Language emergence with a dyslexic receiver

Inès Ben Haj Kacem, Hugo Ninou and Constantin Vaillant-Tenzer

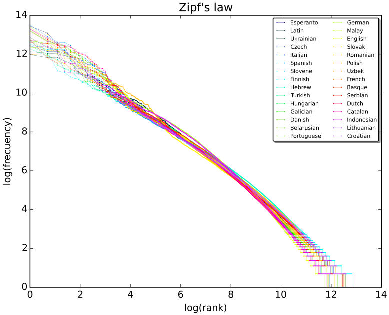
École Normale Supérieure - PSL

April 22, 2022

Presentation of the model

We considered two communication games where a sender chooses a sequence of symbols to send to a receiver. The operators are trained to achieve a successful communication :

- Egg (Chaabouni et al. 2019) : the emergent language does not follow Zipf's law, as all natural languages do.
- Lazimpa (Rita, Chaabouni, and Dupoux 2020), a derivative from Egg: the sender is lazy and the receiver is impatient. With these human-like constraints, the new emergent language respects Zipf's law.



Idea of our project

- More human-like operator : **dyslexia in the receiver**

Idea of our project

- More human-like operator : **dyslexia in the receiver**
- How the overall performance is impacted and how the emergent language differs from before ?

Idea of our project

- More human-like operator : **dyslexia in the receiver**
- How the overall performance is impacted and how the emergent language differs from before ?

Hypothesis

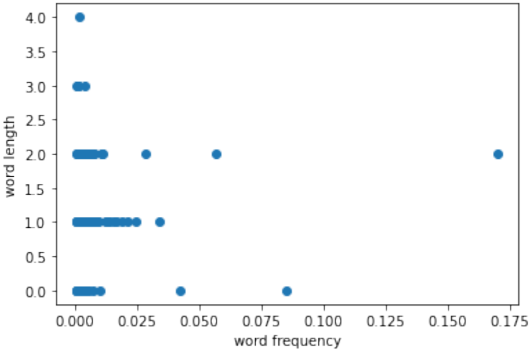
The emergent language will adapt to the dyslexia, with longer words and more repetitions, to allow the agent to communicate successfully.

The agents will be able to communicate, up to a certain permutation frequency.

Anti-efficient coding

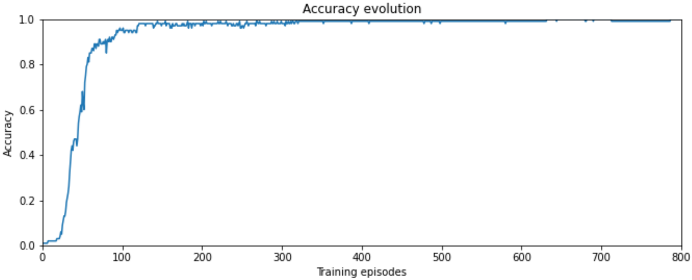
We failed to reproduce the main result of the first paper : the anti-efficient coding of the emergent language.

- Smaller order of magnitude of word length than in the article.
- Qualitatively different result than in the paper.



Lazimpa - accuracy

We reproduced the results of the Lazimpa communication game.



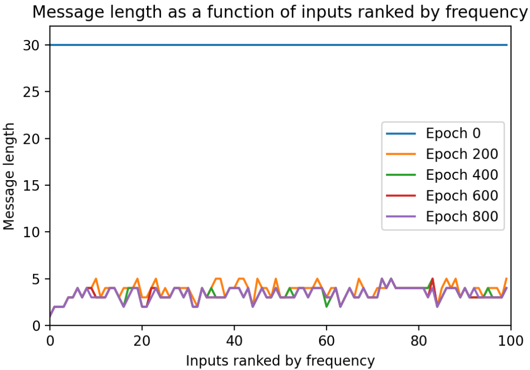
- Increasing accuracy across training.

Lazimpa - informative position

The average position of informative symbols in the message is around 3.5, just like in the paper.

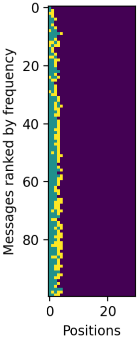
Lazimpa - length and frequencies

The message length decreases with the frequency according to Zip's law.



Lazimpa - useful symbols

Most of the symbols used are useful (in green) in the basic Lazimpa communication way.



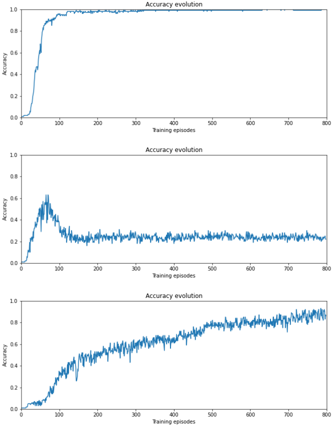
Two dyslexic conditions

Systematic permutation : Every message undergoes a permutation of arbitrary scope during training.

Random permutation : The probability of the permutation happening is proportional to the length of the message.

Dyslexic receiver: Results

Accuracy evolution over training. **High**: Reproduction of the paper's results. **Middle**: Systematic permutation. **Low**: Proportional probability permutation



Similar results in literature

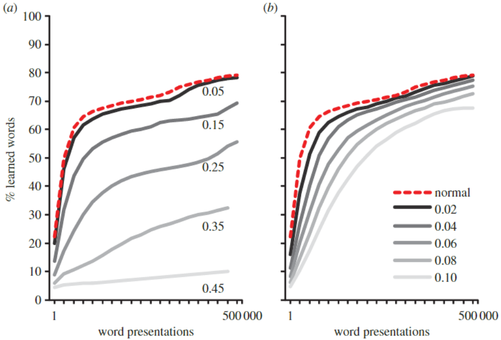


Figure 5. Learning to read with phoneme or visual deficits. (*a*) Phoneme deficits were simulated by changing a correctly assembled phoneme with a phonetically similar but incorrect phoneme with a certain probability (0.05, 0.15, 0.25, 0.35 and 0.45). (*b*) Visual deficits were simulated by switching a letter with the letter next to it with a certain probability (0.02, 0.04, 0.06, 0.08 and 0.10). The dotted line represents the unimpaired network. All simulations were run with a word recognition threshold of 0.15. (Online version in colour.)

Figure: Results obtained by Ziegler et al. Ziegler, Perry, and Zorzi 2014 when performing a computational investigation of reading development thanks to a connectionist model.

Successful message length

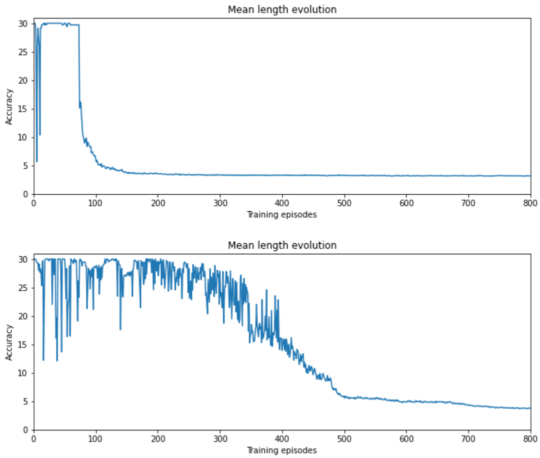


Figure: Evolution of the words' length over 800 training epochs. **High**: Systematic permutation. **Low**: Random permutation

Length and frequency

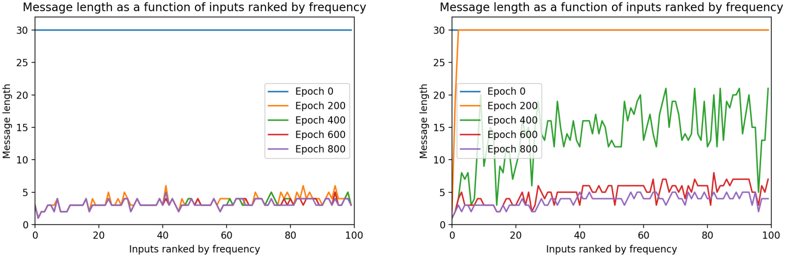
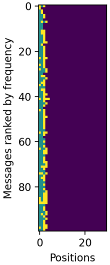


Figure: Message length ranked by frequency for different training epochs. **Left**: Systematic permutation. **Right**: Random permutation

Position of informative symbols



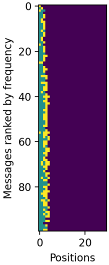


Figure: Position of the informative symbols after training. The green : informative, yellow : non informative symbols.**Left**: Systematic permutation. **Right**: Random permutation.

Average position of the informative symbols

Figure: Average position of the informative symbols after training. **Left**: Systematic permutation. **Right**: Random permutation.

- Longer words for systematic dyslexia

Final words

- Words are longer with more useless symbols and the system takes more time to converge with systematic dyslexia ;
- Two qualitatively different behavior depending of the frequency of the dyslexia : perfect communication or random guess.
- Possible follow-up : Study of the threshold in the permutation frequency between the two possible outcomes.

References

Chaabouni, Rahma et al. (2019). “Anti-efficient encoding in emergent communication”. In: *Advances in Neural Information Processing Systems* 32.

Rita, Mathieu, Rahma Chaabouni, and Emmanuel Dupoux (2020). “" LazImpa": Lazy and Impatient neural agents learn to communicate efficiently”. In: *arXiv preprint arXiv:2010.01878.*

Ziegler, Johannes C., Conrad Perry, and Marco Zorzi (2014). “Modelling reading development through phonological decoding and self-teaching: implications for dyslexia”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, p. 20120397. (Visited on 04/22/2022).

Questions