# *A05 - Analyzing "Arrival" Through the Lens of NLP*

Deep Learning Artificial Intelligence (ITAI-2376)
Prof. Anna Devarakonda
Date: July 01, 2025

**Submitted by: Iffraah Rehman**
**ID – W216165700**

# Table of Contents

# Analyzing "Arrival" Through the Lens of NLP

## 1. Overview - Movie's Plot

The story starts with twelve giant's aka aliens appear across the twelve locations throughout the world and hover in place. Louise Banks, a doctor of linguistics and expert translator, is recruited by American military to help the US Government communicate with the alien which they call 'Heptapods'. The goal is to contact the extraterrestrial beings and decode their language to communicate with them. Working alongside physicist Ian Donnelly, Louise strives to uncover the purpose of the aliens' visit.

As Louise delves deeper into the Heptapods' language, she realizes that mastering it profoundly changes her perception of reality especially her understanding of time. The Heptapod language is non-linear, and through it, Louise comes to see that it's not merely a means of communication, but also a completely different way of experiencing time. The aliens possess the ability to perceive the present and the future simultaneously, revealing a concept of time that is entirely new to her.

Overall, *Arrival* offers a subtle and nuanced exploration of the role of language and communication in shaping our understanding of the world. It highlights the profound influence that language can have on our perception of reality and the way we interpret our experiences. The film's exploration of language reflects the Sapir-Whorf hypothesis (linguistic relativity), which proposes that the structure of a language shapes how its speakers perceive and interpret the world. Arrival reflects on the profound impact language has on human understanding, empathy, and the ability to connect across cultural and even species boundaries

## 2. NLP Challenges (Movie VS Parallels to Real-World NLP Challenges)

The film Arrival presents a range of interesting challenges related to Natural Language Processing (NLP) and the difficulties of communicating with an entirely different species. These challenges closely mirror real-world issues in NLP, the complexity of language and the current limitations of existing technologies. Following are some of the key challenges:

### a. Contextual Ambiguity and Interpretation:

**Arrival**: One of the most exciting moments takes place when Louise Banks translates a Heptapod symbol as "offer weapon," triggering global alarm. She later explains that, depending on context, the symbol could also mean something more neutral or even positive, such as "offer tool."

**Real-world NLP**: Human language ambiguation is a core challenge in NLP. Words often have multiple meanings depending on context. NLP systems must depend heavily on

context to accurately interpret these variations. Building models capable of reliably understanding and disambiguating language based on context remains a major challenge in the field.

### b. Learning from Limited Data

**Arrival**: Louise has to build an understanding of an entirely new language with a very limited set of samples and no prior data.

**Real-world NLP**: Many Natural Language Processing applications rely on domain-specific knowledge and terminology. Acquiring labelled data for these specialized fields is often challenging. This scarcity of tailored data reduces the effectiveness of NLP system.

### c. The Role of Non-linguistic Communication

**Arrival**: Louise primarily focuses on decoding the Heptapods' written language, the film also suggests that non-linguistic cues such as gestures, plays a crucial role in communication, though it presents these elements in a simplified manner compared to their true complexity in real-world interactions.

**Real-world NLP**: Most current NLP systems are centred on text and speech, often overlooking non-verbal elements like body language, facial expressions, or tone. Grasping context is essential for NLP tasks like sentiment analysis, summarization, and language translation. Accurately capturing and representing contextual information remains a significant challenge, particularly in complex or nuanced linguistic settings.

### d. Non-Linear Language Structure

**Arrival**: The Heptapod language is written in circular symbols with no clear start or end, reflecting their non-linear perception of time.

**Real-world NLP**: NLP systems are generally built on sequential, linear models of language, however, data sparsity and inconsistency present major challenges in developing reliable NLP systems, often resulting in reduced performance in real-world applications. Modelling non-linear or highly structured communication systems is still a major challenge in the field of NLP.

## 3. Analysis of the Communication Methods

### a. Symbol Association – Tokenization and Embedding

**Arrival**: Louise must associate abstract, nonlinear symbols with real-world meanings. These symbols have no phonetic or textual parallels in human language, making the process more complex than traditional translation.

**NLP**: This is similar to tokenization and embedding strategies, where symbols (words, subwords, or characters) are mapped to vector representations that carry meaning. The

Heptapod symbols could be thought of as high-dimensional, holistic "tokens" that require advanced embedding to interpret.

### b. Unsupervised Learning

**Arrival**: Louise Banks starts without any prior knowledge of the Heptapod language and builds an understanding from scratch, based solely on exposure, observation, and pattern recognition.

**NLP**: This reflects unsupervised learning or zero-shot learning, where models must detect structure, patterns, and meaning in data without labelled examples. In modern NLP, this is same as how large language models like GPT are trained on vast amounts of unannotated text to learn representations of language.

### c. Sapir-Whorf Hypothesis:

**Arrival**: The movie strongly reflects the Sapir-Whorf hypothesis, that proposes that language shapes how we perceive reality and influences our cognitive processes. Louise's acquisition of the Heptapod language enables her to experience time in a non-linear way, fundamentally altering her perception.

**NLP and Cognitive Modeling**: This concept aligns with areas of NLP research that investigate how language impacts human thought and how such effects can be replicated in artificial intelligence. For example, tasks like sentiment analysis aim to interpret emotional nuances in language, echoing the film's central idea that language doesn't just communicate perception, it shapes it.

### d. Data Scarcity (Low-Resource Language Challenges)

**Arrival**: Louise works with extremely limited samples of the alien language, without parallel data or a lexicon.

**NLP**: This reflects low-resource language processing, where systems must learn to understand or translate languages with few available texts or labeled datasets. Few-shot, zero-shot, and transfer learning approaches aim to address this.

## 4. Self-Reflection:

In the movie Arrival, language is not only a vehicle for communication rather it shapes thought, perception, and even reality. Louise Banks' acquisition of the alien language revamps her perception of time, suggesting that the structure of a language can fundamentally alter cognition. However, most NLP models treat language statistically as patterns in data without embedding the idea that language can alter cognition or worldview.

The language of alien in the movie is non-linear i.e. it is written as a complete whole rather than sequentially. This suggests a different logic to how language can be structured and understood. The film opens the possibility of non-linear language processing, models that

don’t just predict the next word, but can construct meaning from an entire context or intended message. This could inspire new NLP directions focused on global message Modeling, perhaps drawing from techniques in graph-based or top-down generative models.

In the film, misinterpretation of alien communication nearly leads to global war. In the movie the language became a political, existential tool as it required patience, care, and openness. While today NLP systems are now used in sensitive domains such as diplomacy, law, healthcare etc. where misunderstanding can have major consequences. The film warns of the dangers of overconfidence in machine interpretation and reminds us of the importance of human-in-the-loop systems, ethical design, and interpretability. It also highlights the importance of cross-linguistic empathy and cultural understanding, which current NLP models are not yet designed to handle with nuance.