15.31 for English→Assamese, 14.25 for English→Bodo, 10.21 for English→Manipuri (Bengali), and 3.96 for English→Manipuri (Meitei Mayek).

1 Introduction

Machine translation (MT) for low-resource languages remains one of the most challenging problems in natural language processing. While neural machine translation (NMT) has achieved remarkable quality for high-resource language pairs, its data-hungry nature makes it difficult to apply directly to languages with limited parallel corpora (Koehn and Knowles, 2017). This challenge is particularly acute for the languages of Northeastern India, a linguistically diverse region whose languages are severely under-represented in digital text resources and have only recently begun to receive sustained attention in the MT community.

In this paper, we describe the **EROL** submission to the WMT Shared Task on Low-Resource Indic Language Translation (Pakray et al., 2026). We participate in four language pairs covering three languages of Northeastern India in eight translation directions: English–Assamese (Eng↔Asm), English–Bodo (Eng↔Brx), English–Manipuri (Bengali Script) (Eng↔Mni), and English–Manipuri (Meitei Mayek) (Eng↔Mni-Mtei). Notably, the shared task treats the two orthographies of Manipuri - the Bengali script and the indigenous Meitei Mayek script - as two separate translation tasks, reflecting the real-world digital coexistence of both writing systems.

Our approach deliberately favors simplicity and computational efficiency. For each language pair, we train a single bidirectional Transformer-Base model (Gala et al., 2023; Vaswani et al., 2017) that handles both the English→Indic and Indic→English directions. Bidirectional training doubles the effective amount of training signal seen by the model and allows the encoder and decoder to share cross-lingual representations through tied embeddings over a joint BPE vocabulary (Sennrich et al., 2016). Despite using no external pretrained models and no additional monolingual data, our systems achieve competitive results across all eight directions, with BLEU scores of up to 25.10 for Assamese→English and chrF++ scores above 53 for all Indic→English directions.

The remainder of this paper is organized as follows. Section 2 reviews related work on low-resource Indic MT. Section 3 presents the participating language pairs and surveys the literature on their scripts and usage. Section 4 describes the shared task data and our preprocessing pipeline. Section 5 presents the model architecture, workflow, and training configuration. Section 6 reports and analyzes the official evaluation results, and Section 7 concludes.

2 Related Work

The Transformer architecture (Vaswani et al., 2017) is the de facto standard for NMT, but its performance degrades substantially in low-resource conditions, where careful choices of vocabulary size, regularization, and model capacity matter greatly (Sennrich and Zhang, 2019). Subword segmentation with byte-pair encoding (Sennrich et al., 2016) is particularly important for morphologically rich Indic languages, as it mitigates data sparsity at the lexical level.

For the specific languages considered in this work, a growing body of research has emerged around the WMT shared-task ecosystem. Singh et al. (2023b) developed a Transformer-based NMT system for the English-Manipuri (Bengali Script) pair at the WMT 2023 Low-Resource Indic Language Translation task, obtaining 22.75 and 26.92 BLEU for the English→Manipuri and Manipuri→English directions, respectively - results that our from-scratch systems are directly comparable to. In a complementary study, compared from-scratch Transformer training against transfer learning with the multilingual pretrained mT5 model for English-Manipuri Singh et al. (2023a), finding that a well-tuned conventional NMT model can outperform pretrained transfer in this low-resource regime, which motivates our own decision to train compact Transformer-Base models from scratch. Follow-up work at WMT 2024 extended English-Manipuri translation to the MultiIndic22MT (Singh et al., 2024) shared task and highlighted how conversational and general-domain test sets can yield sharply lower surface-metric scores for Manipuri, a phenomenon we also observe for our English→Manipuri (Meitei Mayek) system.

Beyond Manipuri, low-resource Indic MT research has explored multimodal signals, with CLIP-based selective visual grounding on top of strong Indic translation backbones (Kumar et al., 2025), and script-centric evaluation of large multilingual models such as IndicTrans2 (Gala et al., 2023) and ByT5 for the Ol Chiki script of Santali (Singh et al., 2025a), a setting closely analogous to our Manipuri (Meitei Mayek) task. Large multilingual models such as IndicTrans2 (Gala et al., 2023) and NLLB-200 (Team et al., 2022) cover several of our languages and represent a strong transfer-learning alternative; in contrast, our submission investigates how far a compact, from-pretrain bidi-

rectional Transformer can go using only the constrained shared-task data. Finally, our evaluation protocol, which combines the surface-level metrics BLEU, METEOR, and TER with the character-level chrF++ and the embedding-based BERTScore and COMET, follows recent multi-metric evaluation practice for Indic MT, which shows that no single automatic metric adequately characterizes translation quality for these languages.

Previous editions of the WMT Low-Resource Indic Language Translation shared task (Pakray et al., 2025) have established benchmarks for Assamese, Manipuri, and other languages of North-eastern India, and participating systems have explored techniques ranging from transliteration and back-translation to fine-tuning of large pretrained models.

3 Languages and Scripts of the Participating Pairs

This section surveys the four language pairs in our submission, with particular attention to the writing systems involved, since script choice directly affects tokenization, resource availability, and evaluation in low-resource MT.

Assamese: Assamese is an Eastern Indo-Aryan language and the major lingua franca of the Brahmaputra valley, spoken by approximately 15 million people according to the 2011 Census of India. It is the official language of the state of Assam and has been included in the Eighth Schedule of the Indian Constitution since its inception. Assamese is written in the Assamese script, a variant of the Eastern Nagari (Bengali-Assamese) script family that shares most of its character inventory with the Bengali script but differs in a small number of letters (Meetei et al., 2026), most notably the characters for *ra* and *wa*. As an abugida, the script encodes consonant-vowel sequences with inherent vowels, vowel diacritics, and a rich set of consonant conjuncts, all of which increase surface lexical sparsity and make subword segmentation essential for NMT. Among the four pairs in our submission, English-Assamese is the best resourced, and Assamese also enjoys comparatively better coverage in large multilingual corpora and models.

Bodo: Bodo (also Boro) is a Tibeto-Burman language of the Bodo-Garo group, spoken by roughly 1.5 million people, primarily in the Bodoland Territorial Region of Assam. It was added to the

Eighth Schedule of the Indian Constitution in 2003 and serves as an associate official language of Assam. Historically, Bodo has been written in several scripts, including the Assamese and Latin scripts; since the 1960s-70s, the Devanagari script has been officially adopted and is used in education, administration, and print media. Devanagari is well supported digitally, but the available Bodo text is scarce, and English-Bodo is the smallest pair in our submission, with only 15,215 training sentence pairs. Bodo morphology is agglutinative, with productive suffixation, which further increases type-token ratios and motivates our use of a joint BPE vocabulary.

Manipuri: Manipuri (Meiteilon) is a Tibeto-Burman language spoken by approximately 1.8 million people, chiefly in the state of Manipur, and was included in the Eighth Schedule of the Indian Constitution in 1992. It is a tonal, agglutinative language with rich verbal morphology. Its defining characteristic from an MT perspective is its *digraphia*: Manipuri is currently written in different scripts, its original indigenous script called Meitei Mayek (Meetei Mayek) and the Bengali script (Singh et al., 2026a,b, 2025b). The Bengali script was adopted in the early eighteenth century and dominated Manipuri writing, education, and publishing for nearly three centuries. A sustained revival movement in the twentieth century led to the formation of a Meetei Mayek advisory committee in 1973 and, eventually, to the official reintroduction of Meetei Mayek into the school curriculum of Manipur from the mid-2000s onward. Meitei Mayek was encoded in Unicode only in version 5.2 (2009), which partly explains why digital text in the indigenous script remains scarcer than Bengali-script Manipuri text.

The two orthographies are not trivially interconvertible: Meitei Mayek is a more phonemic alphasyllabary with a compact letter inventory, whereas the Bengali script, inherited from Indo-Aryan usage, distinguishes several characters that map to a single Meitei Mayek letter, so rule-based transliteration between the scripts is lossy in one direction. As a consequence, the MT literature has largely treated the two scripts as separate tasks: earlier WMT systems for English-Manipuri were built on Bengali-script data, while later shared tasks moved to Meitei Mayek and reported substantially lower surface-metric scores in the English→Manipuri direction. Analogous observations have been made

Pair	Script	Train	Dev
Eng-Assamese	Assamese	54,000	1,024
Eng-Bodo	Devanagari	15,216	1,024
Eng-Manipuri (Beng.)	Bengali	23,687	1,024
Eng-Manipuri (Mtei.)	Meitei Mayek	17,000	1,024
Total		109,899	4,092

Table 1: Parallel data statistics (sentence pairs) used for training and validation. Beng. denotes Bengali script and Mtei. denotes Meitei Mayek.

for other indigenous Indic scripts, such as Ol Chiki for Santali, where translation into the indigenous script is markedly harder than translation into English. Our shared task mirrors this two-script reality by including both English-Manipuri (Bengali Script) and English-Manipuri (Meitei Mayek) as independent pairs, and our results in Section 6 reproduce the characteristic asymmetry reported in this literature.

4 Task and Data

We use only the parallel corpora released by the shared task organizers, without any additional monolingual or external parallel data. Table 1 summarizes the corpus statistics. The training data ranges from 15,216 sentence pairs for English-Bodo to 54,000 sentence pairs for English-Assamese, placing all four pairs firmly in the low-resource regime. Each pair is accompanied by a development set of 1,024 sentence pairs, and the English side of the development data is shared across all pairs.

Preprocessing: For each language pair, we apply a uniform preprocessing pipeline. We first normalize Unicode representations and remove empty and duplicate sentence pairs. We then learn a *joint* BPE vocabulary (Sennrich et al., 2016) over the concatenated English and Indic sides of the training corpus of each pair. A joint vocabulary enables tied source and target embeddings and allows the model to share subword units, which is especially useful for named entities and code-mixed content.

Bidirectional training: Following common practice for low-resource multilingual NMT (Johnson et al., 2017), we construct the bidirectional training corpus for each pair by concatenating the original (Eng→X) data with its reversed (X→Eng) counterpart, prepending a target-language tag token to each source sentence to indicate the desired translation direction. This doubles the number of train-

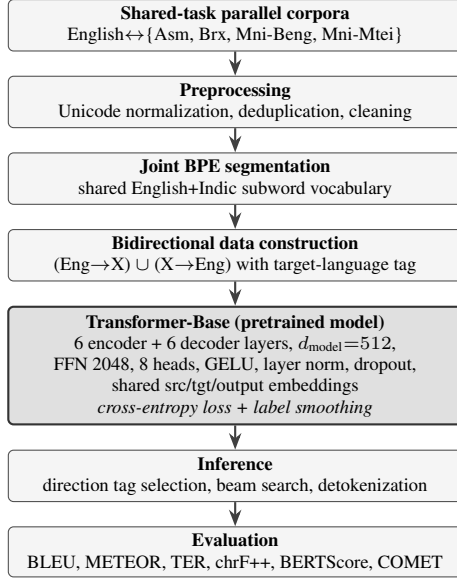


Figure 1: Architecture and workflow of the EROL translation systems.

ing examples each model sees and encourages the emergence of shared cross-lingual representations.

5 System Description

5.1 Workflow Overview

Figure 1 shows the end-to-end architecture and workflow of the EROL systems, from the raw shared-task corpora through preprocessing, joint BPE segmentation, and bidirectional data construction, to Transformer training, beam-search decoding, and multi-metric evaluation. The same pipeline is instantiated once per language pair, producing four independent bidirectional models.

5.2 Architecture

All EROL systems use the Transformer-Base architecture (Gala et al., 2023; Vaswani et al., 2017) with the configuration summarized in Table 2. Each model has 6 encoder and 6 decoder layers, a model dimension of 512, a feed-forward inner dimension of 2048, and 8 attention heads. We use GELU activations (Hendrycks and Gimpel, 2023) instead of ReLU, layer normalization, and dropout for regularization. Source embeddings, target embeddings, and the output projection are shared, which reduces the parameter count and acts as an additional regularizer in the low-resource setting (Press and Wolf, 2017).

Hyperparameter	Value
Encoder / decoder layers	6 / 6
Model dimension (d_{model})	512
Feed-forward dimension	2048
Attention heads	8
Activation	GELU
Embedding sharing	Src + Tgt + Output
Dropout	0.1
Label smoothing	0.1
Tokenization	Joint BPE
Optimizer	Adam
Learning rate schedule	Inverse sqrt, warmup
Beam size (inference)	5

Table 2: Model and training hyperparameters of the EROL systems.

5.3 Training

The models are trained with the standard sequence-to-sequence cross-entropy objective with label smoothing. We optimize with Adam (Kingma and Ba, 2017) using an inverse square-root learning-rate schedule with linear warmup, following Vaswani et al. (2017). Model selection is performed on the official development sets: we monitor validation loss and select the checkpoint with the best development performance for each pair.

5.4 Inference

At test time, the desired translation direction is selected via the target-language tag, and translations are generated with beam search. Generated subwords are merged back and detokenized before evaluation.

6 Results and Analysis

We report the official automatic evaluation results of our primary submissions in Table 3. Following multi-metric evaluation practice for Indic MT, systems are evaluated with BLEU (Papineni et al., 2002), METEOR (Banerjee and Lavie, 2005), TER (Snover et al., 2006), chrF++ (Popović, 2017), BERTScore (Zhang* et al., 2020), and COMET (Rei et al., 2020).

Indic→English: Translation into English is consistently the stronger direction, with BLEU scores ranging from 22.47 to 25.10 and chrF++ scores ranging from 53.04 to 55.39 across all four source languages. This is expected: English is the shared target across all bidirectional models and benefits from being the higher-resource side of every corpus, and n-gram metrics are more reliable for English than for morphologically rich Indic languages. Our

Direction	BLEU↑	METEOR↑	TER↓	chrF++↑	BERTScore↑	COMET↑
<i>Indic → English</i>						
Assamese→English	25.10	59.33	64.72	55.39	94.90	85.86
Bodo→English	23.50	58.14	66.75	54.02	94.55	81.87
Manipuri (Beng.)→English	22.47	56.60	69.41	53.04	94.27	80.15
Manipuri (Mtei.)→English	24.71	58.97	65.96	55.00	94.69	77.82
<i>English → Indic</i>						
English→Assamese	15.31	39.22	65.76	49.42	85.91	81.92
English→Bodo	14.25	35.31	75.40	48.68	84.47	61.38
English→Manipuri (Beng.)	10.21	18.40	85.13	38.83	84.42	67.04
English→Manipuri (Mtei.)	3.96	19.80	118.02	41.39	96.76	84.75

Table 3: Official automatic evaluation results of the EROL primary submissions on the shared-task test sets. Beng. = Bengali script; Mtei. = Meitei Mayek. ↑ indicates higher is better, and ↓ indicates lower is better.

Manipuri (Bengali Script)→English score of 22.47 BLEU is in the same range as the 26.92 BLEU reported by the WMT 2023 NITS-CNLP system for this direction (Singh et al., 2023b), despite our model being trained on a considerably smaller corpus. Assamese→English achieves the best scores overall (25.10 BLEU, 85.86 COMET), which correlates directly with English-Assamese having the largest training corpus (54,000 pairs). Interestingly, Manipuri (Meitei Mayek)→English translates into English nearly as well (24.71 BLEU) despite having far less training data (17,000 pairs), suggesting that the indigenous script provides a clean, unambiguous input representation.

English→Indic: Translation out of English is markedly harder, with BLEU ranging from 3.96 (Eng→Manipuri (Meitei Mayek)) to 15.31 (Eng→Assamese). The gap between the two directions reflects both genuine modeling difficulty - generating fluent text in a morphologically rich, low-resource language is harder than generating English - and a well-known evaluation artifact: word-level n-gram metrics such as BLEU and METEOR heavily penalize agglutinative morphology and script-level tokenization mismatches. This artifact is most visible for Eng→Manipuri (Meitei Mayek), where BLEU is only 3.96, yet the embedding-based metrics remain high (96.76 BERTScore, 84.75 COMET), the latter being the second-best COMET score among all English→Indic directions. Sharply reduced BLEU in the English→Manipuri direction has likewise been reported for Meitei Mayek output, supporting the interpretation that indigenous-script tokenization conventions, rather than translation quality alone, drive the low surface-metric scores. The character-level chrF++ score of 41.39 further supports the view that the translations are

of considerably better quality than BLEU alone suggests.

Effect of data size: Across both directions, ranking by translation quality broadly follows ranking by corpus size (Assamese > Manipuri (Bengali Script) ≈ Manipuri (Meitei Mayek) > Bodo in data; Assamese best in quality), but the correlation is not perfect: Bodo, the smallest corpus at 15,216 pairs, outperforms Manipuri (Bengali Script) in the Indic→English direction. This suggests that corpus quality, domain match, and script characteristics play a role at least as important as raw corpus size in this regime, echoing the finding of that careful configuration of a conventional Transformer can matter more than model scale or pretraining in low-resource English-Manipuri translation.

7 Conclusion

We presented the EROL submission to the WMT Shared Task on Low-Resource Indic Language Translation, covering eight translation directions between English and Assamese, Bodo, Manipuri (Bengali Script), and Manipuri (Meitei Mayek). Using a bidirectional Transformer-Base model per language pair, trained from a pretrained model on only the constrained shared-task data with joint BPE vocabularies and shared embeddings, we obtained competitive results in both directions, including 25.10 BLEU for Assamese→English. Our analysis highlights a substantial asymmetry between the Indic→English and English→Indic directions and shows that surface-level metrics can severely underestimate translation quality for Meitei Mayek script output, consistent with prior findings for Manipuri and other indigenous Indic scripts.

References

- Satanjeev Banerjee and Alon Lavie. 2005. [METEOR: An automatic metric for MT evaluation with improved correlation with human judgments](#). In *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, pages 65–72, Ann Arbor, Michigan. Association for Computational Linguistics.
- Jay Gala, Pranjal A Chitale, A K Raghavan, Varun Gumma, Sumanth Doddapaneni, Aswanth Kumar M, Janki Atul Nawale, Anupama Sujatha, Ratish Pudupully, Vivek Raghavan, Pratyush Kumar, Mitesh M Khapra, Raj Dabre, and Anoop Kunchukuttan. 2023. [Indictrans2: Towards high-quality and accessible machine translation models for all 22 scheduled indian languages](#). *Transactions on Machine Learning Research*.
- Dan Hendrycks and Kevin Gimpel. 2023. [Gaussian error linear units \(gelus\)](#).
- Melvin Johnson, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat, Fernanda Viégas, Martin Wattenberg, Greg Corrado, Macduff Hughes, and Jeffrey Dean. 2017. [Google’s multilingual neural machine translation system: Enabling zero-shot translation](#). *Transactions of the Association for Computational Linguistics*, 5:339–351.
- Diederik P. Kingma and Jimmy Ba. 2017. [Adam: A method for stochastic optimization](#).
- Philipp Koehn and Rebecca Knowles. 2017. [Six challenges for neural machine translation](#). In *Proceedings of the First Workshop on Neural Machine Translation*, pages 28–39, Vancouver. Association for Computational Linguistics.
- Deepak Kumar, Baban Gain, Kshetrimayum Boynao Singh, and Asif Ekbal. 2025. [Does vision still help? multimodal translation with CLIP-based image selection](#). In *Proceedings of the Twelfth Workshop on Asian Translation (WAT 2025)*, pages 115–123, Mumbai, India. Association for Computational Linguistics.
- Loitongbam Sanayai Meetei, Sujit Kumar Das, Kshetrimayum Boynao Singh, Utpal Barman, and Manoj Kumar Deka. 2026. A pragmatic comparison of nmt strategies for assamese: Transfer learning vs. from-scratch.
- Partha Pakray, Reddi Krishna, Santanu Pal, Advaita Vetagiri, Sandeep Dash, Arnab Kumar Maji, Saralin A. Lyngdoh, Lenin Laitonjam, Anupam Jamatia, Koj Sambyo, Ajit Das, and Riyanka Manna. 2025. [Findings of WMT 2025 shared task on low-resource Indic languages translation](#). In *Proceedings of the Tenth Conference on Machine Translation*, pages 532–553, Suzhou, China. Association for Computational Linguistics.
- Partha Pakray, Santanu Pal, Advaita Vetagiri, Kshetrimayum Boynao Singh, Sandeep Kumar Dash, Arnab Kumar Maji, Saralin A. Lyngdoh, Lenin Laitonjam, Anupam Jamatia, Ajit Das, Sunita Warjri, Uzzal Sharma, Koj Sambyo, and Riyanka Manna. 2026. Findings of WMT 2026 shared task on low-resource Indic languages translation. In *Proceedings of the Eleventh Conference on Machine Translation (WMT)*, Budapest, Hungary. Association for Computational Linguistics.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. [Bleu: a method for automatic evaluation of machine translation](#). In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.
- Maja Popović. 2017. [chrF++: words helping character n-grams](#). In *Proceedings of the Second Conference on Machine Translation*, pages 612–618, Copenhagen, Denmark. Association for Computational Linguistics.
- Ofir Press and Lior Wolf. 2017. [Using the output embedding to improve language models](#). In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 157–163, Valencia, Spain. Association for Computational Linguistics.
- Ricardo Rei, Craig Stewart, Ana C Farinha, and Alon Lavie. 2020. [COMET: A neural framework for MT evaluation](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2685–2702, Online. Association for Computational Linguistics.
- Rico Sennrich, Barry Haddow, and Alexandra Birch. 2016. [Neural machine translation of rare words with subword units](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1715–1725, Berlin, Germany. Association for Computational Linguistics.
- Rico Sennrich and Biao Zhang. 2019. [Revisiting low-resource neural machine translation: A case study](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 211–221, Florence, Italy. Association for Computational Linguistics.
- Kshetrimayum Boynao Singh, Asif Ekbal, and Partha Pakray. 2025a. [Evaluating IndicTrans2 and ByT5 for English–Santali machine translation using the Ol Chiki script](#). In *Proceedings of the 1st Workshop on Multimodal Models for Low-Resource Contexts and Social Impact (MMLoSo 2025)*, pages 95–100, Mumbai, India. Association for Computational Linguistics.
- Kshetrimayum Boynao Singh, Asif Ekbal, and Partha Pakray. 2026a. A bidirectional neural machine

- transliteration system for manipuri bengali and latin scripts.
- Kshetrimayum Boynao Singh, Kasar Horlung, Asif Ekbal, and Partha Pakray. 2026b. Byte-level neural machine translation for manipuri and tangkhul: Advancing low-resource language translation.
- Kshetrimayum Boynao Singh, Ningthoujam Avichandra Singh, Loitongbam Sanayai Meetei, Ningthoujam Justwant Singh, Thoudam Doren Singh, and Sivaji Bandyopadhyay. 2023a. [A comparative study of transformer and transfer learning MT models for English-Manipuri](#). In *Proceedings of the 20th International Conference on Natural Language Processing (ICON)*, pages 791–796, Goa University, Goa, India. NLP Association of India (NLP AI).
- Kshetrimayum Boynao Singh, Ningthoujam Avichandra Singh, Loitongbam Sanayai Meetei, Sivaji Bandyopadhyay, and Thoudam Doren Singh. 2023b. [NITS-CNLP low-resource neural machine translation systems of English-Manipuri language pair](#). In *Proceedings of the Eighth Conference on Machine Translation*, pages 967–971, Singapore. Association for Computational Linguistics.
- Ningthoujam Avichandra Singh, Boynao Kshetrimayum, Ningthoujam Justwant Singh, Thoudam Doren Singh, and Shilpa Sharma. 2025b. Quantifying the impact of data scale and quality on neural machine translation for low-resource language pair. In *2025 OITS International Conference on Information Technology (OCIT)*, pages 1–6. IEEE.
- Ningthoujam Justwant Singh, Kshetrimayum Boynao Singh, Ningthoujam Avichandra Singh, Sanjita Phijam, and Thoudam Doren Singh. 2024. [WMT24 system description for the MultiIndic22MT shared task on Manipuri language](#). In *Proceedings of the Ninth Conference on Machine Translation*, pages 797–803, Miami, Florida, USA. Association for Computational Linguistics.
- Matthew Snover, Bonnie Dorr, Rich Schwartz, Linnea Micciulla, and John Makhoul. 2006. [A study of translation edit rate with targeted human annotation](#). In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*, pages 223–231, Cambridge, Massachusetts, USA. Association for Machine Translation in the Americas.
- NLLB Team, Marta R. Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Hefernan, 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](#).
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. [Attention is all you need](#). In *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.
- Tianyi Zhang*, Varsha Kishore*, Felix Wu*, Kilian Q. Weinberger, and Yoav Artzi. 2020. [Bertscore: Evaluating text generation with bert](#). In *International Conference on Learning Representations*.