

# 《自然语言生成与机器翻译》 阅读推荐列表

## Machine Translation

### Statistics Machine Translation (SMT)

1. Knight K. A statistical MT tutorial workbook[C]//Prepared for the 1999 JHU Summer Workshop. 1999.
2. Och F J. Minimum error rate training in statistical machine translation[C]//Proceedings of the 41st annual meeting of the Association for Computational Linguistics (ACL' 03). 2003: 160-167.
3. Chiang D. A hierarchical phrase-based model for statistical machine translation[C]//Proceedings of the 43rd annual meeting of the association for computational linguistics (ACL' 05). 2005: 263-270.
4. Cho K, van Merriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2014: 1724-1734.

### Vanilla Neural Machine Translation (NMT)

1. Bahdanau D, Cho K H, Bengio Y. Neural machine translation by jointly learning to align and translate[C]//3rd International Conference on Learning Representations, ICLR 2015. 2015.
2. Gehring J, Auli M, Grangier D, et al. Convolutional sequence to sequence learning[C]//International Conference on Machine Learning. PMLR, 2017: 1243-1252.
3. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 6000-6010.

### Word / Subword

1. Lee J, Cho K, Hofmann T. Fully character-level neural machine translation without explicit segmentation[J]. Transactions of the Association for Computational Linguistics, 2017, 5: 365-378.
2. Sennrich R, Haddow B, Birch A. Neural Machine Translation of Rare Words with Subword Units[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016: 1715-1725.
3. Kudo T, Richardson J. SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing[C]//Proceedings of the 2018

Conference on Empirical Methods in Natural Language Processing: System Demonstrations. 2018: 66-71.

4. Kudo T. Subword Regularization: Improving Neural Network Translation Models with Multiple Subword Candidates[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018: 66-75.
5. Chen H, Huang S, Chiang D, et al. Combining character and word information in neural machine translation using a multi-level attention[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018: 1284-1293.
6. He X, Haffari G, Norouzi M. Dynamic Programming Encoding for Subword Segmentation in Neural Machine Translation[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 3042-3051.

## Syntax (or structure information)

1. Eriguchi A, Hashimoto K, Tsuruoka Y. Tree-to-Sequence Attentional Neural Machine Translation[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016: 823-833.
2. Chen H, Huang S, Chiang D, et al. Improved Neural Machine Translation with a Syntax-Aware Encoder and Decoder[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017: 1936-1945.
3. Zhou H, Tu Z, Huang S, et al. Chunk-Based Bi-Scale Decoder for Neural Machine Translation[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2017: 580-586.
4. Wang X, Tu Z, Wang L, et al. Self-Attention with Structural Position Representations[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 1403-1409.
5. Bugliarello E, Okazaki N. Enhancing Machine Translation with Dependency-Aware Self-Attention[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 1618-1627.

## Domain Adaptation

1. Luong M T, Manning C D. Stanford neural machine translation systems for spoken language domains[C]//Proceedings of the international workshop on spoken language translation. 2015: 76-79.
2. Freitag M, Al-Onaizan Y. Fast domain adaptation for neural machine translation[J]. arXiv preprint arXiv:1612.06897, 2016.

3. Chu C, Dabre R, Kurohashi S. An empirical comparison of domain adaptation methods for neural machine translation[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). 2017: 385-391.
4. Miceli-Barone A V, Haddow B, Hermann U, et al. Regularization techniques for fine-tuning in neural machine translation[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 1489-1494.
5. Kothur S S R, Knowles R, Koehn P. Document-level adaptation for neural machine translation[C]//Proceedings of the 2nd Workshop on Neural Machine Translation and Generation. 2018: 64-73.
6. Hu J, Xia M, Neubig G, et al. Domain Adaptation of Neural Machine Translation by Lexicon Induction[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019: 2989-3001.

## Document Translation

1. Longyue Wang, Zhaopeng Tu, Andy Way, and Qun Liu. 2017. Exploiting Cross-Sentence Context for Neural Machine Translation. In Proceedings of EMNLP 2017.
2. Rachel Bawden, Rico Sennrich, Alexandra Birch, and Barry Haddow. 2018. Evaluating Discourse Phenomena in Neural Machine Translation. In Proceedings of NAACL 2018.
3. Jörg Tiedemann, and Yves Scherrer. 2017. Neural Machine Translation with Extended Context. In Proceedings of the Third Workshop on Discourse in Machine Translation.
4. Elena Voita, Pavel Serdyukov, Rico Sennrich, and Ivan Titov. 2018. Context-Aware Neural Machine Translation Learns Anaphora Resolution. In Proceedings of ACL 2018.
5. Jiacheng Zhang, Huanbo Luan, Maosong Sun, Feifei Zhai, Jingfang Xu, Min Zhang and Yang Liu. 2018. Improving the Transformer Translation Model with Document-Level Context. In Proceedings of EMNLP 2018.
6. Zhaopeng Tu, Yang Liu, Shuming Shi, and Tong Zhang. 2018. Learning to Remember Translation History with a Continuous Cache. Transactions of the Association for Computational Linguistics.
7. Elena Voita, Rico Sennrich, and Ivan Titov. 2019. Context-Aware Monolingual Repair for Neural Machine Translation. In Proceedings of EMNLP 2019.
8. Zaixiang Zheng, Xiang Yue, Shujian Huang, Jiajun Chen, Alexandra Birch. 2020. Towards Making the Most of Context in Neural Machine Translation. In Proceedings of IJCAI 2020.

## Quality Estimation

1. Shah K, Avramidis E, Biçici E, et al. QuEst–design, implementation and extensions of a framework for machine translation quality estimation[J]. The Prague Bulletin of Mathematical Linguistics, 2013, 100(2013): 19-30.

2. Kim H, Jung H Y, Kwon H, et al. Predictor-estimator: Neural quality estimation based on target word prediction for machine translation[J]. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 2017, 17(1): 1-22.
3. Martins A F T, Junczys-Dowmunt M, Kepler F N, et al. Pushing the limits of translation quality estimation[J]. Transactions of the Association for Computational Linguistics, 2017, 5: 205-218.
4. Fan K, Wang J, Li B, et al. “Bilingual Expert” Can Find Translation Errors[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(01): 6367-6374.
5. Sun S, Guzmán F, Specia L. Are we Estimating or Guesstimating Translation Quality? [C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 6262-6267.
6. Kepler F, Trénous J, Treviso M, et al. OpenKiwi: An Open Source Framework for Quality Estimation[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations. 2019: 117-122.

## Latent Variable

1. Zhang B, Xiong D, Su J, et al. Variational Neural Machine Translation[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016: 521-530.
2. Shah H, Barber D. Generative neural machine translation[C]//Proceedings of the 32nd International Conference on Neural Information Processing Systems. 2018: 1353-1362.
3. Shen T, Ott M, Auli M, et al. Mixture models for diverse machine translation: Tricks of the trade[C]//International Conference on Machine Learning. PMLR, 2019: 5719-5728.
4. Zheng Z, Zhou H, Huang S, et al. Mirror-generative neural machine translation[C]//International Conference on Learning Representations. 2020.

## Non-autoregressive Translation (NAT)

1. Gu J, Bradbury J, Xiong C, et al. Non-autoregressive neural machine translation[C]//International Conference on Learning Representations. 2018.
2. Lee J, Mansimov E, Cho K. Deterministic Non-Autoregressive Neural Sequence Modeling by Iterative Refinement[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018: 1173-1182.
3. Ghazvininejad M, Levy O, Liu Y, et al. Mask-Predict: Parallel Decoding of Conditional Masked Language Models[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 6114-6123.
4. Kaiser L, Bengio S, Roy A, et al. Fast decoding in sequence models using discrete latent variables[C]//International Conference on Machine Learning. PMLR, 2018: 2390-2399.

5. Ma X, Zhou C, Li X, et al. FlowSeq: Non-Autoregressive Conditional Sequence Generation with Generative Flow[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 4273-4283.
6. Qian L, Zhou H, Bao Y, et al. Glancing transformer for non-autoregressive neural machine translation[J]. arXiv preprint arXiv:2008.07905, 2020.
7. Gu J, Kong X. Fully Non-autoregressive Neural Machine Translation: Tricks of the Trade[J]. arXiv preprint arXiv:2012.15833, 2020.

## Other

1. **UNMT:** Artetxe M, Labaka G, Agirre E, et al. Unsupervised neural machine translation. In Proceedings of the International Conference on Learning Representations, 2018.
2. **MERT:** Shen S, Cheng Y, He Z, et al. Minimum Risk Training for Neural Machine Translation[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016: 1683-1692.
3. Cheng S, Huang S, Chen H, et al. PRIMT: A pick-revise framework for interactive machine translation[C]//Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016: 1240-1249.
4. Weng R, Zhou H, Huang S, et al. Correct-and-Memorize: Learning to Translate from Interactive Revisions[C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19): 5255-5263.

## Natural Language Generation

### Summarization [[reference](#)]

1. Li C, Liu F, Weng F, et al. Document summarization via guided sentence compression[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. 2013: 490-500.
2. Rush A M, Chopra S, Weston J. A Neural Attention Model for Abstractive Sentence Summarization[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015: 379-389.
3. Cheng J, Lapata M. Neural Summarization by Extracting Sentences and Words[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016: 484-494.
4. Li P, Lam W, Bing L, et al. Deep Recurrent Generative Decoder for Abstractive Text Summarization[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. 2017: 2091-2100.



5. See A, Liu P J, Manning C D. Get To The Point: Summarization with Pointer-Generator Networks[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017: 1073-1083.
6. Paulus R, Xiong C, Socher R. A Deep Reinforced Model for Abstractive Summarization[C]//International Conference on Learning Representations. 2018.
7. Zhang Y, Merck D, Tsai E, et al. Optimizing the Factual Correctness of a Summary: A Study of Summarizing Radiology Reports[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 5108-5120.

## Paraphrase Detection/Generation [[reference](#)]

1. Huang P S, He X, Gao J, et al. Learning deep structured semantic models for web search using clickthrough data[C]//Proceedings of the 22nd ACM international conference on Information & Knowledge Management. 2013: 2333-2338.
2. Chen Q, Zhu X, Ling Z H, et al. Enhanced LSTM for Natural Language Inference[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017: 1657-1668.
3. Gupta A, Agarwal A, Singh P, et al. A deep generative framework for paraphrase generation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2018, 32(1).
4. Iyyer M, Wieting J, Gimpel K, et al. Adversarial Example Generation with Syntactically Controlled Paraphrase Networks[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers). 2018: 1875-1885.
5. Qian L, Qiu L, Zhang W, et al. Exploring diverse expressions for paraphrase generation[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 3164-3173.
6. Liu X, Mou L, Meng F, et al. Unsupervised Paraphrasing by Simulated Annealing[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020: 302-312.
7. Huang K H, Chang K W. Generating Syntactically Controlled Paraphrases without Using Annotated Parallel Pairs[J]. arXiv preprint arXiv:2101.10579, 2021.

## Style Transfer

1. Shen T, Lei T, Barzilay R, et al. Style transfer from non-parallel text by cross-alignment[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017: 6833-6844.
2. Fu Z, Tan X, Peng N, et al. Style transfer in text: Exploration and evaluation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2018, 32(1).

3. Yin D, Huang S, Dai X Y, et al. Utilizing Non-Parallel Text for Style Transfer by Making Partial Comparisons[C]//IJCAI. 2019: 5379-5386.
4. He J, Wang X, Neubig G, et al. A Probabilistic Formulation of Unsupervised Text Style Transfer[C]//International Conference on Learning Representations. 2019.

## Image Captioning

1. Vinyals O, Toshev A, Bengio S, et al. *Show and tell: A neural image caption generator*[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3156-3164.
2. Xu K, Ba J, Kiros R, et al. *Show, attend and tell: Neural image caption generation with visual attention*[C]//International conference on machine learning. PMLR, 2015: 2048-2057.
3. Rennie S J, Marcheret E, Mroueh Y, et al. *Self-critical sequence training for image captioning*[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 7008-7024.
4. Lu J, Xiong C, Parikh D, et al. *Knowing when to look: Adaptive attention via a visual sentinel for image captioning*[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 375-383.
5. Dai B, Fidler S, Urtasun R, et al. *Towards diverse and natural image descriptions via a conditional gan*[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2970-2979.
6. Anderson P, He X, Buehler C, et al. Bottom-up and top-down attention for image captioning and visual question answering[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 6077-6086.
7. Cornia M, Baraldi L, Cucchiara R. *Show, control and tell: A framework for generating controllable and grounded captions*[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 8307-8316.
8. Zhou L, Palangi H, Zhang L, et al. *Unified vision-language pre-training for image captioning and vqa*[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(07): 13041-13049.

## Table2text (or Data2text)

1. Jain P, Laha A, Sankaranarayanan K, et al. A Mixed Hierarchical Attention Based Encoder-Decoder Approach for Standard Table Summarization[C]//Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers). 2018: 622-627.
2. Song L, Zhang Y, Wang Z, et al. A Graph-to-Sequence Model for AMR-to-Text Generation[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018: 1616-1626.

3. Gong H, Feng X, Qin B, et al. Table-to-Text Generation with Effective Hierarchical Encoder on Three Dimensions (Row, Column and Time)[C]//Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019: 3134-3143.
4. Ye R, Shi W, Zhou H, et al. Variational Template Machine for Data-to-Text Generation[C]//International Conference on Learning Representations. 2020.
5. Xu K, Wu L, Wang Z, et al. SQL-to-Text Generation with Graph-to-Sequence Model[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018: 931-936.

## Other

1. **VAE:** Bowman S, Vilnis L, Vinyals O, et al. Generating Sentences from a Continuous Space[C]//Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning. 2016: 10-21.
2. **Sequence GAN:** Yu L, Zhang W, Wang J, et al. Seqgan: Sequence generative adversarial nets with policy gradient[C]//Proceedings of the AAAI conference on artificial intelligence. 2017, 31(1).
3. **CTRL:** Keskar N S, McCann B, Varshney L R, et al. Ctrl: A conditional transformer language model for controllable generation[J]. arXiv preprint arXiv:1909.05858, 2019.
4. **PPLM:** Dathathri S, Madotto A, Lan J, et al. Plug and Play Language Models: A Simple Approach to Controlled Text Generation[C]//International Conference on Learning Representations. 2020.
5. Zhao T, Lee K, Eskenazi M. Unsupervised Discrete Sentence Representation Learning for Interpretable Neural Dialog Generation[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2018: 1098-1107.

## Large Language Model (LLM)

### Pre-training

1. Language models are few-shot learners. Brown, Tom and Mann, Benjamin and Ryder, Nick and Subbiah, Melanie and Kaplan, Jared D and Dhariwal, Prafulla and Neelakantan, Arvind and Shyam, Pranav and Sastry, Girish and Askell, Amanda and others. Advances in neural information processing systems. 2020.
2. Opt: Open pre-trained transformer language models. Zhang, Susan and Roller, Stephen and Goyal, Naman and Artetxe, Mikel and Chen, Moya and Chen, Shuohui and Dewan, Christopher and Diab, Mona and Li, Xian and Lin, Xi Victoria and others. 2022.
3. mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer. Xue, Linting and Constant, Noah and Roberts, Adam and Kale, Mihir and Al-Rfou, Rami and Siddhant, Aditya



and Barua, Aditya and Raffel, Colin. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2021.

4. Few-shot Learning with Multilingual Generative Language Models. Lin, Xi Victoria and Mihaylov, Todor and Artetxe, Mikel and Wang, Tianlu and Chen, Shuohui and Simig, Daniel and Ott, Myle and Goyal, Naman and Bhosale, Shruti and Du, Jingfei and others. Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing. 2022.
5. Palm: Scaling language modeling with pathways. Chowdhery, Aakanksha and Narang, Sharan and Devlin, Jacob and Bosma, Maarten and Mishra, Gaurav and Roberts, Adam and Barham, Paul and Chung, Hyung Won and Sutton, Charles and Gehrmann, Sebastian and others. arXiv preprint arXiv:2204.02311. 2022.

## Instruct-tuning

1. Finetuned language models are zero-shot learners. Wei, Jason and Bosma, Maarten and Zhao, Vincent Y and Guu, Kelvin and Yu, Adams Wei and Lester, Brian and Du, Nan and Dai, Andrew M and Le, Quoc V. arXiv preprint arXiv:2109.01652. 2021.


## Alignment

1. ChatGPT: Optimizing Language Models for Dialogue. <https://openai.com/blog/chatgpt>. OpenAI. 2022.

## LLM for machine translation

1. Prompting PaLM for Translation: Assessing Strategies and Performance. Vilar, David and Freitag, Markus and Cherry, Colin and Luo, Jiaming and Ratnakar, Viresh and Foster, George. arXiv preprint arXiv:2211.09102. 2022.
2. In-context Examples Selection for Machine Translation. Agrawal, Sweta and Zhou, Chunting and Lewis, Mike and Zettlemoyer, Luke and Ghazvininejad, Marjan. arXiv preprint arXiv:2212.02437. 2022.
3. Adaptive Machine Translation with Large Language Models. Moslem, Yasmin and Haque, Rejwanul and Way, Andy. arXiv preprint arXiv:2301.13294.
4. A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity. Bang, Yejin and Cahyawijaya, Samuel and Lee, Nayeon and Dai, Wenliang and Su, Dan and Wilie, Bryan and Lovenia, Holy and Ji, Ziwei and Yu, Tiezheng and Chung, Willy and others. arXiv preprint. 2023.

# Reference:



NLG & MT--2020.txt

6.76KB

