

## DAFTAR PUSTAKA

- [1] K. P. K. R. dan Kebudayaan. Direktorat Kursus dan Pelatihan Direktorat Jenderal Pendidikan Vokasi, “622667104-2020-SKL-Bahasa-Inggris-English-for-Communication,” 2021, Accessed: Sep. 13, 2025. [Online]. Available: <https://www.scribd.com/document/622667104/2020-SKL-Bahasa-Inggris-English-for-Communication>
- [2] K. D. A. P. K. P. K. R. D. T. R. I. 2022 BADAN STANDAR, “Capaian Pembelajaran Mata Pelajaran Bahasa Inggris Fase A-Fase F,” Jakarta, 2022. Accessed: Sep. 13, 2025. [Online]. Available: <https://www.studocu.id/id/document/universitas-pgri-mahadewa-indonesia/english-education/cp-bahasa-inggris/64329629>
- [3] K. D. A. P. K. P. K. R. D. T. R. I. BADAN STANDAR, “Pembelajaran dan Asesmen,” Jakarta, 2025. Accessed: Sep. 13, 2025. [Online]. Available: <https://repositori.kemendikdasmen.go.id/33290/1/Panduan%20Pembelajaran%20Asesmen.pdf>
- [4] D. De-Fitero-Dominguez, E. Garcia-Lopez, A. Garcia-Cabot, J. A. Del-Hoyo-Gabaldon, and A. Moreno-Cediel, “Distractor Generation Through Text-to-Text Transformer Models,” *IEEE Access*, vol. 12, pp. 25580–25589, 2024, doi: 10.1109/ACCESS.2024.3361673.
- [5] G. Kurdi, J. Leo, B. Parsia, U. Sattler, and S. Al-Emari, “A Systematic Review of Automatic Question Generation for Educational Purposes,” *Int J Artif Intell Educ*, vol. 30, no. 1, pp. 121–204, Mar. 2020, doi: 10.1007/s40593-019-00186-y.
- [6] T. Andersson and P. Picazo-Sanchez, “Closing the Gap: Automated Distractor Generation in Japanese Language Testing,” *Educ Sci (Basel)*, vol. 13, no. 12, Dec. 2023, doi: 10.3390/educsci13121203.
- [7] H. Setiawan, I. Hidayah, and S. S. Kusumawardani, “Automatic Distractor Generation with Paradigmatic Relation for English Vocabulary Tests,” *Institute of Electrical and Electronics Engineers (IEEE)*, Jul. 2025, pp. 1–7. doi: 10.1109/siml65326.2025.11081171.
- [8] Y. Susanti, T. Tokunaga, H. Nishikawa, and H. Obari, “Automatic distractor generation for multiple-choice English vocabulary questions,” *Res Pract Technol Enhanc Learn*, vol. 13, no. 1, Dec. 2018, doi: 10.1186/s41039-018-0082-z.
- [9] H. F. Yusuf, I. Hidayah, and T. B. Adji, “Syntagmatic Distractor Generation for Vocabulary Multiple Choice Questions,” in *Proceedings - International Conference on Education and Technology, ICET*, Institute of Electrical and Electronics Engineers, 2023, pp. 94–99. doi: 10.1109/ICET59790.2023.10453362.
- [10] J. Hill and R. Simha, “Automatic generation of context-based fill-in-the-blank exercises using co-occurrence likelihoods and google n-grams,” in *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications, BEA 2016 at the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2016*, Association for Computational Linguistics (ACL), 2016, pp. 23–30. doi: 10.18653/v1/w16-0503.
- [11] A. Putri, S. Fauziati, and A. A. A. Rizqi, “Syntagmatic Distractor Generation for Multiple-Choice Language Tests: A Large Language Model-Based Approach,” *Institute of Electrical and Electronics Engineers (IEEE)*, Jul. 2025, pp. 617–622. doi: 10.1109/icera66156.2025.11087340.



- [12] Y. Hoshino, "Relationship between types of distractor and difficulty of multiple-choice vocabulary tests in sentential context," *Language Testing in Asia*, vol. 3, no. 1, Dec. 2013, doi: 10.1186/2229-0443-3-16.
- [13] P. Lewis *et al.*, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," Apr. 2021, [Online]. Available: <http://arxiv.org/abs/2005.11401>
- [14] Y. Susanti, R. Iida, and T. Tokunaga, "Automatic Generation of English Vocabulary Tests." [Online]. Available: <https://wordnet.princeton.edu/wordnet/documentation/>
- [15] Y. Susanti, T. Tokunaga, H. Nishikawa, and H. Obari, "Automatic distractor generation for multiple-choice English vocabulary questions," *Res Pract Technol Enhanc Learn*, vol. 13, no. 1, Dec. 2018, doi: 10.1186/s41039-018-0082-z.
- [16] H. F. Yusuf, I. Hidayah, and T. B. Adji, "Syntagmatic Distractor Generation for Vocabulary Multiple Choice Questions," in *Proceedings - International Conference on Education and Technology, ICET*, Institute of Electrical and Electronics Engineers, 2023, pp. 94–99. doi: 10.1109/ICET59790.2023.10453362.
- [17] Q. Wang, R. Rose, N. Orita, and A. Sugawara, "Automated Generation of Multiple-Choice Cloze Questions for Assessing {E}nglish Vocabulary Using {GPT}-turbo 3.5," *Proceedings of the Joint 3rd International Conference on Natural Language Processing for Digital Humanities and 8th International Workshop on Computational Linguistics for Uralic Languages*, pp. 52–61, 2023.
- [18] Y.-C. Liu and A.-Z. Yen, "ISSR: Iterative Selection with Self-Review for Vocabulary Test Distractor Generation," Jan. 2025, [Online]. Available: <http://arxiv.org/abs/2501.03462>
- [19] Q. Wang, R. Rose, N. Orita, and A. Sugawara, "Automated Generation of Multiple-Choice Cloze Questions for Assessing English Vocabulary Using GPT-turbo 3.5," 2024. [Online]. Available: <https://github.com/openai/openai-python>
- [20] M. Anastasia, H. Setiawan, Asmarini, T. B. Adji, and I. Hidayah, "Advancing Automatic Distractor Generation for English Vocabulary Tests Using Generative Artificial Intelligence," Institute of Electrical and Electronics Engineers (IEEE), Jul. 2025, pp. 1–7. doi: 10.1109/siml65326.2025.11081103.
- [21] R. Circi, J. Hicks, and E. Sikali, "Automatic item generation: foundations and machine learning-based approaches for assessments," 2023, *Frontiers Media S.A.* doi: 10.3389/feduc.2023.858273.
- [22] J. R. Bormuth, *Development of Readability Analysis: Final Report*, Project No. 7-0052, Contract No. OEC-3-7-070052-0326, U.S. Office of Education, Bureau of Research, Washington, DC, USA, Mar. 1969.
- [23] S. Panda, F. P. Gomez, M. Flor, and A. Rozovskaya, "Automatic Generation of Distractors for Fill-in-the-Blank Exercises with Round-Trip Neural Machine Translation." [Online]. Available: <https://www.esl-lounge.com/student/>
- [24] Y. Tomikawa, A. Suzuki, and M. Uto, "Adaptive Question-Answer Generation With Difficulty Control Using Item Response Theory and Pretrained Transformer Models," *IEEE Transactions on Learning Technologies*, vol. 17, pp. 2240–2252, 2024, doi: 10.1109/TLT.2024.3491801.
- [25] U. Lee *et al.*, "Few-shot is enough: exploring ChatGPT prompt engineering method for automatic question generation in english education," *Educ Inf Technol (Dordr)*, vol. 29, no. 9, pp. 11483–11515, Jun. 2024, doi: 10.1007/s10639-023-12249-8.
- [26] Z. Lin and H. Chen, "Investigating the capability of ChatGPT for generating multiple-choice reading comprehension items," *System*, vol. 123, Jul. 2024, doi: 10.1016/j.system.2024.103344.



- [27] U. Bezirhan and M. von Davier, "Automated reading passage generation with OpenAI's large language model," *Computers and Education: Artificial Intelligence*, vol. 5, Jan. 2023, doi: 10.1016/j.caeai.2023.100161.
- [28] Balai Bahasa Universitas Pendidikan Indonesia, *TOEFL PREPARATION COURSE*. Bandung: The Language Center, Indonesia University of Education, 2023.
- [29] Ismi Narulita Firda, Intan Azkiyah, and Albiansyah, "Testing Breadth and Depth of Vocabulary Knowledge and Their Relationship with Vocabulary Size of EFL Students," *JET (Journal of English Teaching)*, vol. 7, no. 1, pp. 89–100, Feb. 2021, doi: 10.33541/jet.v7i1.2434.
- [30] A. Chaer, "Pendekatan sintagmatik dan paradigmatis dalam kajian linguistik," *Jurnal Linguistik*, vol. 4, no. 2, pp. 45–58, 2007.
- [31] J. Hill and R. Simha, "Automatic Generation of Context-Based Fill-in-the-Blank Exercises Using Co-occurrence Likelihoods and Google n-grams."
- [32] N. Afzal and R. Mitkov, "Automatic generation of multiple-choice questions using dependency-based semantic relations," *Soft comput*, vol. 18, no. 7, pp. 1269–1281, 2014, doi: 10.1007/s00500-013-1141-4.
- [33] X. Du, M. Afzaal, and H. Al Fadda, "Collocation Use in EFL Learners' Writing Across Multiple Language Proficiencies: A Corpus-Driven Study," *Front Psychol*, vol. 13, Feb. 2022, doi: 10.3389/fpsyg.2022.752134.
- [34] D. Gablasova, V. Brezina, and T. McEnery, "Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence," Jun. 01, 2017, *Blackwell Publishing Ltd*. doi: 10.1111/lang.12225.
- [35] S. Dipper, "POS-Tagging of Historical Language Data: First Experiments." [Online]. Available: <http://www.llc.manchester.ac.uk/>
- [36] M. T. Machado and E. E. Seron Ruiz, "Evaluating large language models for the tasks of PoS tagging within the Universal Dependency framework." [Online]. Available: <https://openai.com/blog/openai-api>
- [37] M. Alekseeva, A. Myachykov, and Y. Shtyrov, "Inflectional zero morphology – Linguistic myth or neurocognitive reality?" *Front Psychol*, vol. 13, Dec. 2022, doi: 10.3389/fpsyg.2022.1015435.
- [38] K. Peter, "583 Peter K. Exploring Morphology an Overview and Current Perspectives," *Int J Anat Var*, vol. 17, no. 5, pp. 583–584, 2024, doi: 10.37532/1308-4038.17(5).399.
- [39] M. Hasan, "Transformers in Natural Language Processing," 2022, doi: 10.13140/RG.2.2.18062.84809.
- [40] Y. Chang *et al.*, "A Survey on Evaluation of Large Language Models," Dec. 2023, [Online]. Available: <http://arxiv.org/abs/2307.03109>
- [41] T. Xiao and J. Zhu, "Foundations of Large Language Models," Jun. 2025, [Online]. Available: <http://arxiv.org/abs/2501.09223>
- [42] W. X. Zhao *et al.*, "A Survey of Large Language Models," Mar. 2025, [Online]. Available: <http://arxiv.org/abs/2303.18223>
- [43] L. Zheng *et al.*, "Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena," Dec. 2023, [Online]. Available: <http://arxiv.org/abs/2306.05685>
- [44] J. Gu *et al.*, "A Survey on LLM-as-a-Judge," Oct. 2025, [Online]. Available: <http://arxiv.org/abs/2411.15594>
- [45] M. Son and S. Lee, "Performance Analysis of Prompt-Engineering Techniques for Large Language Model," in *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*, Institute of Electrical and Electronics Engineers Inc., 2025. doi: 10.1109/ICCE63647.2025.10930066.
- [46] K. Sampath Kini, S. Bose, B. Tyagi, P. Sharma, and A. Sharma, "Prompt Engineering: An efficient way for Content Creation using Prompt Chaining," in *2024 IEEE 4th*

- International Conference on ICT in Business Industry and Government, ICTBIG 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICTBIG64922.2024.10911772.
- [47] G. Beri and V. Srivastava, “Advanced Techniques in Prompt Engineering for Large Language Models: A Comprehensive Study,” in *2024 IEEE 4th International Conference on ICT in Business Industry and Government, ICTBIG 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ICTBIG64922.2024.10911672.
- [48] J. Wei *et al.*, “Finetuned Language Models Are Zero-Shot Learners,” Feb. 2022, [Online]. Available: <http://arxiv.org/abs/2109.01652>
- [49] T. B. Brown *et al.*, “Language Models are Few-Shot Learners,” Jul. 2020, [Online]. Available: <http://arxiv.org/abs/2005.14165>
- [50] M. R. Sudarshan and C. Kumbhar, “Comparative Analysis of Zero-shot, Few-shot, and One-shot Learning in Natural Language Processing,” in *2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation, ARIIA 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ARIIA63345.2024.11051380.
- [51] J. Kaplan *et al.*, “Scaling Laws for Neural Language Models,” Jan. 2020, [Online]. Available: <http://arxiv.org/abs/2001.08361>
- [52] H. Touvron *et al.*, “LLaMA: Open and Efficient Foundation Language Models,” Feb. 2023, [Online]. Available: <http://arxiv.org/abs/2302.13971>
- [53] P. Lewis *et al.*, “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,” Apr. 2021, [Online]. Available: <http://arxiv.org/abs/2005.11401>
- [54] Z. Jiang *et al.*, “Active Retrieval Augmented Generation,” Oct. 2023, [Online]. Available: <http://arxiv.org/abs/2305.06983>
- [55] N. Shinn, F. Cassano, E. Berman, A. Gopinath, K. Narasimhan, and S. Yao, “Reflexion: Language Agents with Verbal Reinforcement Learning,” Oct. 2023, [Online]. Available: <http://arxiv.org/abs/2303.11366>
- [56] Y. S. Kiyak, Ö. Coşkun, I. İ. Budakoğlu, and C. Uluoğlu, “ChatGPT for generating multiple-choice questions: Evidence on the use of artificial intelligence in automatic item generation for a rational pharmacotherapy exam,” *Eur J Clin Pharmacol*, vol. 80, no. 5, pp. 729–735, May 2024, doi: 10.1007/s00228-024-03649-x.
- [57] K. W. Church, “WORD ASSOCIATION NORMS, ] /UTUAL INFORMATION, AND LEXICOGRAPHY,” 1990.
- [58] M. E. Nahak, I. Hidayah, and T. B. Adji, “Automatic Distractor Generation with Syntagmatic Relation for Vocabulary Multiple Choice Questions using LLM,” in *2025 International Conference on Information Technology and Computing (ICITCOM)*, IEEE, Aug. 2025, pp. 362–367. doi: 10.1109/ICITCOM66635.2025.11265136.
- [59] S. Papageorgiou, R. J. Tannenbaum, B. Bridgeman, and Y. Cho, “The Association Between TOEFL iBT ® Test Scores and the Common European Framework of Reference (CEFR) Levels,” 2015.
- [60] J. B. Heaton, *Writing English Language Tests*, London, UK: Longman, 1988.