

On the Word Boundaries of Emergent Languages

Ryo Ueda

2023/06/29 @ SNL2023

The University of Tokyo

Self-introduction

Self-introduction

name **Ryo Ueda (D1)**
aff **Miyao Lab, The University of Tokyo**
field **Emergent Communication**

Interest

How **statistical** properties
of **languages** emerge?

News

- Organized a session at JSAI2023
- Released a pre-print

<https://psyarxiv.com/rz5ng>

my homepage



pre-print



Paper I introduce today

Paper I introduce today

Today, I would like to introduce this paper:

Published as a conference paper at ICLR 2023

ON THE WORD BOUNDARIES OF EMERGENT LANGUAGES BASED ON HARRIS'S ARTICULATION SCHEME

Ryo Ueda, Taiga Ishii & Yusuke Miyao

The University of Tokyo

{ryoryueda, taigarana, yusuke}@is.s.u-tokyo.ac.jp

accepted to **ICLR2023** (held in **Kigali, Rwanda**)

Pictures in Rwanda



Overview of our paper 1/2

Research Question 🤔

Harris's articulation scheme (HAS)
holds for **emergent languages** ?



Emergent languages have **meaningful words** ?

Emergent Language

communication protocol
emerging among agents

HAS

statistical property on
words in natural language

Results Implication

HAS does not hold for **emergent languages** 😞



Emergent languages lack **meaningful words** 😞

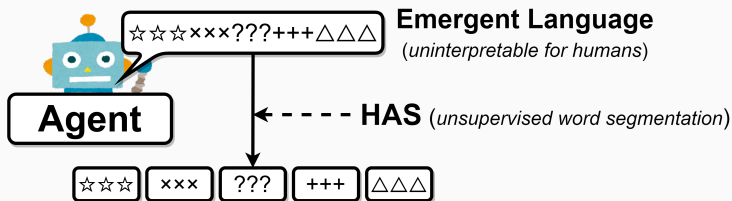
Overview of our paper 2/2

Research Question 🤔

Harris's articulation scheme (HAS)
holds for **emergent languages** ?



Emergent languages have **meaningful words** ?



HAS is (anyway) applicable to **emergent languages**
to obtain **segments**

But does it really make sense ? 🙋 That's the problem 5/20

Background: Emergent Language

Background: Emergent Language

Emergent Language

Communication protocol emerging among agents
a.k.a **Emergent Communication**

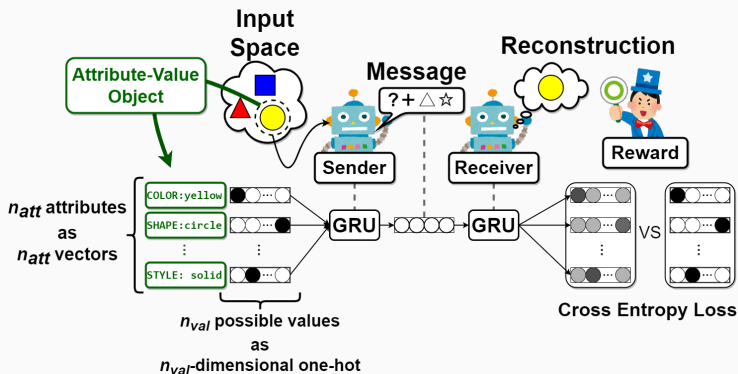
Typical motivations in this area

1. Develop interactive AI
[e.g., [Lazaridou et al., 2018](#)]
2. Simulate language evolution
[e.g., [Kirby, 2001](#)]
3. Clarify diff between emergent/human languages
[e.g., [Chaabouni et al., 2020](#)]

Signaling Game

Typical Environmental Setting in this area

Signaling Game [Lewis, 1969]



Attribute-Value Setting

#attributes $n_{att} \in \mathbb{N}$, **#values** $n_{val} \in \mathbb{N}$

Background: Harris's Articulation Scheme (HAS)

Next character uncertainty and word boundaries

Example: “natural”

n... 🤔 ? I have no idea on the next
nat... 🤔 !? The next may be “u” or “i”
natura... 😄 !! The next must be “l”
natural... 🤔 ?

Harris’s hypothesis [Harris, 1955]

Next character gets uncertain again

👉 there tends to be a **word boundary** there

Harris’s Articulation Scheme (HAS) [Tanaka-Ishii, 2021]

Information-theoretic re-formulation

👉 **Uncertainty as Entropy**

Conditional Entropy & Branching Entropy 1/2

A set of character sequences Σ^* ,
a sequence $x_{1:n} = x_1 \cdots x_n \in \Sigma^*$,
a random variable $X_{1:n} = X_1 \cdots X_n$ over Σ^*

Conditional Entropy $H(n)$

$$\mathcal{H}(X_{n+1}|X_{1:n})$$

Branching Entropy $h(x_{1:n})$

$$\mathcal{H}(X_{n+1}|X_{1:n} = x_{1:n})$$

Its intuition

next character **uncertainty**
on average

Its intuition

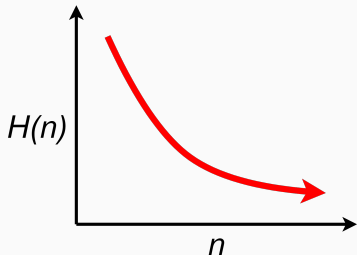
next character **uncertainty**
in a specific context

Conditional Entropy is the average of Branching Entropy

$$\therefore H(n) = \sum_{x_{1:n} \in \Sigma^*} p(x_{1:n}) h(x_{1:n})$$

Conditional Entropy & Branching Entropy 2/2

Conditional Entropy



Branching Entropy



Their behavior in natural languages

Conditional Entropy $H(n)$ monotonically decreases

VS

Branching Entropy $h(x_{1:n})$ repeatedly falls and rises

👉 In natural languages, uncertainty decreases on average, but jitters in specific contexts

Harris's Articulation Scheme (HAS)

Harris's Articulation Scheme (HAS) [Tanaka-Ishii, 2021]

There tend to be **word boundaries**
at the **increasing points** of **Branching Entropy h**

Branching Entropy h decreases on average
(\because Conditional Entropy H monotonically decreases),
but sometimes increases



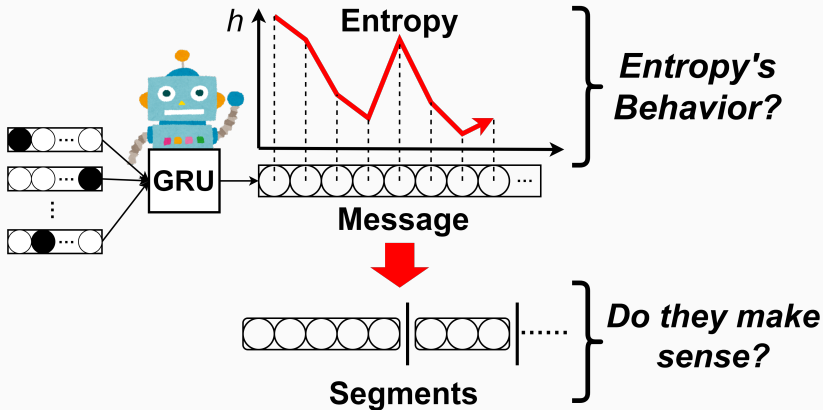
There tend to be a **word boundary** 🙌

Problem Definition

Problem Definition 1

Re: Research Question 🤔

HAS holds for **emergent languages**?



Problem Definition 2

Re: Research Question 🤔

HAS holds for emergent languages?

Three questions to answer:

In emergent languages...

Q1 Conditional Entropy monotonically decreases ?

Q2 Branching Entropy jitters ?

Q3 Tentatively obtained segments are meaningful ?

Entropies behave like natural languages? 👉 Q1, Q2

Obtained segments are really “words”? 👉 Q3

Problem Definition 3

Tentatively obtained segments are meaningful?

Difficulty of Q3

No ground truth data of word segmentation

Emergent languages have words in the first place?

To mitigate this issue,
replace Q3 with the following questions:

Additional questions alternative to Q3

Q3-1 #attributes \propto the number of boundaries ?

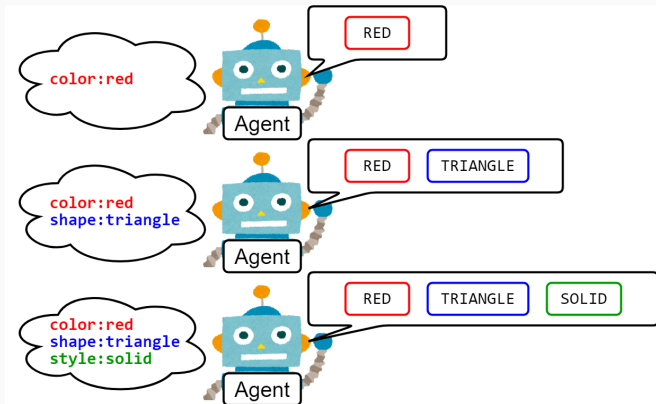
Q3-2 #values \propto the number of distinct segments ?

Q3-3 segment-level compositionality $>$
character-level compositionality ?

Problem Definition 4

Intuition of Q3-1

#attributes \propto the number of boundaries ?



For Q3-2 & Q3-3, refer to our paper!

Experimental Results

Experimental Results

To Q1, Q2, & Q3

Q1 Conditional Entropy H monotonically decreases?

👉 YES ✓

Q2 Branching Entropy h jitters?

👉 YES ✓

Q3 Tentatively obtained segments are meaningful?

👉 Probably NO 😞

Alternatives to Q3

Q3-1 #attributes \propto the number of boundaries?

Q3-2 #values \propto the number of distinct segments?

Q3-3 segment-level compositionality $>$

char-level compositionality? 👉 NO for all 😞

Summary

Summary

Research Question 🤔

Harris's articulation scheme (HAS)
holds for **emergent languages** ?



Emergent languages have **meaningful words** ?

Emergent Language

communication protocol
emerging among agents

HAS

statistical property on
words in natural language

Results Implication

HAS does not hold for **emergent languages** 😞



Emergent languages lack **meaningful words** 😞

Additional Remarks

Entanglement and Disentanglement

“Emergent languages have meaningful segments”

👉 implies...

Disentanglement of Segments

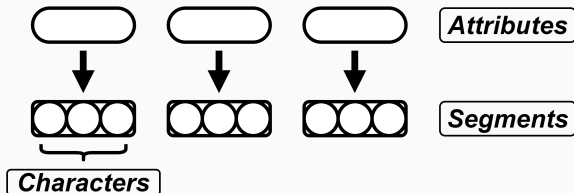
1-to-1 correspondence between attributes/segments

Segments are **disentangled**

Entanglement of Characters

Segments can be obtained via **characters' entropy**

Characters are **entangled** (statistically correlated)



References i

- Rahma Chaabouni, Eugene Kharitonov, Diane Bouchacourt, Emmanuel Dupoux, and Marco Baroni. Compositionality and generalization in emergent languages. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel R. Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 4427–4442. Association for Computational Linguistics, 2020. doi: 10.18653/v1/2020.acl-main.407.
- Zellig S. Harris. From phoneme to morpheme. *Language*, 31(2):190–222, 1955. ISSN 00978507, 15350665.
- Simon Kirby. Spontaneous evolution of linguistic structure—an iterated learning model of the emergence of regularity and irregularity. *IEEE Trans. Evol. Comput.*, 5(2):102–110, 2001. doi: 10.1109/4235.918430.

References ii

Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. Emergence of linguistic communication from referential games with symbolic and pixel input. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018.

David K. Lewis. *Convention: A Philosophical Study*. Wiley-Blackwell, 1969.

Kumiko Tanaka-Ishii. *Articulation of Elements*, pages 115–124. Springer International Publishing, Cham, 2021.