

6 Conclusion

In this paper, we described Lingua, an application that is capable of interpreting live speeches and creating synchronized dubbings for videos. It mitigates the shortcomings of cascaded STS systems by following an optional uploaded script, although it can revert to the default cascade if the speaker goes off script. We assessed the performance of the fuzzy matching classifier, and we found that it achieves F1 scores > 0.95 on a difficult test set. We discussed the primary hyperparameters and demonstrated how they affect the performance of the fuzzy matching algorithm. Finally, we proposed future avenues of research to improve the TTS component and prepare Lingua for deployment.

We anticipate that Lingua will increase accessibility to church multimedia content for underserved linguistic communities. It will decrease the time, cost, and expertise required to perform live interpretation and produce professional dubbings.

References

- Di Gangi, M., Rossenbach, N., Pérez, A., Bahar, P., Beck, E., Wilken, P., and Matusov, E. (2022). Automatic video dubbing at AppTek. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation*, pages 349–350.
- Federico, M., Enyedi, R., Barra-Chicote, R., Giri, R., Isik, U., Krishnaswamy, A., and Sawaf, H. (2020). From speech-to-speech translation to automatic dubbing. In *Proceedings of the 17th International Conference on Spoken Language Translation*, pages 257–264.
- Frankenberg-Garcia, A. (2009). Are translations longer than source texts. *A corpus-based study of explicitation* In: Beeby, A., Rodríguez P., & Sánchez-Gijón, P.(eds.) *Corpus use and learning to translate (CULT): An Introduction*. Amsterdam & Philadelphia: John Benjamins, pages 47–58.
- Google (2022). Aloud. <https://aloud.area120.google.com/>, Accessed 07/25/2022.
- Lakew, S. M., Federico, M., Wang, Y., Hoang, C., Virkar, Y., Barra-Chicote, R., and Enyedi, R. (2021). Machine translation verbosity control for automatic dubbing. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7538–7542. IEEE.
- Maestra (2021). Maestra Video Dubber. <https://maestrasuite.com/video-dubber>, Accessed 07/25/2022.
- Riccardi, A. (2005). On the evolution of interpreting strategies in simultaneous interpreting. *Meta: Journal des traducteurs/Meta: Translators' Journal*, 50(2):753–767.
- Sperber, M., Neubig, G., Niehues, J., and Waibel, A. (2019). Attention-passing models for robust and data-efficient end-to-end speech translation. *Transactions of the Association for Computational Linguistics*, 7:313–325.