UNIVERSITAS
GADJAH MADA
**Analisis Perbandingan Performa Library Rhubarb dan Azure untuk Real-Time Lipsync pada Sistem Avatar Interaktif**
Shyra Athaya, Divi Galih Prasetyo Putri, S.Kom., M.Kom, Ph.D.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# DAFTAR PUSTAKA

[1] M. Mori, K. MacDorman, and N. Kageki, "The Uncanny Valley [From the Field]," *IEEE Robot. Autom. Mag.*, vol. 19, no. 2, pp. 98–100, Jun. 2012, doi: 10.1109/MRA.2012.2192811.

[2] C. G. Fisher, "Confusions Among Visually Perceived Consonants," *J. Speech Hear. Res.*, vol. 11, no. 4, pp. 796–804, Dec. 1968, doi: 10.1044/jshr.1104.796.

[3] P. Edwards, C. Landreth, E. Fiume, and K. Singh, "JALI: an animator-centric viseme model for expressive lip synchronization," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–11, Jul. 2016, doi: 10.1145/2897824.2925984.

[4] DanielSWolf, *Rhubarb Lip Sync*. (2025). [Online]. Available: https://github.com/DanielSWolf/rhubarb-lip-sync

[5] Microsoft, "Get Facial Position with Viseme – Azure AI Speech." Accessed: May 28, 2026. [Online]. Available: https://learn.microsoft.com/en-us/azure/ai-services/speech-service/how-to-speech-synthesis-viseme

[6] M. X. Sampaio *et al.*, "Evaluation of Automatic Speech Recognition Systems," in *Anais do XXXVI Simpósio Brasileiro de Banco de Dados (SBBD 2021)*, Brasil: Sociedade Brasileira de Computação - SBC, Oct. 2021, pp. 301–306. doi: 10.5753/sbbd.2021.17889.

[7] F. P. Lovely and A. Wicaksana, "Rule-based lip-syncing algorithm for virtual character in voice chatbot," *TELKOMNIKA Telecommun. Comput. Electron. Control*, vol. 19, no. 5, p. 1517, Oct. 2021, doi: 10.12928/telkomnika.v19i5.19824.

[8] A. Rachman, R. Hidayat, and H. A. Nugroho, "Pembandingan Tiga Nada Vokal /e/ untuk Animasi Gerak Bibir," *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 5, no. 2, p. 209, Aug. 2019, doi: 10.26418/jp.v5i2.32667.

[9] W. Mattheyses and W. Verhelst, "Audiovisual speech synthesis: An overview of the state-of-the-art," *Speech Commun.*, vol. 66, pp. 182–217, Feb. 2015, doi: 10.1016/j.specom.2014.11.001.

[10] H. L. Bear and R. Harvey, "Phoneme-to-viseme mappings: the good, the bad, and the ugly," *Speech Commun.*, vol. 95, pp. 40–67, Dec. 2017, doi: 10.1016/j.specom.2017.07.001.

UNIVERSITAS
GADJAH MADA
**Analisis Perbandingan Performa Library Rhubarb dan Azure untuk Real-Time Lipsync pada Sistem Avatar Interaktif**
Shyra Athaya, Divi Galih Prasetyo Putri, S.Kom., M.Kom, Ph.D.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

[11] N. Li, S. Liu, Y. Liu, S. Zhao, M. Liu, and M. Zhou, "Neural Speech Synthesis with Transformer Network," Jan. 30, 2019, *arXiv*: arXiv:1809.08895. doi: 10.48550/arXiv.1809.08895.

[12] R. Kannojia, A. K. Singh, I. Sharma, and S. Gupta, "Gen AI driven multilingual audio dubbing and synthesis system for cross-language video platforms," *Results Eng.*, vol. 27, p. 106241, Sep. 2025, doi: 10.1016/j.rineng.2025.106241.

[13] R. Fitriyanti and A. Aryanti, "STUDI LITERATUR MEAN OPINION SCORE MENGGUNAKAN MOVING PICTURE QUALITY METRICS (MPQM) DI JARINGAN LTE," 2018.

[14] R. Sekarjati and Palupi, "Presentasi Diri Online Pengguna Zepeto Melalui Avatar," Universitas Muhammadiyah Surakarta, 2023. [Online]. Available: https://eprints.ums.ac.id/id/eprint/113504

[15] Freepik, *3D illustration with online avatar*. [Digital illustration / PSD]. Available: https://www.freepik.es/psd-gratis/ilustracion-3d-avatar-linea_171612680.htm#fromView=keyword&page=1&position=0&uuid=46db497b-a7d4-406d-ad94-3fe7b6a899d0&query=Plantilla+avatar+3d

[16] A. S. E. Fathin, M. Suyanto, and A. F. Sofyan, "Penerapan Teknik Rotoscoping pada Karakter Animasi 2D Pembelajaran Makhraj Huruf Hijaiyah," *Creat. Inf. Technol. J.*, vol. 8, no. 1, p. 1, Mar. 2021, doi: 10.24076/citec.2021v8i1.262.

[17] Pinterest, *3D Avatar Illustration*. [Digital image]. Available: https://id.pinterest.com/pin/797911258952939355/

[18] A. Rachman, R. Hidayat, and H. A. Nugroho, "Pembandingan Tiga Nada Vokal /e/ untuk Animasi Gerak Bibir," *J. Edukasi Dan Penelit. Inform. JEPIN*, vol. 5, no. 2, p. 209, Aug. 2019, doi: 10.26418/jp.v5i2.32667.

[19] Roheen Qamar and Baqar Ali Zardari, "Artificial Neural Networks: An Overview," *Mesopotamian J. Comput. Sci.*, vol. 2023, pp. 124–133, Aug. 2023, doi: 10.58496/MJCSC/2023/015.

[20] V. Macketanz, E. Avramidis, A. Burchardt, J. Helcl, and A. Srivastava, "Machine Translation: Phrase-Based, Rule-Based and Neural Approaches with

UNIVERSITAS
GADJAH MADA
**Analisis Perbandingan Performa Library Rhubarb dan Azure untuk Real-Time Lipsync pada Sistem Avatar Interaktif**
Shyra Athaya, Divi Galih Prasetyo Putri, S.Kom., M.Kom, Ph.D.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

Linguistic Evaluation," *Cybern. Inf. Technol.*, vol. 17, no. 2, pp. 28–43, Jun. 2017, doi: 10.1515/cait-2017-0014.

[21] Blender Foundation, "Features — Blender." Blender, 2025. [Online]. Available: https://www.blender.org/features/

[22] R. S. Martin, D. Y. Dewanto, and U. M. Suryadarma, "PROTOTIPE KUNCI PINTU OTOMATIS MENGGUNAKANSENSOR KAMERA BERBASIS RASPBERRY," vol. 12, 2023.

[23] Ready Player Me, "Ready Player Me." [Online]. Available: https://readyplayer.me/

[24] Microsoft, "Azure Speech Text-to-Speech." Microsoft, 2025. [Online]. Available: https://ai.azure.com/catalog/models/Azure-Speech-Text-to-speech

[25] NVIDIA Corporation, "Grapheme-to-Phoneme Models," NVIDIA NeMo Framework User Guide. [Online]. Available: https://docs.nvidia.com/nemo-framework/user-guide/latest/nemotoolkit/tts/g2p.html

[26] Indonesian Online, "The International Phonetic Alphabet (IPA)." Accessed: May 28, 2026. [Online]. Available: https://indonesian-online.com/courses/mastering-indonesian-pronunciation/lesson/ipa/

[27] P. Roach, *English phonetics and phonology: a practical course*, 4. ed., Reprinted. Cambridge: Cambridge Univ. Press, 2010.

[28] K. L. Pike, *The intonation of American English*, 12th pr. in University of Michigan publications Linguistics, no. Vol. 1. Ann Arbor, Mich: Univ. of Michigan Press, 1972.

[29] Y. Wang, "Automatic Speech Recognition Model for Swedish using Kaldi".

[30] S. Pezzulli, M. G. Martini, and N. Barman, "Estimation of Quality Scores From Subjective Tests-Beyond Subjects' MOS," *IEEE Trans. Multimed.*, vol. 23, pp. 2505–2519, 2021, doi: 10.1109/TMM.2020.3013349.

[31] Arifin, S. Sumpeno, M. Hariadi, and A. M. Syarif, "Development of Indonesian Text-to-Audiovisual Synthesis System Using Syllable Concatenation Approach to Support Indonesian Learning," *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 2, pp. 166–184, 2017, doi: 10.3991/ijet.v12i02.6384.

[32] dalageo, "Random English Sentences." 2022. [Online]. Available:

UNIVERSITAS
GADJAH MADA
**Analisis Perbandingan Performa Library Rhubarb dan Azure untuk Real-Time Lipsync pada Sistem Avatar Interaktif**
Shyra Athaya, Divi Galih Prasetyo Putri, S.Kom., M.Kom, Ph.D.
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

https://www.kaggle.com/datasets/dalageo/random-english-sentences

[33] P. A. Mesiano, H. Innes-Brown, T. May, and J. Zaar, “A Method for Measuring Speech Intelligibility Using Connected Speech,” *J. Speech Lang. Hear. Res.*, vol. 69, no. 6, pp. 2874–2886, Jun. 2026, doi: 10.1044/2026_JSLHR-25-00235.