JPC1: Programmable Semantic Shift Analysis using ChainNet and CoreLex

Barbora Mahdalová (bmahdalova)
mahdba00@upol.cz

Dominik Fránek (dfranek161)
frando02@upol.cz

February 26, 2026

1 Introduction

The goal of this task was to investigate lexical–conceptual mappings in a programmable way, with a focus on semantic relationships between categories such as PERSON and ANIMAL. While the assignment suggested approaches such as manual inspection of dictionary entries or cross-linguistic comparison, we approached the problem computationally.

Our objective was to automatically detect semantic category shifts (e.g. ANIMAL → HUMAN) using structured lexical resources. Specifically, we integrated:

- ChainNet – which encodes metaphorical relations between WordNet senses
- CoreLex – which assigns WordNet noun senses to coarse semantic types

By combining these resources, we aimed to identify not only whether a semantic shift occurs, but also what conceptual mapping it represents.

2 Background

This work builds on recent research showing that tropes such as metaphor can be analyzed computationally using structured lexical data.

Resources

We used three primary resources:

- **OMW English 1.4 (ChainNet)** – a version of WordNet extended with metaphor relations between senses.
- **CoreLex** – an ontology mapping WordNet noun senses into 39 semantic types such as ANIMAL, HUMAN, ARTIFACT.
- **Python wn library** – used to load and query the ChainNet XML lexicon.

These resources allowed us to move from manual analysis toward a fully programmable pipeline.

3 Approach

Our approach evolved during development.

An earlier version relied on extracting synset offsets and aligning them via the NLTK interface. This proved unreliable due to version mismatches between resources.

The final pipeline instead uses a sense–index mapping strategy directly within the wn environment.

The implemented system (analyze_shifts_final.py) performs the following steps:

1. Loads the ChainNet XML lexicon into the wn database
2. Retrieves noun senses for a set of polysemous words
3. Extracts metaphor relations between senses
4. Dynamically resolves metaphor targets, even when they belong to different lemmas
5. Maps sense index positions to CoreLex semantic types

Only forward metaphor relations were traced in order to avoid double-counting bidirectional links.

4 Results

The system identified 11 distinct semantic shifts.

Word	Source Type	Target Type
Chestnut	NATURAL_BODY	PSYCHOLOGICAL_FEATURE
Star	NATURAL_BODY	FORM
Mouth	PART	SPACE
Mouth	PART	PHYSICAL_OBJECT
Mouth	PART	ARTIFACT
Arm	PART	ARTIFACT
Leg	PART	ARTIFACT
Dog	ANIMAL	HUMAN
Rat	ANIMAL	ARTIFACT
Snake	NATURAL_BODY	PHYSICAL_OBJECT
Hawk	ANIMAL	HUMAN

The most frequent shift patterns were:

- PART → ARTIFACT
- ANIMAL → HUMAN
- NATURAL_BODY → PHYSICAL_OBJECT / FORM

5 Discussion

The results reveal systematic conceptual mappings.

Mappings from PART to ARTIFACT illustrate how human anatomy serves as a structural model for objects. For example, containers may have a “mouth” and furniture may have “legs”.

Mappings from ANIMAL to HUMAN rely on behavioral traits rather than physical resemblance, as seen in metaphorical uses of words such as *hawk*.

An inconsistency was observed in the classification of *chestnut*, where a physical anatomical feature was mapped to a psychological category. This highlights limitations of fully automated semantic ontologies.

All observed shifts moved from more concrete domains toward more abstract or socially interpreted ones, supporting the directional nature of metaphor.

6 Conclusions

This project demonstrates that semantic shifts can be detected automatically by integrating structured lexical resources. By refining the mapping between ChainNet and CoreLex using sense-index alignment, we developed a robust pipeline capable of identifying conceptual mappings such as ANIMAL $\rightarrow$ HUMAN.

Future work could extend this approach cross-linguistically or explore connections between semantic shift and morphological processes such as zero-derivation.

Author Contributions

The code represents collaborative work resulting from extensive brainstorming and problem-solving. The collaboration included understanding the assignment objectives, analyzing and reconciling the ChainNet and CoreLex datasets, designing and implementing the semantic-shift detection pipeline, and verifying the results to ensure accuracy and completeness.

The original report was written by Dominik Fránek and subsequently refined to its final form by Barbora Mahdalová.

While the computational analysis and report content were developed independently by the authors, minor assistance from AI tools was used to improve language clarity and formatting.

Code Availability

The script `analyze_shifts_final.py` and associated data files are included in the project repository.