

# “The Time of Man” is Synthetic: A Spoken-word Poetry Album in South American Spanish using Neural Networks

Paola Torres Núñez del Prado

Independent Researcher and Artist

<http://www.singingtextiles.com>

[autodios@gmail.com](mailto:autodios@gmail.com)

## Abstract

Machine Intelligence is an innovation that is already reshaping art [1] and society in general [2], thus, the decolonization of its algorithms becomes imperative, so not to reproduce the ethnic and cultural biases that prevail in contemporary society. By using a pre-trained Transformer-based language model (GPT2) retrained with poetry in Spanish, fine-tuned to examples of South American poetry, and then using this output to feed a text-to-speech synthesis system (Tacotron 2 + Waveglow, trained on an Argentinean voice data set and then on voice examples of Peruvian poet Jorge Eduardo Eielson) [3] I came up with a selection of spoken-word poems in a distinctly Latin American voice that ended up printed on a set of 4 7" lathe cut stereo vinyls albums.

## 1 Introduction

In the text called "Horrifying Sculpture" (set of instructions for an hypothetical artwork) [4] J.E. Eielson spoke of a talking doll made up of, among other things, a magnetic tape recorded with important poetic texts that will allow it to "continually recite the most beautiful poems conceived by man." With the support of Google's Artists + Machine Intelligence program, what I call the AIELSON (neologism mixing "A.I." and "Eielson") spoken-word poetry generation system ended up reciting poems that were taken as aesthetically pleasing for its audience. It is relevant to point out that this album is not a *deepfake* (fictitious audiovisual content made with machine learning tools), as it is not presented as if it were the late poet's work, but as an artificial entity that emulates the poet's voice when reciting. Intrigued by the possibility that new TTS systems offers regarding the recollection (picking-up) of stylistic qualities that could be considered typical of the region (South America), I intended to

as stress and rhythm, in this case, the particular way the poet had of speaking beyond his origin.

## 1.1 Poetry Generation

The 124M pre-trained GPT-2 English model was first fine-tuned on a selection of poetry in Spanish from various authors until it had learned how to properly generate text in Spanish. This retrained model was then fine-tuned with the poetic texts of Jorge Eduardo Eielson, Cecilia Vicuña (Chilean artist and poet) and my own poetry: the reason for this was that all three of us are from South America, visual artists and had written poetry, also having a shared interest in the pre-Columbian knot-based notation devices called quipus, in this way highlighting a contemporary branch of South American art, of particular Andean heritage.

## 1.2 Spoken-word Generation

For this section, two deep learning models were trained that consisted of the text-to-speech synthesis pair Tacotron 2 + Waveglow, first, trained from scratch with a Spanish voice, and then retrained (fine-tuned) with a very small data set of Jorge Eduardo Eielson's voice clips found online [5]. The total amount of time of these clips was about 20 minutes, hence the need for training the model beforehand with the larger M-AIILABS Spanish Argentinian data set. Not all the words from the generated speech are perfectly understandable, but neither were these in the sound files used for training, as most of the recordings used were from when the poet was advanced in years. This was kept as such intentionally in the tracks of the album: for our ears used to the linguistic perfection of computer-generated speech, an unclear voice can contradictorily feel familiar and human.

## 2 Conclusion

The GPT-2 generated unique poetry such as the following:

*Ciencia Oye al sol del cielo.*

*Ya todo es ciencia.*

*Haz que se me abandona.*

*No más a qué es tu noche que no es más, tu ciencia.*

*Necesitada, oye.*

The album's name, "The Time of Man" was generated by the system, and was printed as a 7" lathe cut vinyl record (4 in total), keeping all the irregularities and errors produced by the Tacotron 2 + Waveglow system, which output is still far from perfect, especially with such a small data set as the one used for this album. Nevertheless, according to the people that have heard the album [6], the audio sounds convincingly human and the voice with a distinct accent from the city of Lima, Peru. The tracks can be heard online on Bandcamp [7].

### 3 Ethical Implications

Using the voice of a person, without his or her authorization, can be highly controversial. For this project, I have received the Eielson Estate representative's approval, Martha Canfield, as the poet could obviously not give consent.

The speech generated has never been presented as that of Jorge Eduardo Eielson to the public (as noted above): the intention is not the achieving of a *deepfake*, but to expose and highlight through art the fact that most Machine Learning models that involve human language/speech are mainly trained in English, and in lesser degrees, in other languages (if any). This mirrors the hegemonic structure where English has been imposed as the *lingua franca* of the modern globalized world, while also taking into account the social conflicts linked to the language, for example, within the United States, where speaking Spanish can even lead to the speaker being harassed (if spoken in public places in certain areas of the country) due to prejudice.

The album produced is probably the first of its kind beyond this linguistic concerns, nevertheless, it references both an artistic tradition and a way of speaking that is deeply rooted in South American heritage.

### Acknowledgments

This project was funded and initially developed as part of Google's Artists + Machine Intelligence 2020 residency program. I would like to thank Holly Grimm who helped me with the training of the Machine Learning models used in the generation of both text and speech required for the creation of this album.

### References

- [1] Agüera y Arcas, B. (2017) *Art in the Age of Machine Intelligence*. In *Arts Open Access Journal*, <https://doi.org/10.3390/arts6040018>.
- [2] Vera Cubas, R. (2017) *Un lugar para ningún objeto: Las Esculturas subterráneas de J. E. Eielson*, Meier Ramirez Editores
- [3] Zabjek, A. (2018) *How artificial intelligence is reshaping our lives* In *phys.org*, <https://phys.org/news/2018-04-artificial-intelligence-reshaping.html>.
- [4] Herrera, A., Castellanos, W. (2016) *Jorge Eduardo Eielson: On the Other Side of Languages* In *Aluma Art Foundation Exhibitions*, <http://www.alunartfoundation.com/jorge-eduardo-eielson-on-the-other-side-of-languages/>.
- [5] *Various web sources*
- [6] *Opinions informally recollected from diverse audiences*
- [7] <https://khipumancer.bandcamp.com/album/el-tiempo-del-hombre>