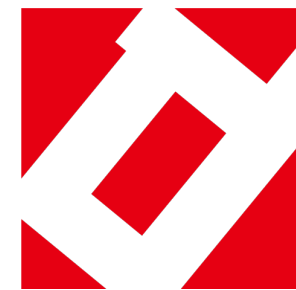


# Discovering Universal Geometry in Embeddings with ICA

Hiroaki Yamagiwa\*<sup>1</sup>, Momose Oyama\*<sup>1,2</sup>, Hidetoshi Shimodaira<sup>1,2</sup>

<sup>1</sup>Kyoto University, <sup>2</sup>RIKEN



**EMNLP  
2023**

# Overview

## Purpose

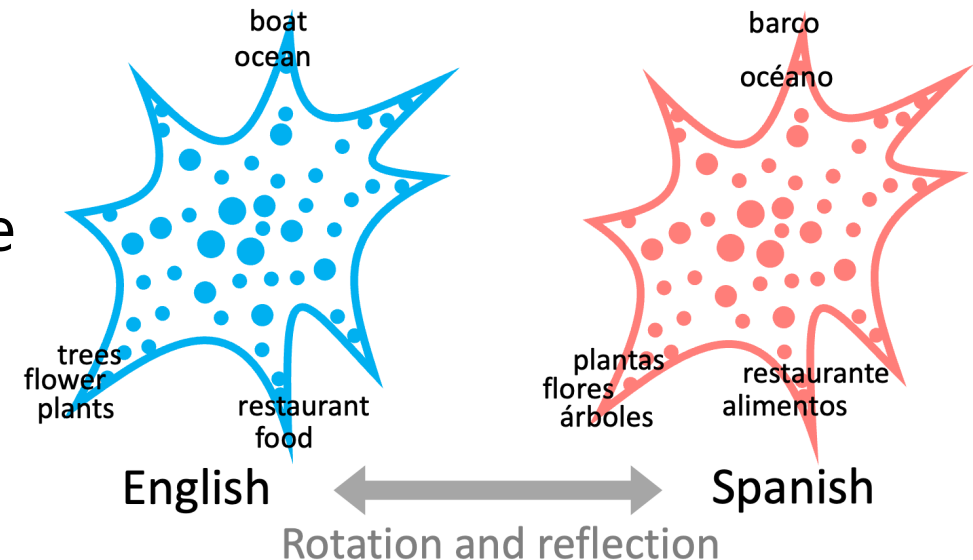
1. Understanding How Embedding Geometry Encodes Meaning
2. Exploring the Universality of Geometric-Meaning Relationships in Embeddings

## We use ICA

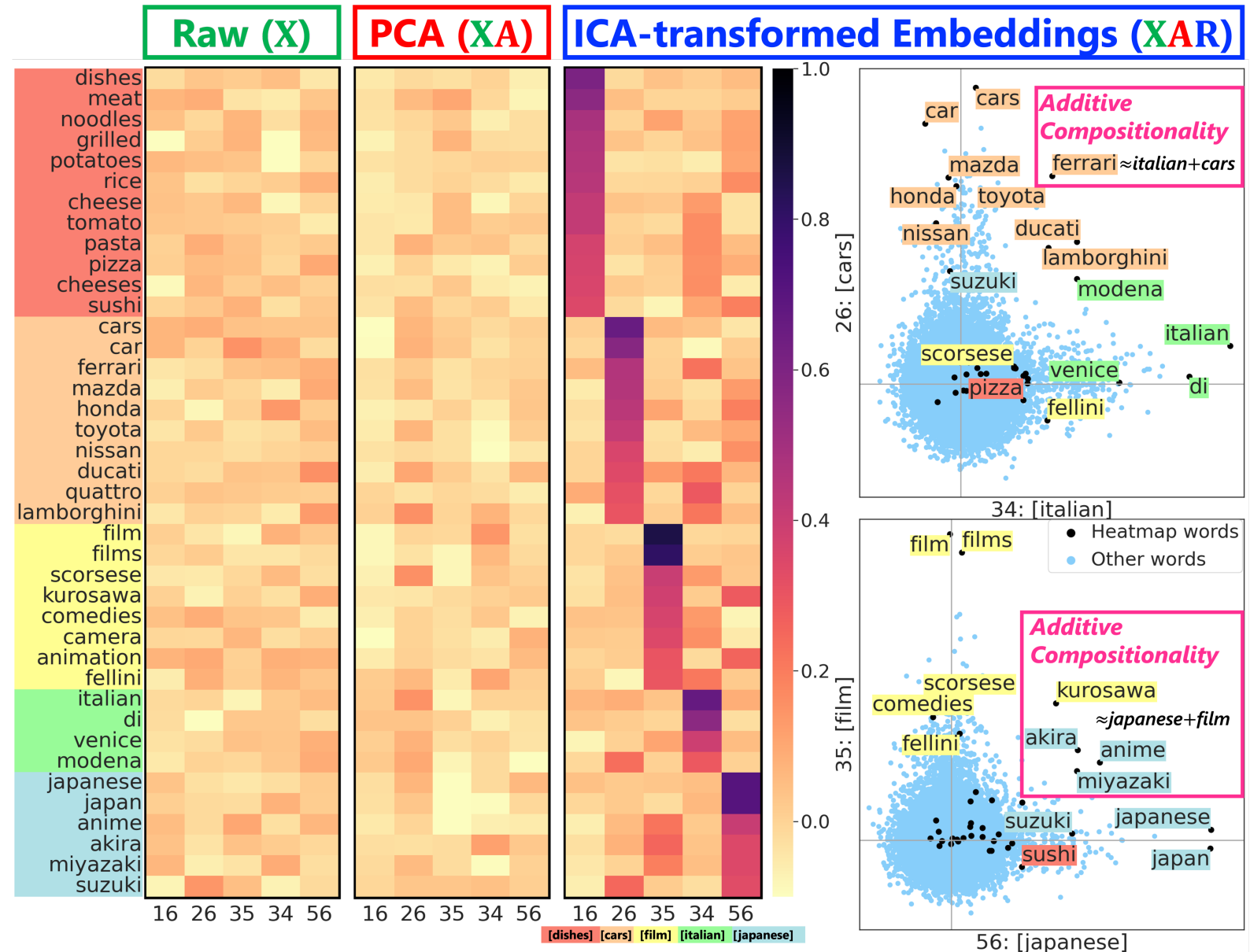
**Independent Component Analysis** (ICA)  
Finds Statistically Independent Axes

## Results

1. **Independent Axes** in Embeddings are **"Spiky"** and **Interpretable**
2. **"Spiky"** and **Interpretable** Axes are **Universal** in Various Embeddings



# ICA Discovers "Spiky" and Interpretable Axes

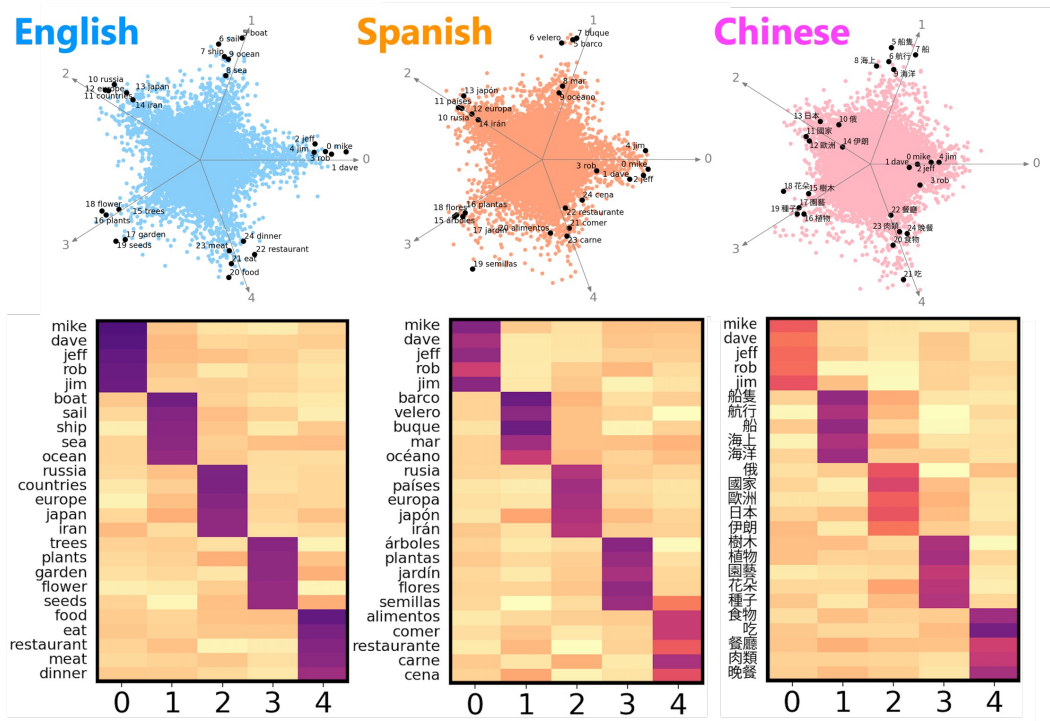


# Universal Geometry in Embeddings

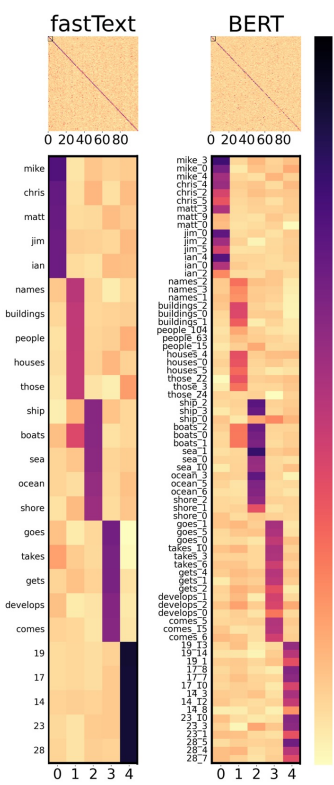
## ① Cross-Lingual

## ② Model

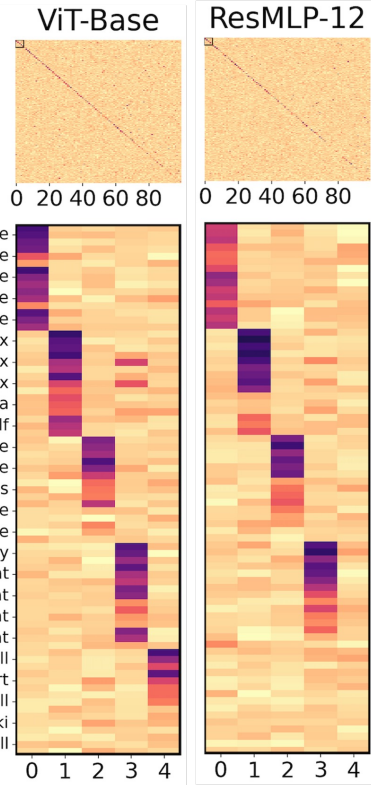
## ③ Modality



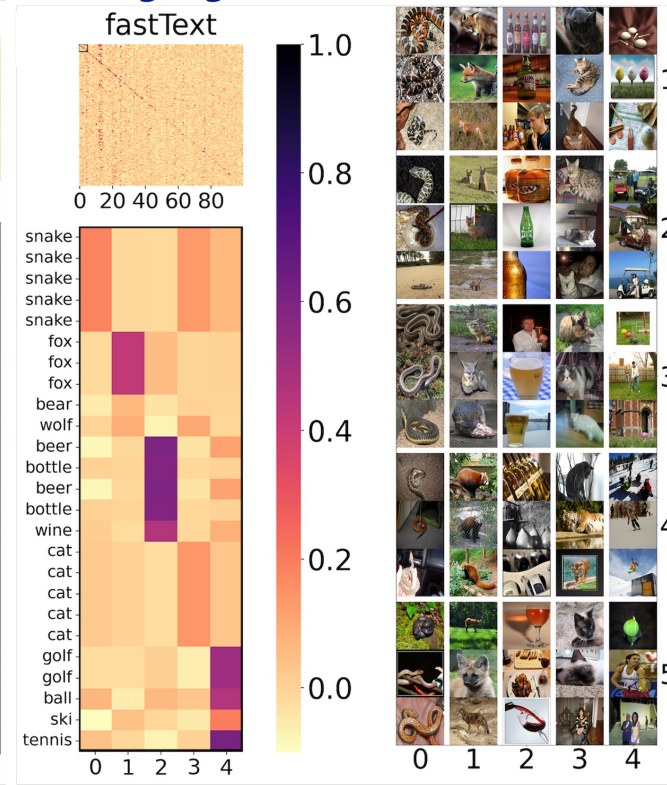
### Static Contextualized



### Vision



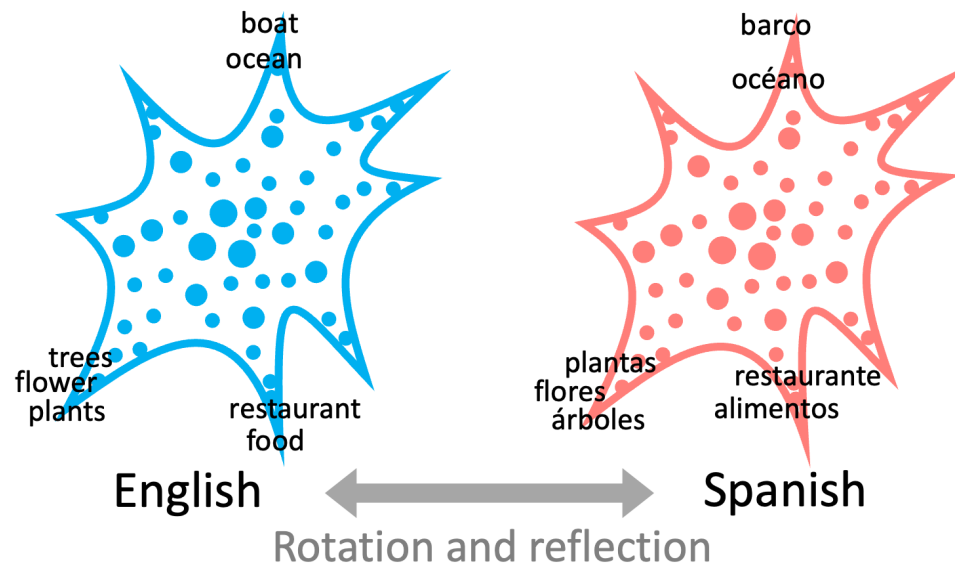
### Language



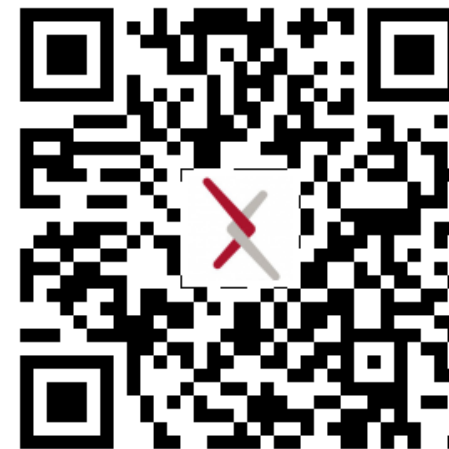
1. Apply ICA to each embedding
2. Permute axes by correlation for translation pairs

# Summary

1. **Independent Axes** in Embeddings are **"Spiky"** and **Interpretable**
2. **"Spiky"** and **Interpretable** Axes are **Universal** in Various Embeddings



Paper



Codes



● Paper: <https://arxiv.org/abs/2305.13175>

● Codes: <https://github.com/shimo-lab/Universal-Geometry-with-ICA>