

A Vocal Learning Model: Development And Analysis

Silvia Pagliarini, Arthur Leblois, Xavier Hinaut

▶ To cite this version:

Silvia Pagliarini, Arthur Leblois, Xavier Hinaut. A Vocal Learning Model: Development And Analysis. IMOL 2019 - Fourth International Workshop on Intrinsically Motivated Open-ended Learning, Jul 2019, Frankfurt, Germany. hal-02413140

HAL Id: hal-02413140 https://inria.hal.science/hal-02413140

Submitted on 16 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



output.

K

E

Y

W

Ο

R

D

S

A Vocal Learning Model: Development And Analysis

Silvia Pagliarini^{1,2,3}, Arthur Leblois³, Xavier Hinaut^{1,2,3}

Mnemosyne, Inria Bordeaux Sud-Ouest, UMR 5800, CNR, Bordeaux INP, Talence, France

- LaBRI, UMR 5800, CNRS, University of Bordeaux, Talence, France
- IMN, UMR 5293, CNRS, University of Bordeaux, Bordeaux, France 3)





(which cheode the goar of the rearining).	Da Cunha et al.,2010	Zebra finches 25 60 Day 90	O Sparrows Spirry Propring Provide Autumn Propring Provide Brainard and Doupe,2002
SENSORIMOTOR MODEL COMPONENTS [1]			
I) Architecture II) Learning III) Exploration			OEC: modelled as a Helmholtz oscillator



Motor control model

- Respiratory system and vocal organs.
- Anatomical properties and small size of birds make the investigation of vocal fold mechanisms difficult.
- The vocal output is driven by a complex gesture-dependent control scheme, and the brain does not control each motor control parameter independently [2].

Perceptual space

- How the brain encodes sensory stimulus.
- Highly nonlinear dynamics.
- Low dimensional representation of the sensory space.



Glottis

MOTOR CONTROL MODEL

AIM: learn the semantic modes in high-dimensional audio signals, and build a latent space useful for exploration. In addition, have sound production in the model.

WaveGAN [3]

PERCEPTUAL SPACE

AIM: have a lower dimensional representation of the physical acoustic space, which is the sensory space (sound).

Deep scattering transform [6]

• From MFCC coefficients introducing wavelets. • Scattering coefficients can be seen as a convolutional neural networks.

- Two player minimax game: generator VS discriminator [4].
- Inspired by DCGAN architecture [5].







GENERATED SOUND AFTER TRAINING

- Batch size = 64
- Epochs = 517

Inception Score (IS) = 1.95 ± 0.02



2nd Principal Component

PERSPECTIVES

Motor control model

Ep. 1662

- Explore the possibility to use a dynamical system model.
- More exaustive evaluation of the generative modle.

Exploration strategy

- Intrinsic motivation to drive motor exploration.
- Investigate the geometrical properties of latent space generated with the GAN.

BIBLIOGRAPHY

[1] Oudeyer, P. Y. (2005). The self-organization of speech sounds. Journal of Theoretical Biology, 233(3), 435-449. [2] Elemans, C. P. H., et al. Universal mechanisms of sound production and control in birds and mammals. Nature communications, 2015, 6: 8978.

[3] Donahue, C., McAuley, J., & Puckette, M. (2019). Adversarial Audio Synthesis. ICLR.

[4] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).

[5] Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

[6] Andén, J., & Mallat, S. (2014). Deep scattering spectrum. IEEE Transactions on Signal Processing, 62(16), 4114-4128.

AKNOWLEDGMENT: this work was supported by the Inria CORDI-S PhD fellowship grant.