# The first sentence the second sentence

## a smaller subtitle

Dewi E. Timman
12419273

Bachelor thesis
Credits: 12 EC

Bachelor *Cognition, Language and Communication*

University of Amsterdam
Faculty of Humanities
Spuistraat 134
1012 VB Amsterdam

*Supervisor*
Dr. R. G. Alhama

Institute for Logic, Language and Computation
Faculty of Science
University of Amsterdam
Science Park 107
1098 XG Amsterdam

April 17, 2024

# Abstract

# Contents

# Chapter 1

# Introduction

We are able to talk and communicate with each other. Our Languages are very complex. But how did language originate in the first place, and how do we acquire (new) languages?

## 1.1 Literature review

When researching the emergence of language, usually ... To do this, it is important to define a few concepts. First, about what kind of communication are we talking? Second, how can we research this communication? And Third, in what state do we start the research?

### 1.1.1 Agent-based linguistic communication

There are different kinds of communication. In this research, there will be looked at agent-based linguistics communication. This means that there is communication between two or more agents with the use of language. Usually, that means the agents need to solve a task together. To solve the task, agents need to succesfully communicate to each other.

### 1.1.2 Referential communication games

To research this agent-based linguistic communication, agents usually participate into a referential communication game. In such a game, agents talk about objects or other entities in a specified world. To do this, they need to come up with a language to communicate about they world. Usually, the agents do not have any prior linguistic knowledge.

### 1.1.3 Word embedding vectors

In this research, however, agents do start with prior linguistics knowledge. Agents use the word embedding vectors of a large language model (LLM) in order to ...

#### 1.1.3.1 LLMs

The LLM from which the word embedding vectors are chosen, is .... Because ...

## 1.2 Current research

Gap: Research with prior linguistic knowledge -> Do the agents do something with the structure/semantics already present in the word embedding vectors and can this say something about language emergence?
RQ: What is the effect of word embeddding vectors from a language model on the emergence of agent-based linguistic communication from referential communication games?

# Chapter 2

# Method

Programming a model and see how it behaves and what makes it behave like that.
Dependent variable: output meassured by the model and its evaluation metric -> how often is the communication succesfull?
Independent variable: input -> words/sentences

## 2.1 Data

What data? Data preprocessing? Data split?

### 2.1.1 (Data analyse)

What does the data look like? Is there a data imbalance?

## 2.2 Experimental Design

What is the setup?

## 2.3 Model description

What does the model look like? Which equations are used? Diagrams/pseudocode?

### 2.3.1 Hyperparameters

What are the hyperparameters?

## 2.4 (Optimization)

What optimizations were done?

# Chapter 3

# Results

# Chapter 4

# Conclusion

# Chapter 5

# Discussion