

# The Interplay of Linguistic Relativity and Natural Language Processing in "Arrival"

*By*

Olugbenga Adegoroye, Jazmine Brown, Varit (Henry) Kobutra, Kimberly Navarrete,  
Phillip Torres

Houston Community College

February 2025

Prof. Patricia McManus

ITAI 2376

## **Introduction**

The 2016 Denis Villeneuve's film *Arrival* dive into the challenges of language and communication when extraterrestrial beings known as the Heptapods arrive on Earth. The movie's exploration of linguistic barriers, perception, and cross-cultural understanding offers a thought-provoking perspective on the complexities of communication. At its core, *Arrival* addresses fundamental questions about how language shapes human thought and perception, mirroring the challenges encountered in the field of Natural Language Processing (NLP). This report examines the film's portrayal of linguistic challenges and their parallels to real-world natural language processing (NLP) obstacles. We will explore the ambiguity, cultural context, and syntactic complexity. By analyzing the methods used to bridge human-alien communication, we draw connections to rule-based, statistical, and deep learning approaches in NLP. The film's narrative underscores the Sapir-Whorf hypothesis, emphasizing how language structures a concept that challenges and expands contemporary NLP paradigms.

## **Linguistic Relativity and the Sapir-Whorf Hypothesis in "Arrival"**

### **Non-Linear Language and Temporal Perception**

The heptapods' circular logograms, which lack a linear sequence, directly challenge human linguistic frameworks. Their written language embodies a holistic understanding of time, where past, present, and future coexist. This aligns with the *strong* version of the Sapir-Whorf hypothesis linguistic determinism which posits that language shapes cognitive processes. Louise's acquisition of this language rewires her brain to perceive time non-linearly, enabling precognitive visions. In NLP, this mirrors the challenge of processing languages with free word order (e.g., Latin) or context-dependent syntax, where traditional rule-based systems struggle without explicit structural cues.

### **Ambiguity and Contextual Dependency**

The film highlights the ambiguity inherent in deciphering alien symbols. Early misinterpretations of the heptapods' intent such as mistaking "tool" for "weapon" reflect real-world NLP issues with polysemy and contextual disambiguation. Human languages often rely on context to resolve meanings (e.g., "bank" as a financial institution vs. a riverbank), a task requiring advanced neural models to replicate. Louise's iterative dialogue with the heptapods, akin to reinforcement learning, demonstrates how feedback loops refine interpretations, paralleling modern NLP's use of transformer models to infer context.

## **NLP Challenges Exemplified in "Arrival"**

### **Cross-Cultural and Cross-Linguistic Barriers**

The heptapods' lack of human-like sensory organs and cultural frameworks complicates communication. Their logograms convey entire concepts rather than phonetic sounds, necessitating a shift from sequential to spatial analysis. This mirrors NLP's difficulties with low-resource languages, where limited datasets and cultural nuances hinder model training. For instance, idiomatic expressions (e.g., "break a leg") require cultural knowledge that rule-based systems often lack. The film's global coordination breakdown as nations withhold data parallels the "data scarcity" problem in NLP, where siloed resources impede multilingual model development.

### **Syntax vs. Semantics**

The heptapods' language divorces syntax from semantics, as logograms encapsulate meaning independent of word order. This contrasts with human languages, where syntax heavily influences interpretation (e.g., "dog bites man" vs. "man bites dog"). NLP systems must balance syntactic rules with semantic intent, a challenge evident in sentiment analysis. For example, sarcasm detection requires understanding tonal cues absent in text, a task poorly handled by rule-based models. Louise's focus on semantic intent over syntactic structure mirrors hybrid NLP approaches that combine statistical patterns with deep learning to capture nuanced meaning.

### **Communication Methods and NLP Parallels**

#### **Rule-Based Foundations and Iterative Learning**

Louise begins with rule-based methods, establishing basic vocabulary through repeated exchanges.

This mirrors early NLP systems like ELIZA, which used predefined templates to simulate conversation. However, rule-based models falter with complexity; the heptapods' logograms require probabilistic reasoning akin to statistical NLP, where patterns emerge from frequency data. The film's breakthrough recognizing logograms as unified concepts reflects the shift to neural networks, which process language holistically rather than through isolated rules.

## **Empathy and Cultural Context**

Louise's success hinges on understanding the heptapods' intent rather than merely translating symbols. This underscores NLP's growing emphasis on contextual embeddings (e.g., BERT), which capture situational and cultural nuances. The film critiques humanity's hostile attribution bias, a reminder that NLP systems must mitigate biases ingrained in training data. For instance, sentiment analysis models often misclassify texts from underrepresented groups due to skewed datasets.

## **Implications for NLP: Beyond Syntax to Cognitive Integration**

### **Temporal Modeling and Non-Linear Processing**

The heptapods' language challenges linear temporal models, suggesting NLP systems could benefit from recursive neural networks (RNNs) that handle sequential data dynamically.

Current models like GPT-4 struggle with long-term context retention, a limitation highlighted by Louise's ability to integrate past, present, and future insights. Future NLP architectures might incorporate temporal coherence mechanisms, akin to the film's circular logograms, to improve contextual understanding.

## **Ethical and Collaborative Frameworks**

The film's geopolitical tensions underscore the need for ethical NLP practices, including transparent data sharing and bias mitigation. Just as global cooperation resolves the heptapod crisis, multilingual NLP requires collaborative efforts to pool resources and cultural expertise. Projects like Meta's No Language Left Behind exemplify this, yet challenges persist in scaling for 7,000+ languages.

## **Conclusion: Language as a Bridge and a Lens**

In conclusion, *Arrival* presents a profound exploration of the complexities of language and communication, highlighting challenges that are deeply relevant to the field of Natural Language Processing. The film's portrayal of the difficulties in understanding an extraterrestrial language reflects the real-world obstacles that NLP faces, such as ambiguity, idiomatic expressions, and cultural differences. Furthermore, the movie's depiction of language as a tool that shapes perception and cognition challenges current NLP paradigms and suggests that future systems may need to account for these deeper aspects of human language. By analyzing these parallels, we can gain a better understanding of both the potential and the limitations of NLP technologies, as well as

how they may evolve to meet the challenges of more complex, human-like communication.

## References

Paramount Pictures. (2016). *Arrival* [DVD]. Hollywood, CA.

Chen, M., Zhang, X., & Li, L. (2020). Domain adaptation in natural language processing: A survey. *ACM Computing Surveys (CSUR)*, 53(1), 1-28.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26.

Navigli, R. (2009). Word sense disambiguation: A survey. *ACM Computing Surveys (CSUR)*, 41(2), 1-69.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.

Wu, Y., Zhang, Y., & Chen, D. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *Google Research Blog*.

Villeneuve, D. (Director). (2016). *Arrival* [Film]. Paramount Pictures.