

Language emergence as representation learning

Marco Baroni
RepL4NLP 2019



Universitat
Pompeu Fabra
Barcelona



Facebook AI Research

The FAIR Paris EViL team



Emmanuel Dupoux



Rahma Chaabouni

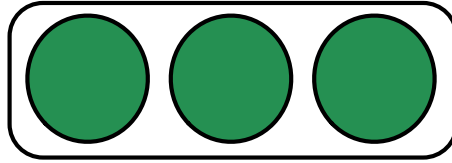


Diane Bouchacourt



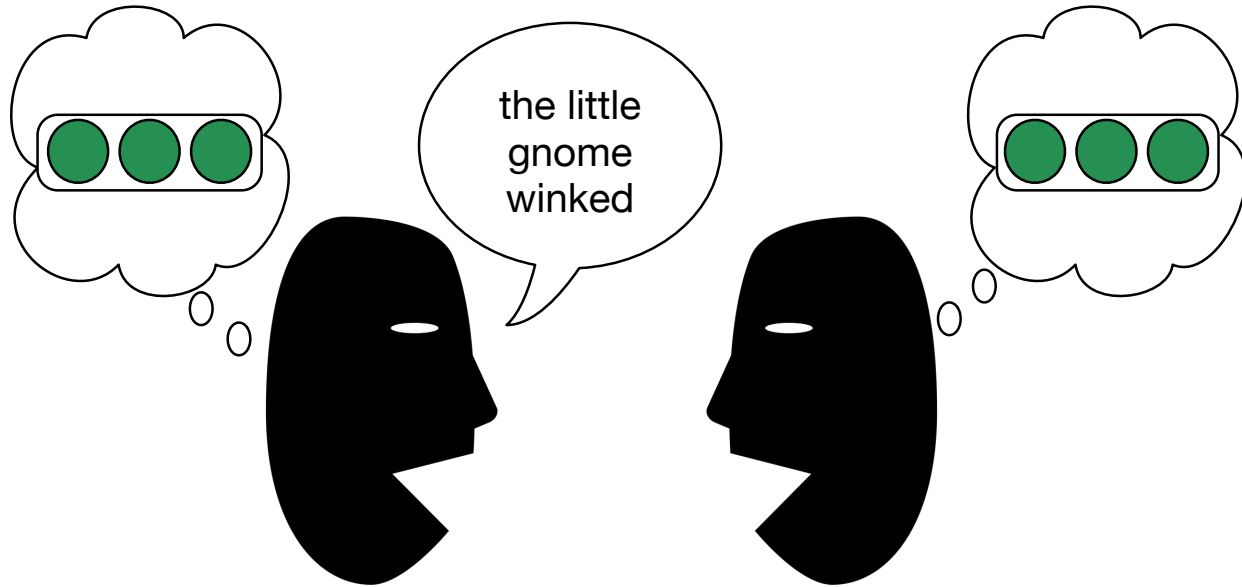
Eugene Kharitonov

Representation learning for language



the little gnome winked

Language as representation

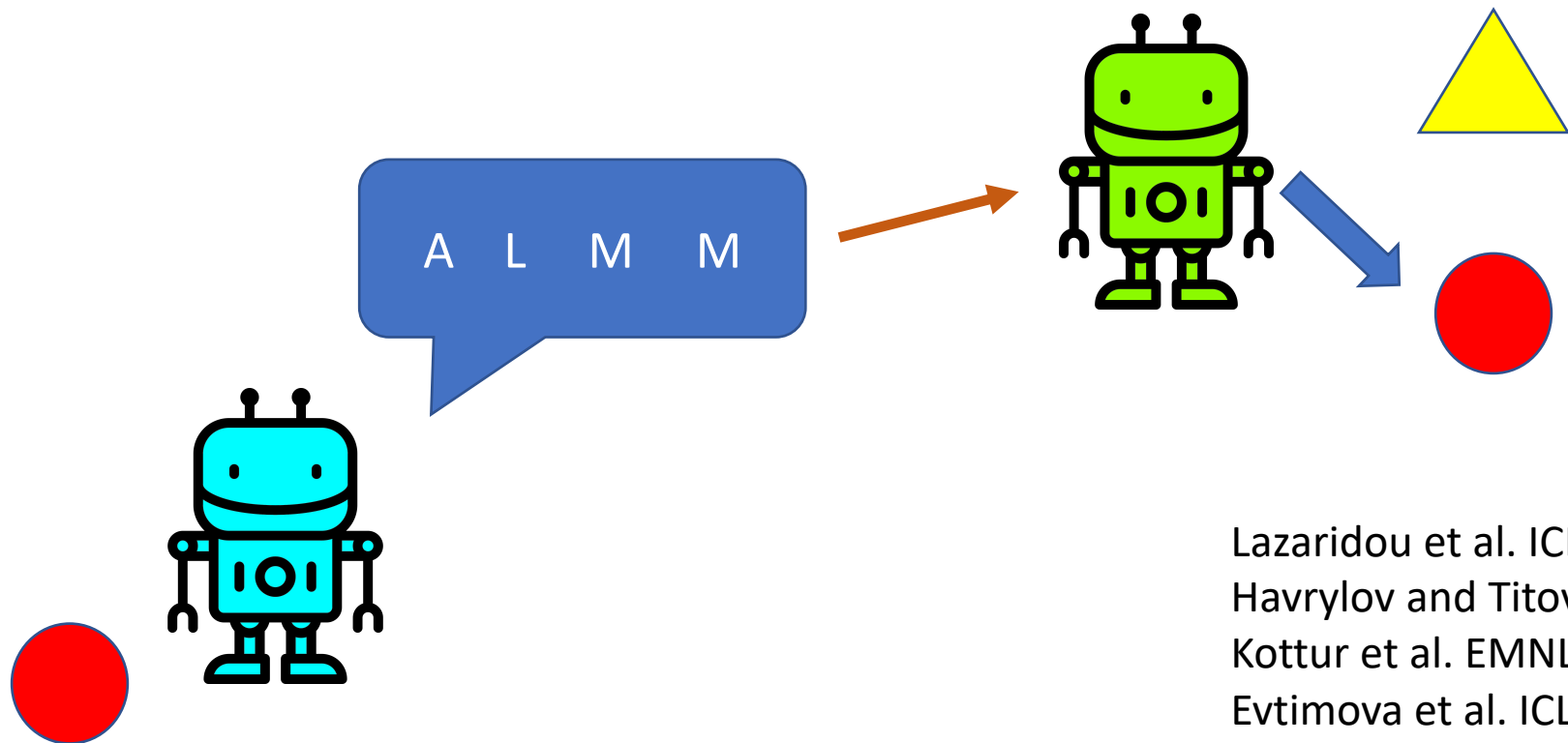


Head source: Pixabay

Language intrinsically adds value! e.g., Lupyan and Bergen TopiCS 2014

Language emergence as representation learning

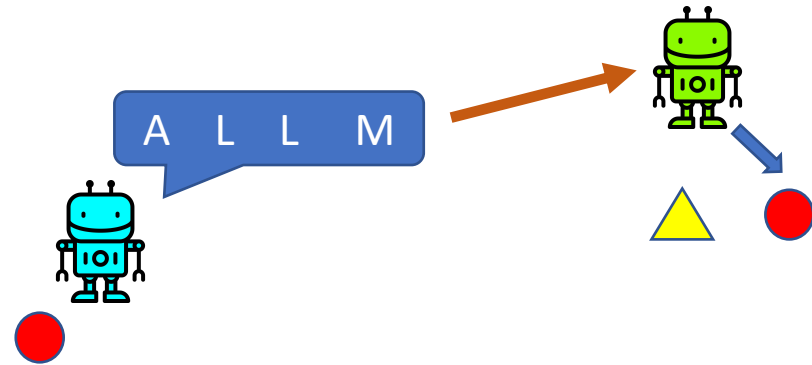
Image sources: Diane Bouchacourt,
<http://www.publicdomainfiles.com/>



Lazaridou et al. ICLR 2017,
Havrylov and Titov NIPS 2017,
Kottur et al. EMNLP 2017,
Evtimova et al. ICLR 2018,

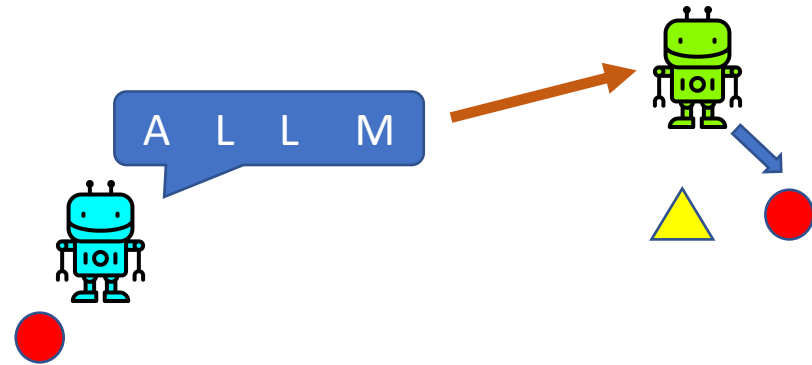
...

Our setup



- Two networks must jointly solve a task
- *Sender* network sees some input (e.g., a target image) and sends a single (possibly multiple-symbol) message to *Receiver* network
- Receiver has to perform an action (e.g., choose right target image) to complete the task
- Networks trained only to solve the task
- **No** supervision on the messages generated by Sender

Our setup: a few technical details



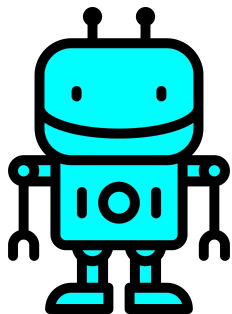
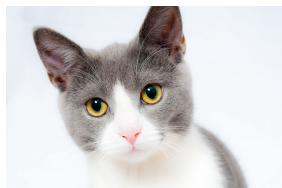
- Agent networks have task-specific input and output layers (MLPs, CNNs)
- Multi-symbol messages produced/read by LSTM cells
- Task-related cost function is of course task specific (reward-based, cross-entropy...)
- Must back-propagate through sampled discrete symbols: Reinforce or Gumbel-Softmax approximation
- All reported results robust across multiple initialization and hyperparameters

Outline

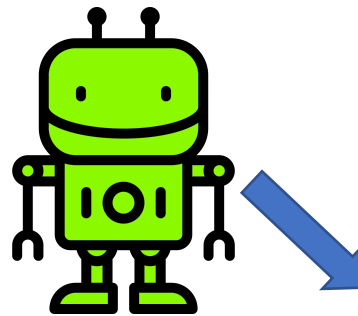
- Language emergence as representation learning
- ***Proto-meaning: What does emergent language represent?***
- Proto-form: How does emergent language represent?
- Take-home messages

Learning words for high-level concepts

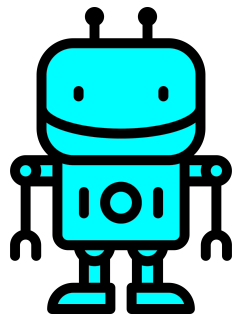
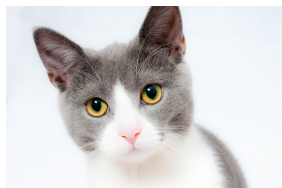
Picture source: Pexels



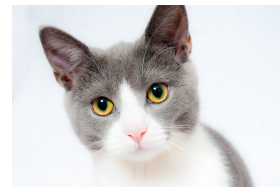
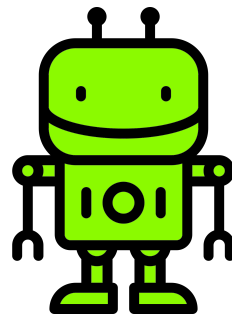
127



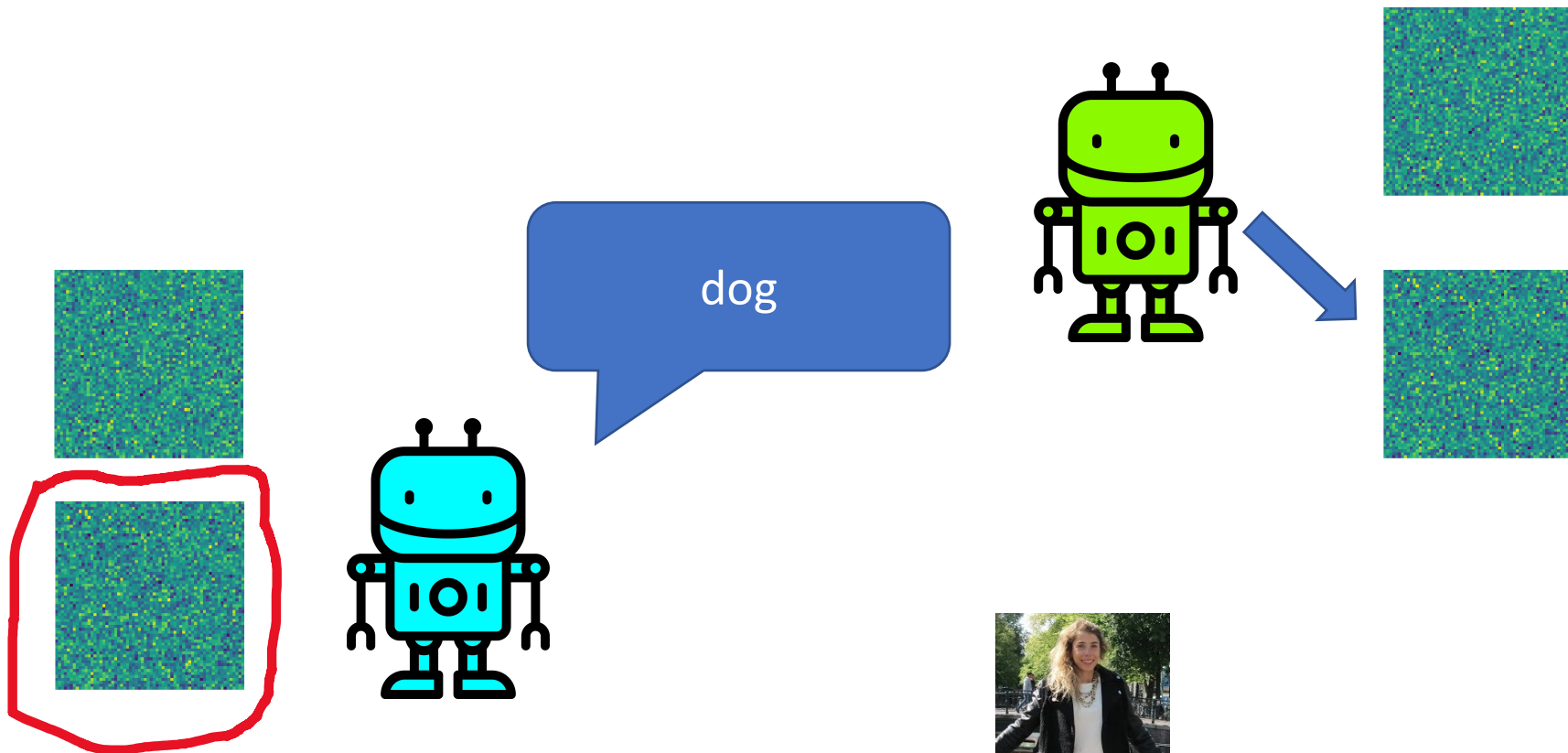
Learning words for high-level concepts



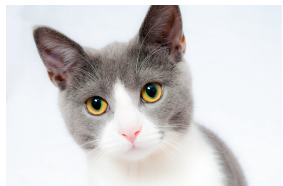
dog



Learning words for high-level concepts???



What would you talk about if all you had to do was to discriminate picture pairs?



"dog" (= "concept 157 of 500")

9 bits



"larger average intensity in
16 pixels at image center"

1 bit

saves time and
effort!

Efficient encoding of input information

- The cost function of human language?

$$\text{Cost} = -\alpha \text{ Task Accuracy} + \beta \text{ Word Entropy}$$

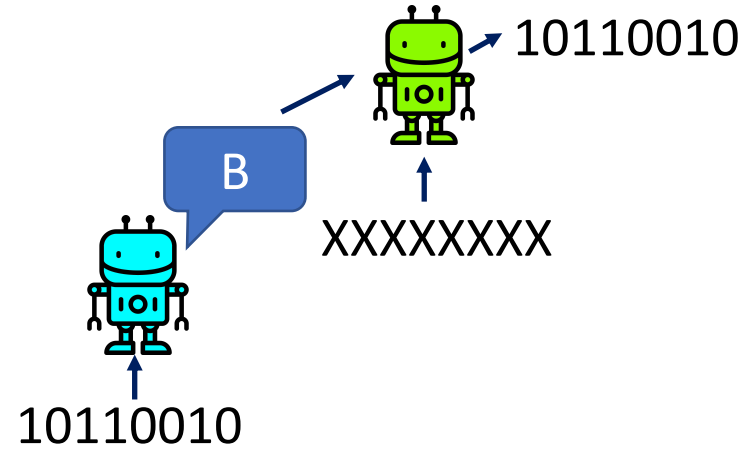
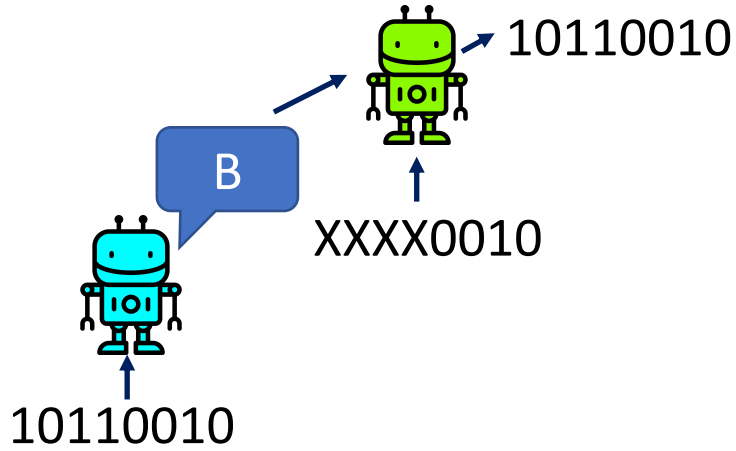
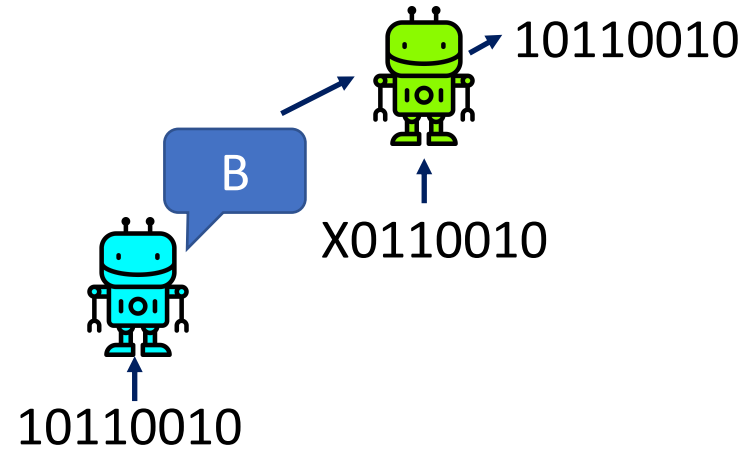
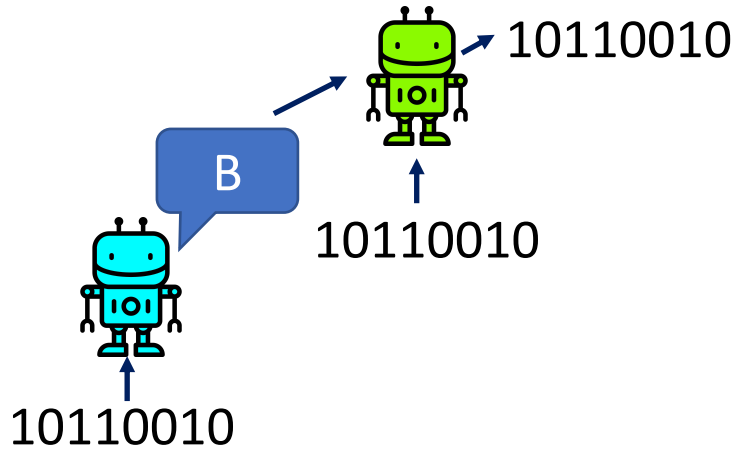
evidence from naming in domains such as color and kinship (Zavlasky/Regier/Kemp/Tishby), might even account for Zipf's law (Ferrer-i-Cancho)

- Are the agents in emergent language simulations optimizing for a similar trade-off (without need for explicit Word Entropy term minimization)?
- Kharitonov et al. 2019: Information minimization in emergent languages

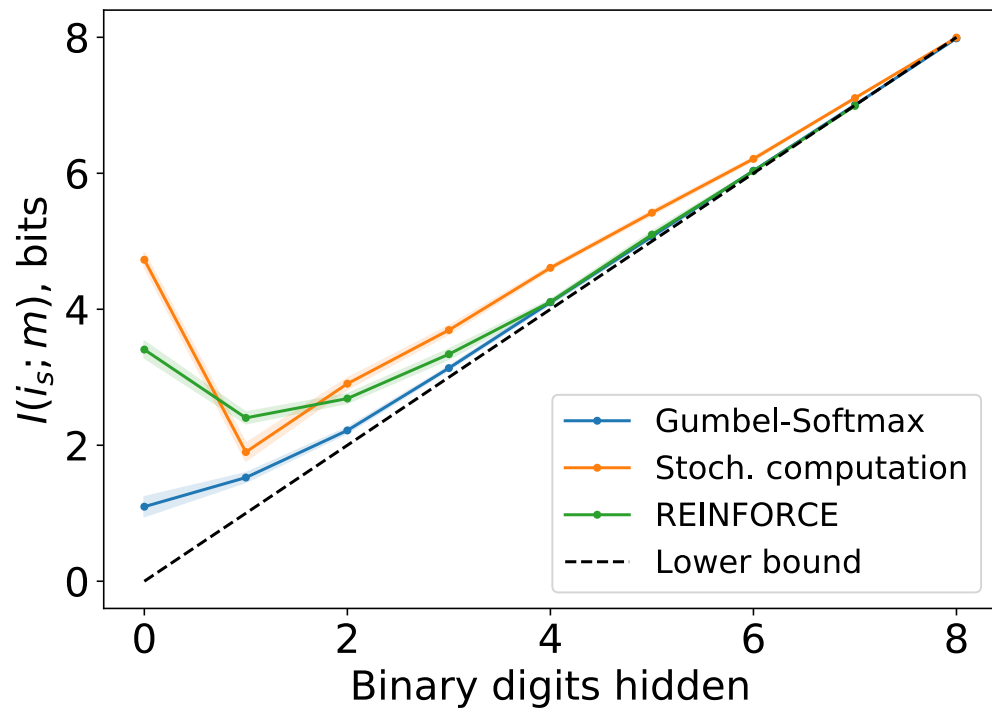
<https://arxiv.org/abs/1905.13687>



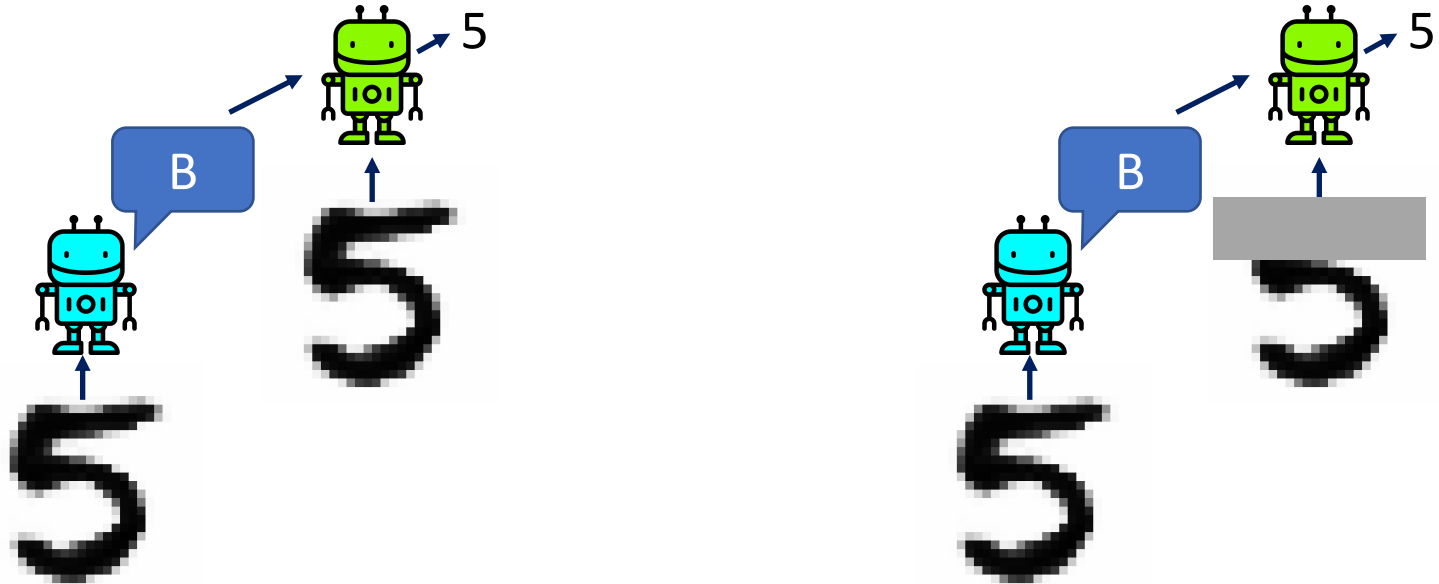
The number game



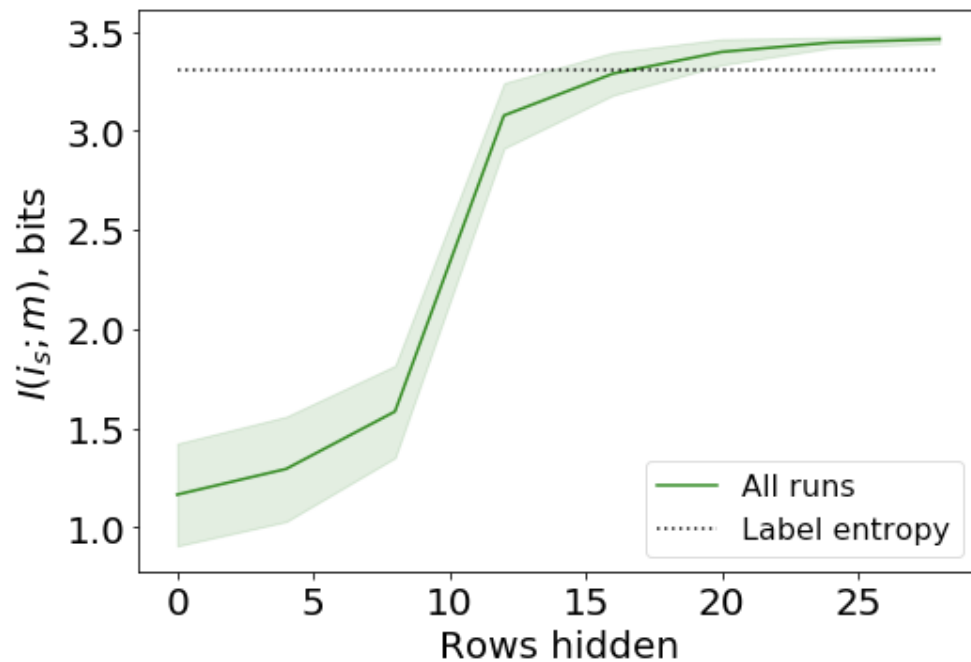
The number game: results



The MNIST game



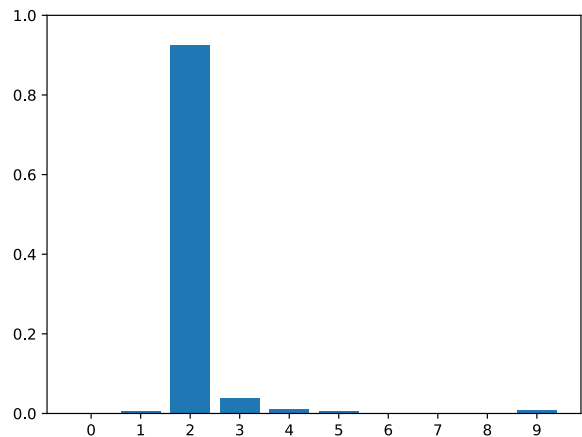
The MNIST game: results



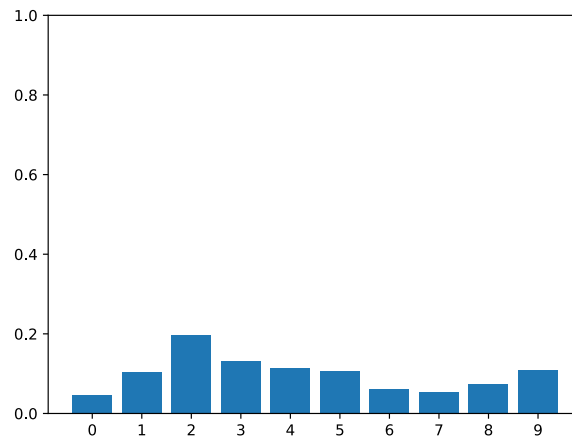
The MNIST game: effects of a looser discrete approximation

Approximating discrete one-hot vectors at training:

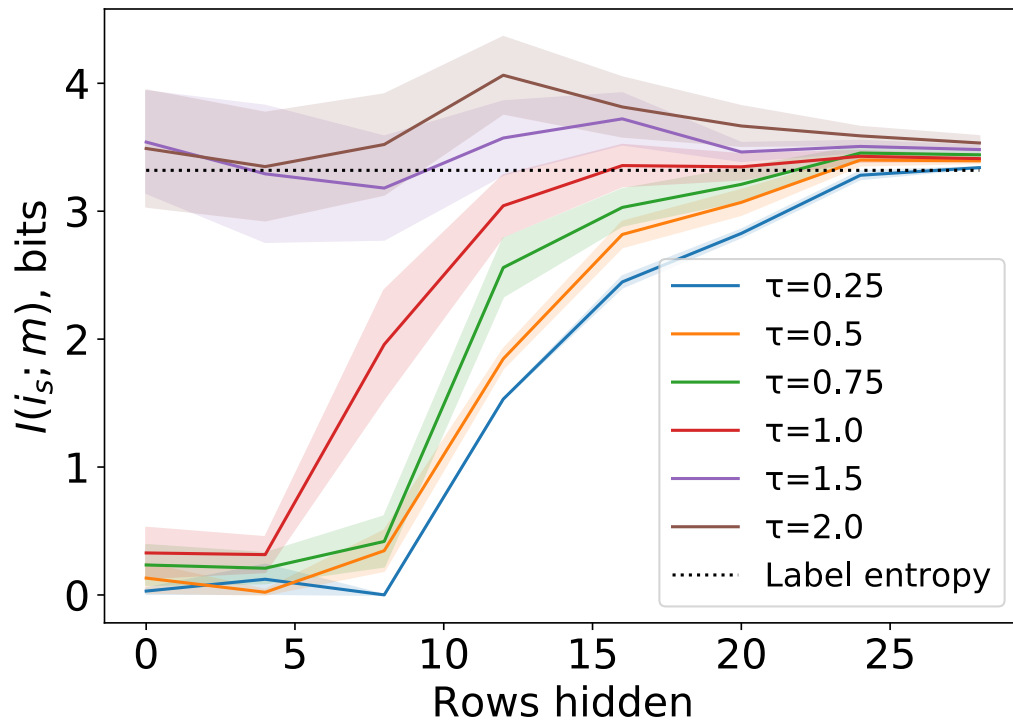
temperature (τ) at 0.25



temperature (τ) at 2



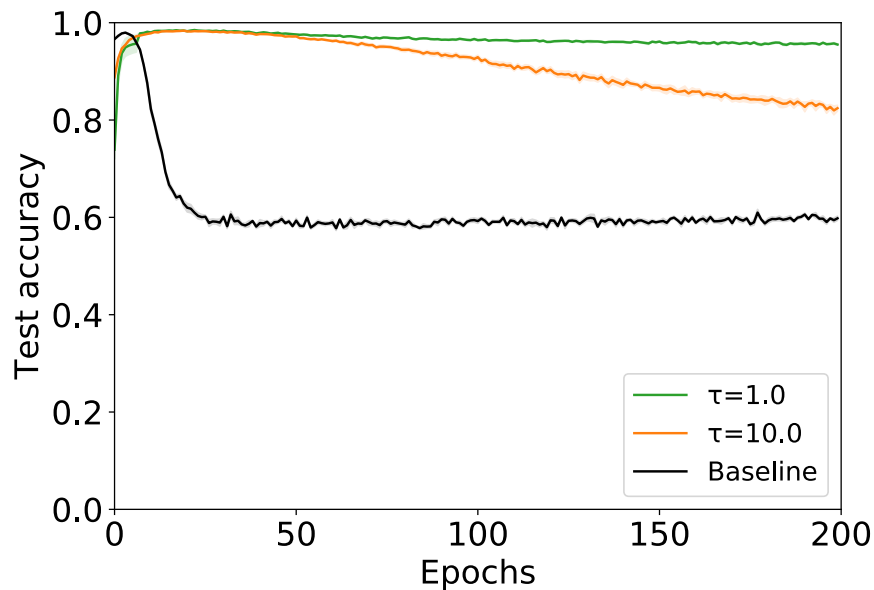
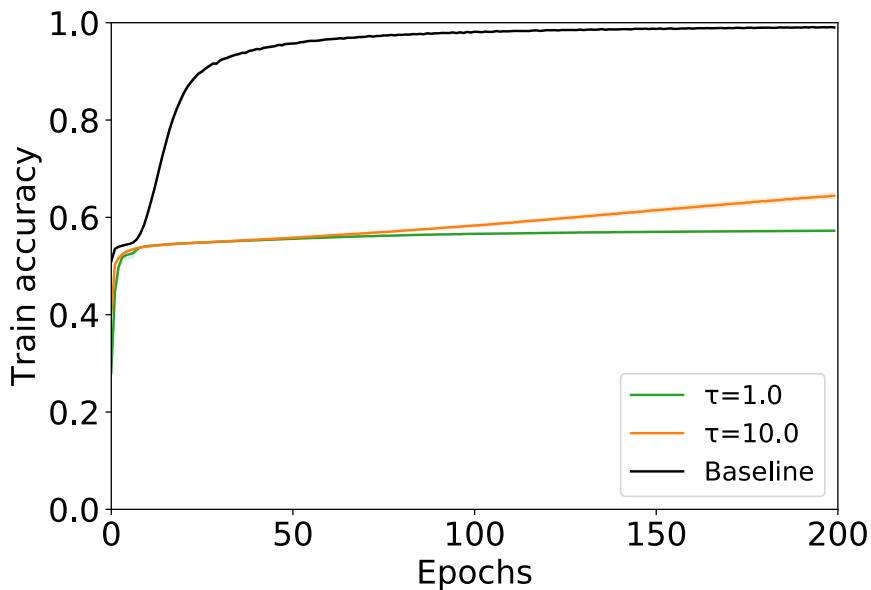
The MNIST game: effects of a looser discrete approximation



The discrete bottleneck is good for you

50% MNIST labels shuffled at training

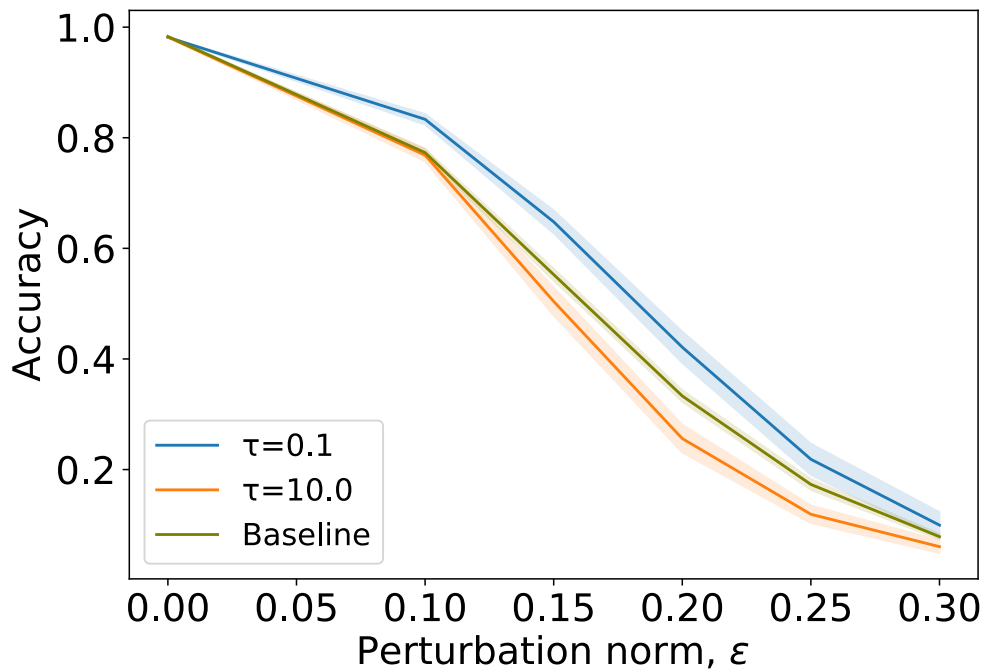
lower τ = more discrete = less noise fitting



The discrete bottleneck is good for you

Adversarial perturbation after Goodfellow et al. 2014

lower τ = more discrete = more robust to adversarial examples



Ad-interim conclusion

- Information minimization might be a natural effect of having to communicate through a discrete channel
 - emerging in human *and neural network* communication
- ... and it is good for you!
 - is this why language is discrete in the first instance?
- However, it also implies there is no wishful room for emergence of a more expressive language than strictly needed
 - agents will find the minimum-information solution that solves their task: if that involves talking about a single pixel, that's what their language will revolve around

Outline

- Language emergence as representation learning
- *Proto-meaning*: What does emergent language represent?
- **Proto-form**: How does emergent language represent?
- Take-home messages

Zipf's law of abbreviation

Human language is not only efficient in how it encodes meanings, but also in the form of the codewords

Source: Gutenberg books frequency list, from Wiktionary

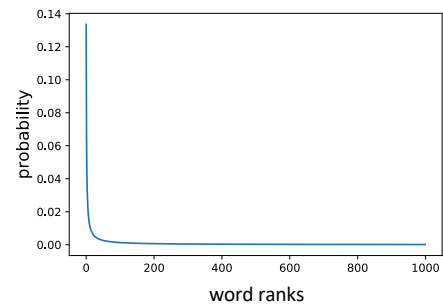
top 10 most frequent words

- the
- of
- and
- to
- in
- I
- that
- was
- his
- he

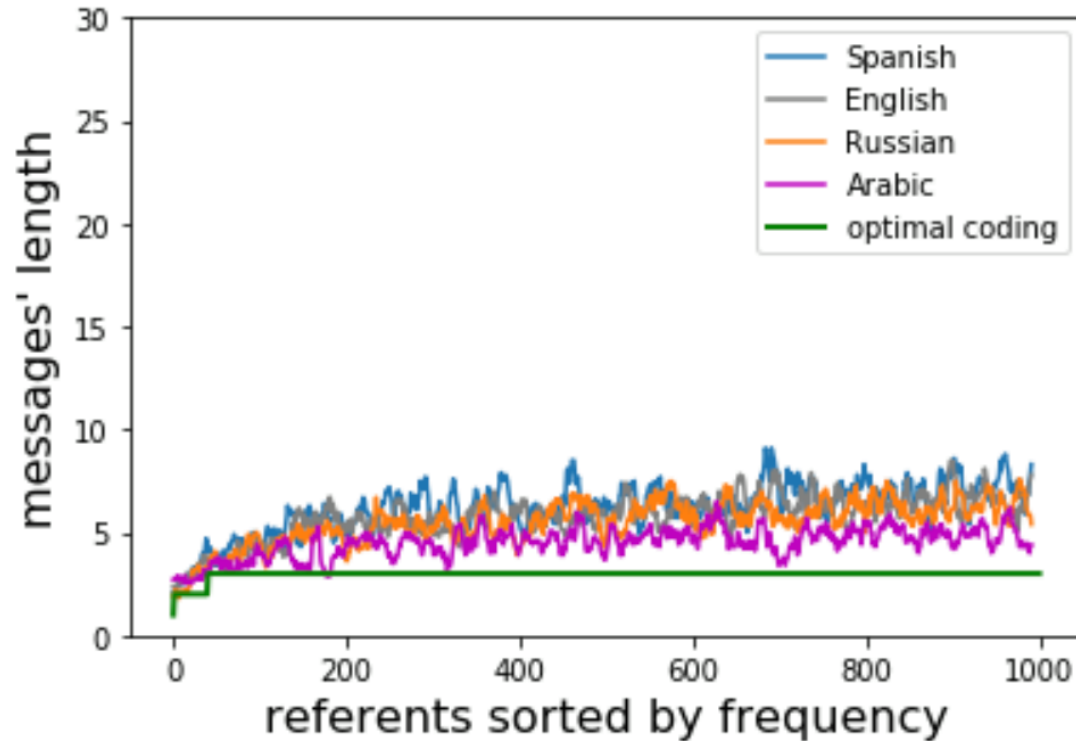
10 random rare(st) words

- anadromous
- barmaster
- cruddy
- gemstone
- gonzo
- idolization
- pigling
- sanguinity
- unpredictability
- walkman

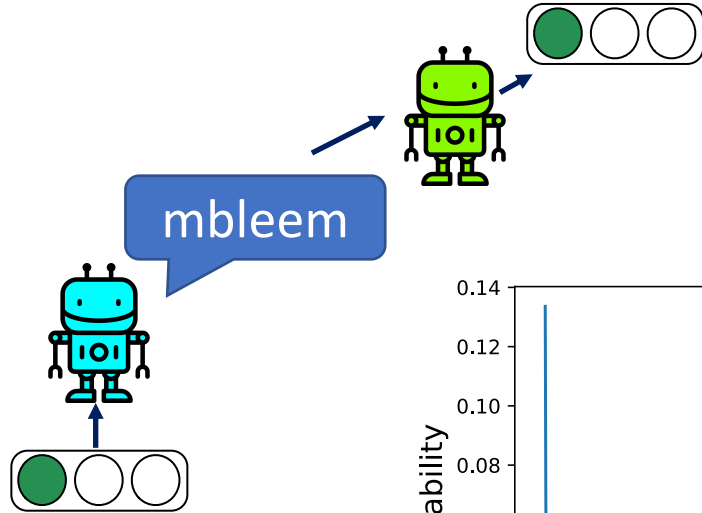
recall Zipf's law
of word frequencies!



ZLA and optimal encoding

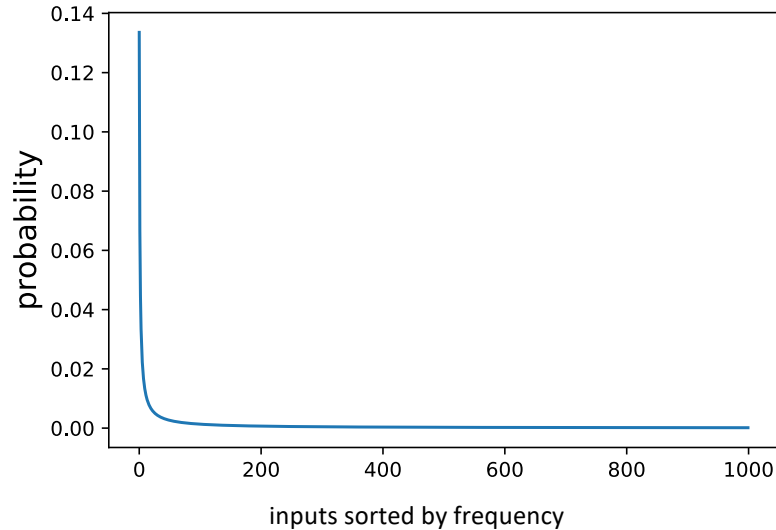


The Power-law Referent Game

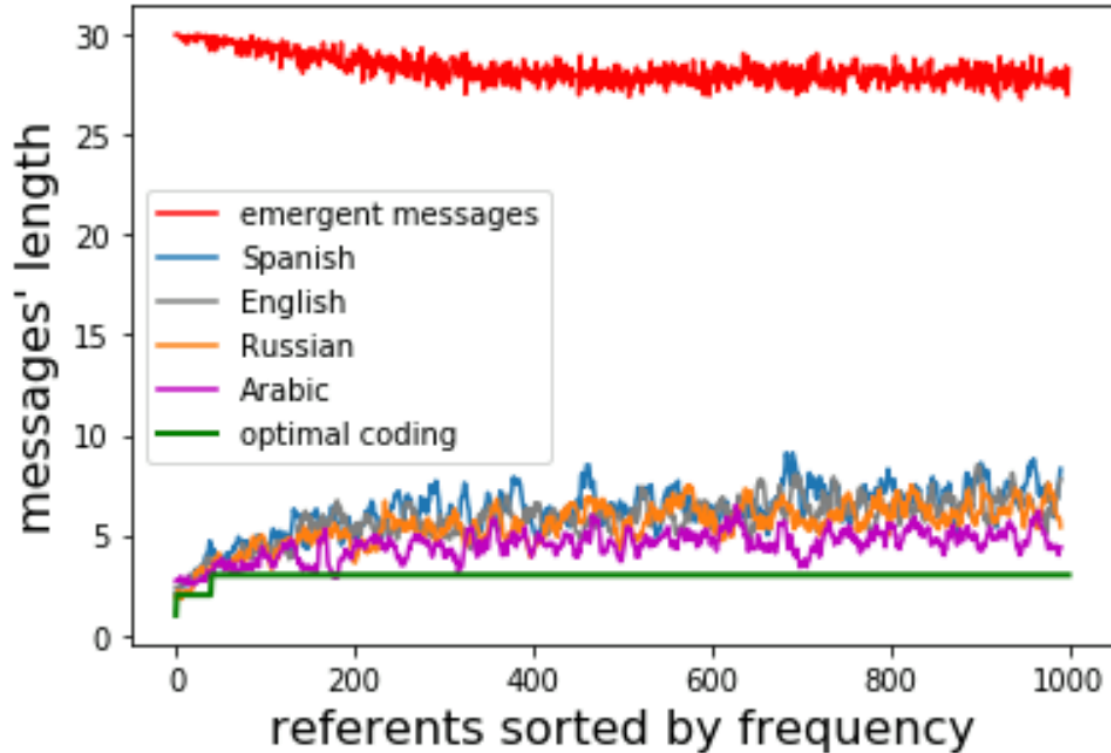


Chaabouni et al. Anti-efficient encoding in emergent communication

<https://arxiv.org/abs/1905.12561>



Anti-efficient message encoding in emergent language



more frequent messages
significantly longer,
unique case compared to
human language and
animal communication
systems (Ferrer-i-Cancho
et al., Cognitive Science
2013)

Most frequent "words": English vs emergent language

English

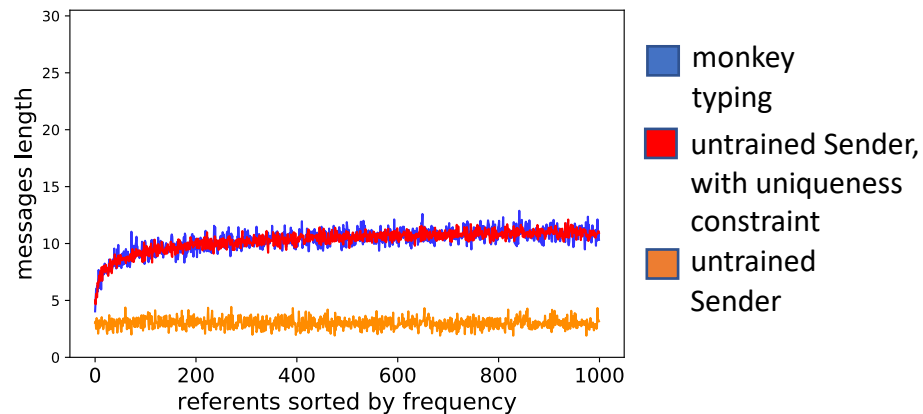
- the
- of
- and
- to
- in
- I
- that
- was
- his
- he

Emergent language

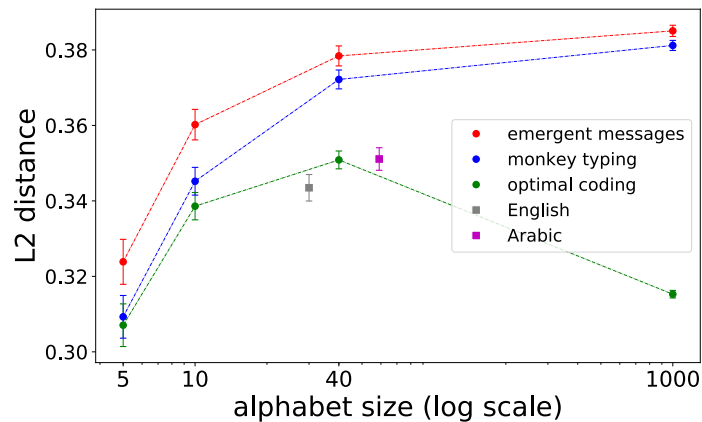
- llmuhhmummmmmzuyuuyzzvqzplan
- naauhhhhhhhhhpucczzpbqaoaqpln
- shmhhhhxxxxxxzummyyttlclgl
- jahhhhhhuxxzzzzzzzzzzceyoawb

The origin of anti-efficient encoding

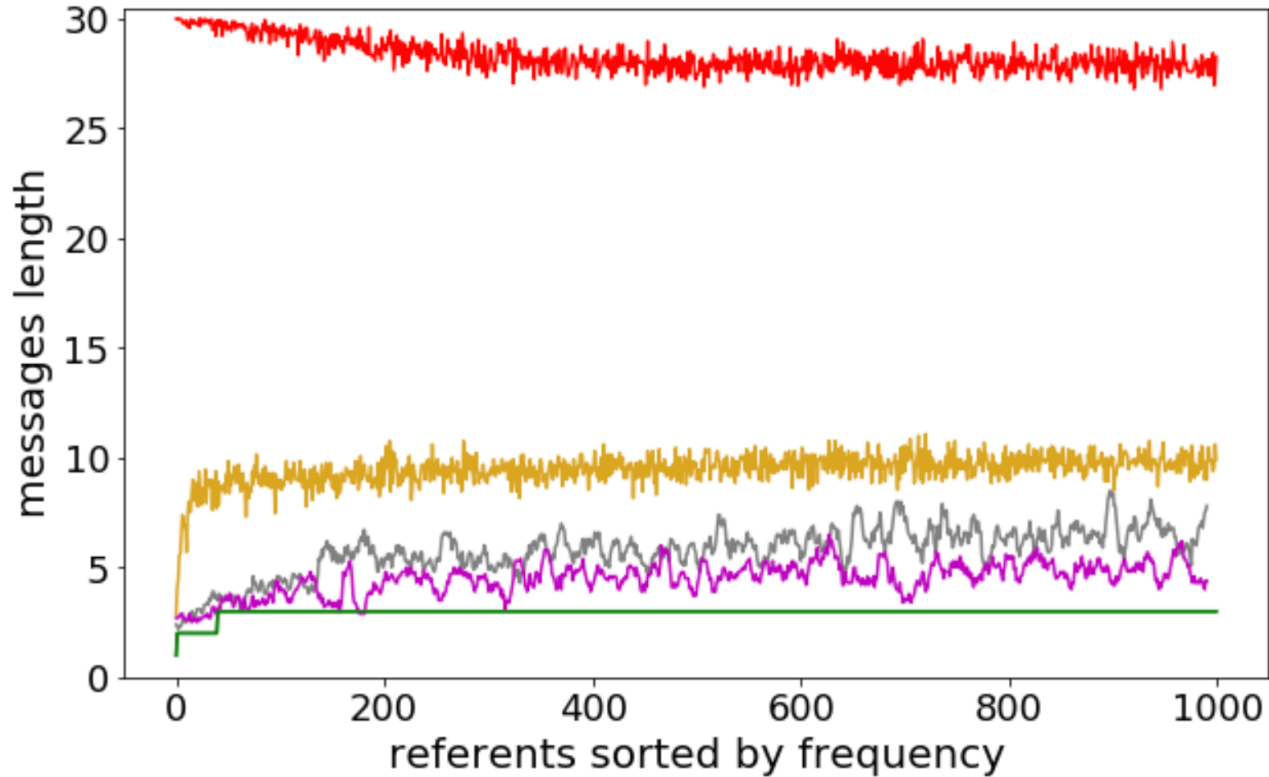
Sender initially obeys ZLA, but has no "articulatory" pressure to stick to it



Longer messages are easier to discriminate for Receiver



Explicitly penalizing longer messages



— Emergent messages — Emergent messages / length pressure — Optimal coding — English — Arabic

Outline

- Language emergence as representation learning
- *Proto-meaning*: What does emergent language represent?
- Proto-form: How does emergent language represent?
- **Take-home messages**

The general framework

- Language is a discrete representation code
- It is a powerful tool to represent what we perceive and what's in our head
- Simulations of language emergence in neural agents are a form of representation learning
 - They might help us to understand core properties of human language and...
 - to endow interactive AIs with a similarly powerful tool

Lessons learned and food for thought

- Given a task, the corresponding emergent language will not be more generally expressive than what is needed to solve the task
- This is not a bug, it's an (occasionally annoying) feature!
- It crucially depends on the discrete nature of communication...
- and might support more robust problem solving

- On the other hand, neural agents develop positively *anti*-efficient message forms
- ... due to lack of pressure to minimize production effort
- Does this phenomenon account for other undesirable properties of the emergent code (e.g., lack of compositionality, Kottur et al. EMNLP 2017)?
 - And would making the code more efficient lead to other benefits?

- What's the relation between these two results?

Now try it yourself!

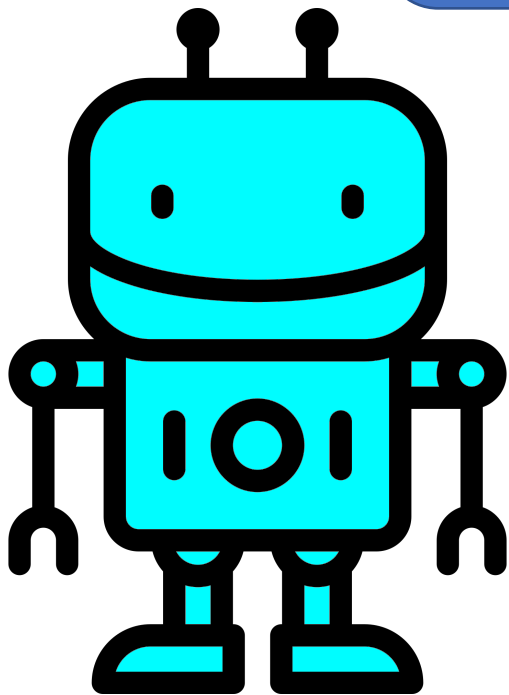
<https://github.com/facebookresearch/EGG>

EGG : Emergence of lanGuage in Games

Introduction

EGG is a toolkit that allows researchers to quickly implement multi-agent games with discrete channel communication. In such games, the agents are trained to communicate with each other and jointly solve a task. Often, the way they communicate is not explicitly determined, allowing agents to come up with their own 'language' in order to solve the task. Such setup opens a plethora of possibilities in studying emergent language and the impact of the nature of task being solved, the agents' models, etc. This subject is a vibrant area of research often considered as a prerequisite for general AI. The purpose of EGG is to offer researchers an easy and fast entry point into this research area.

t h a n k y o u



<https://github.com/facebookresearch/EGG>