Semantic subcategorisation for creative generation of light verb constructions

Lin de Huybrecht

Promotor: Prof. Dr. Dr. Geraint Wiggins

Vrije Universiteit Brussel

computational creativity lab

Flanders AI Research Program

ARTIFICIAL INTELLIGENCE RESEARCH GROUP

Artificial Intelligence Vlaanderen/Flanders

Why do we say

*"I will **grab a shower**"*,

but not

*"I will **grab a bath**"*?

We need transparent language models
to enable us to reason about them

grounded in **linguistics**

language = syntax + semantics

and **cognitive science**

meaning = use

(Wittgenstein, 1953)

Distributional

Compositional

Categorical

framework

The DisCoCat framework explicitly models both the syntax and semantics of language

DisCoCat framework

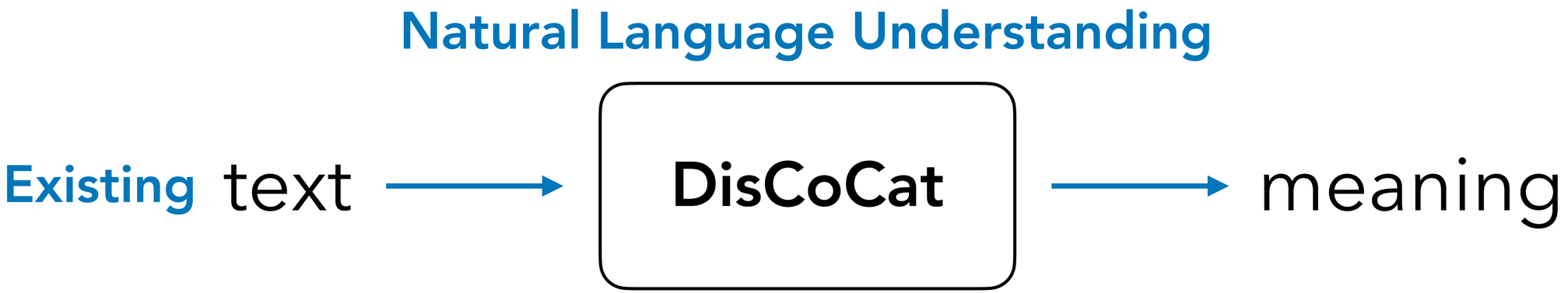
language = syntax × semantics

sentence embedding

word type

word embedding

The DisCoCat framework has been successfully used for natural language understanding



I want to extend the DisCoCat framework
for creative natural language generation

I am studying which kinds of semantic representations are needed for text generation

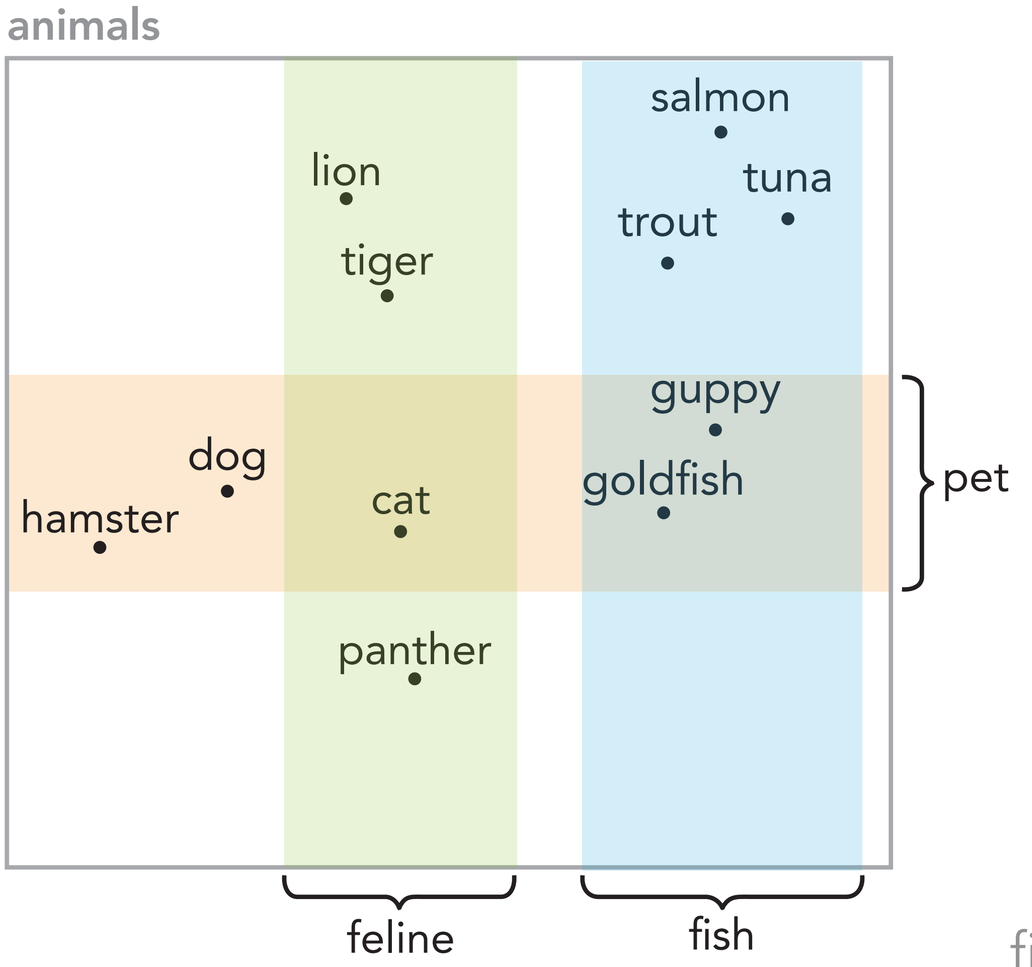


figure based on McGregor et al. (2020)

Semantic subcategorisation for creative generation of light verb constructions

1. **What are light verbs?**
2. What is my goal?
3. What is my approach?

full verb construction

Alice **read** the paper.

full verb construction

Alice **read** the paper.

light verb construction

Alice **gave** the paper **a read**.

Light verb constructions have a verb that is semantically bleached or "lightened"

Alice **read** the paper.

Alice **gave** the paper **a read**.

Why should we care about light verb constructions?

Light verb constructions impose various syntactic restrictions

give + { kiss, kick, hug } — contact nouns

give + { scare, fright, tickle } — stimulus-experiencer nominalisations

give an interest

take an interest

(Wittenberg, 2016)

Light verb constructions are perceived differently than their full verb counterparts (Wittenberg & Levy, 2017)

Durative event

light verb construction

Alice **gave** the paper **a read**.

mass syntax

<

full verb construction

Alice **read** the paper.

transitive form

perceived as taking less time

Application in second language learning (UC5):

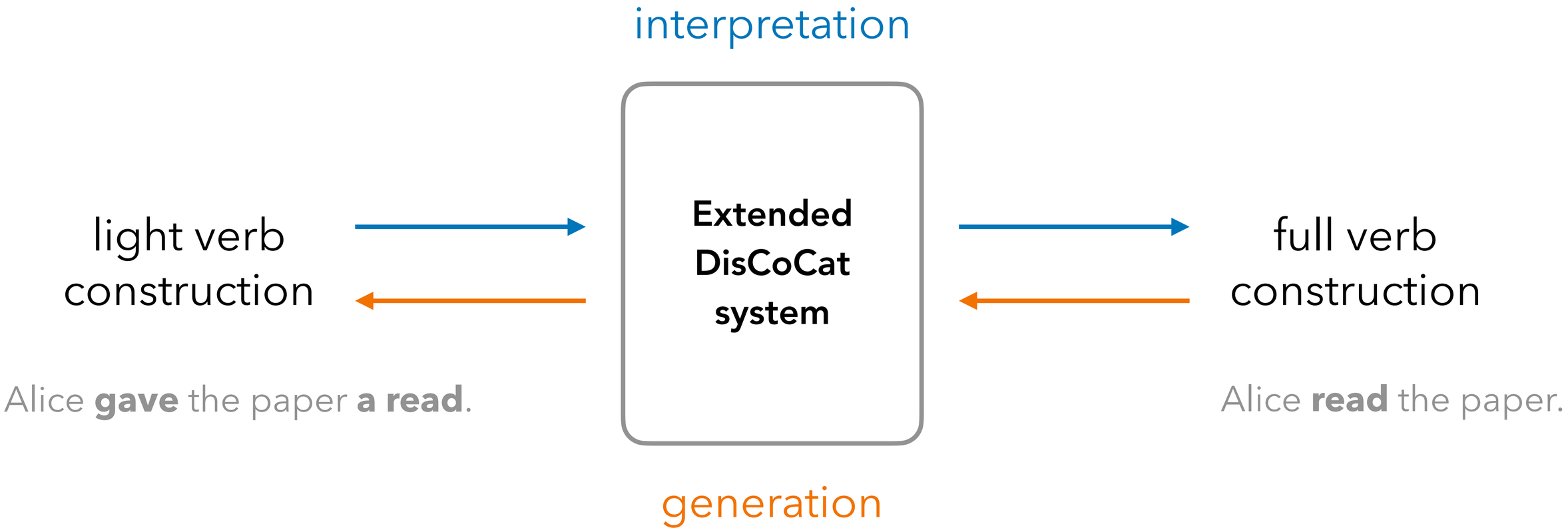
Non-native English speakers use more light verb constructions as their English proficiency increases

(Guilquin, 2019)

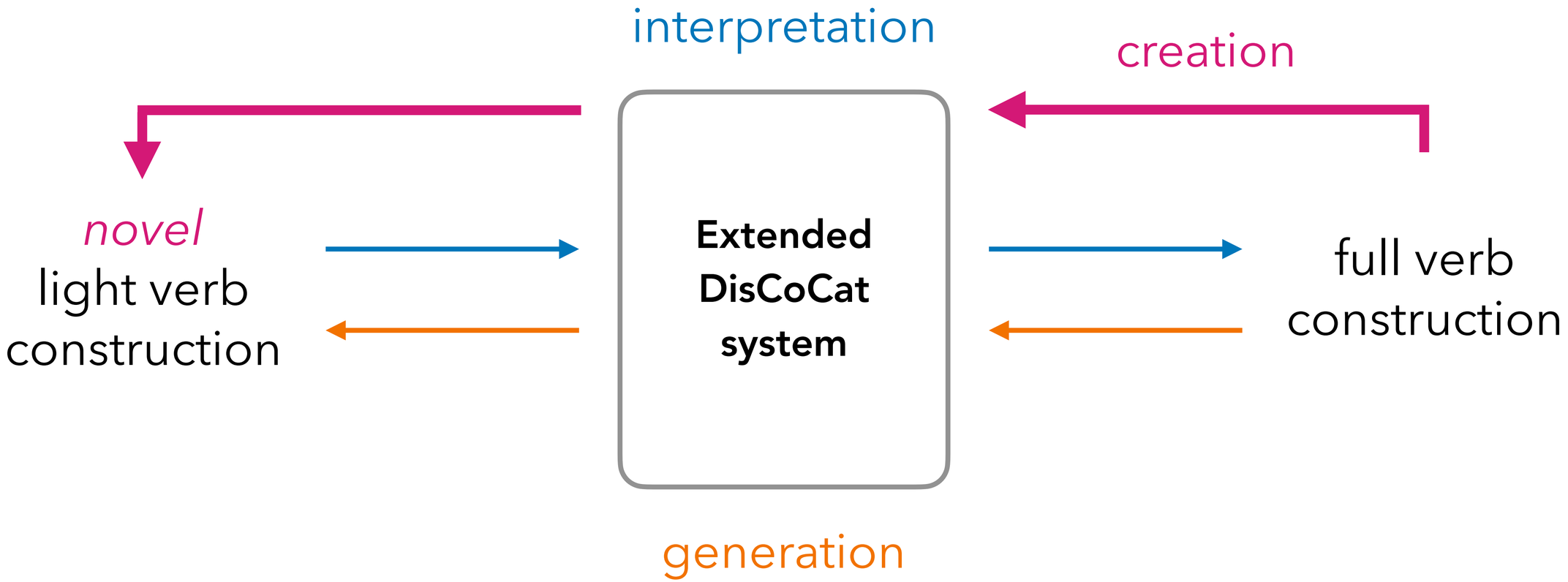
Semantic subcategorisation for creative generation of light verb constructions

1. What are light verbs?
2. **What is my goal?**
3. What is my approach?

First we want to convert between light verb & full verb constructions



The ultimate goal is to develop a system that can generate novel light verb constructions



Semantic subcategorisation for creative generation of light verb constructions

1. What are light verbs?
2. What is my goal?
3. **What is my approach?**

Classical part-of-speech categories are insufficient for text generation

Selectional restrictions

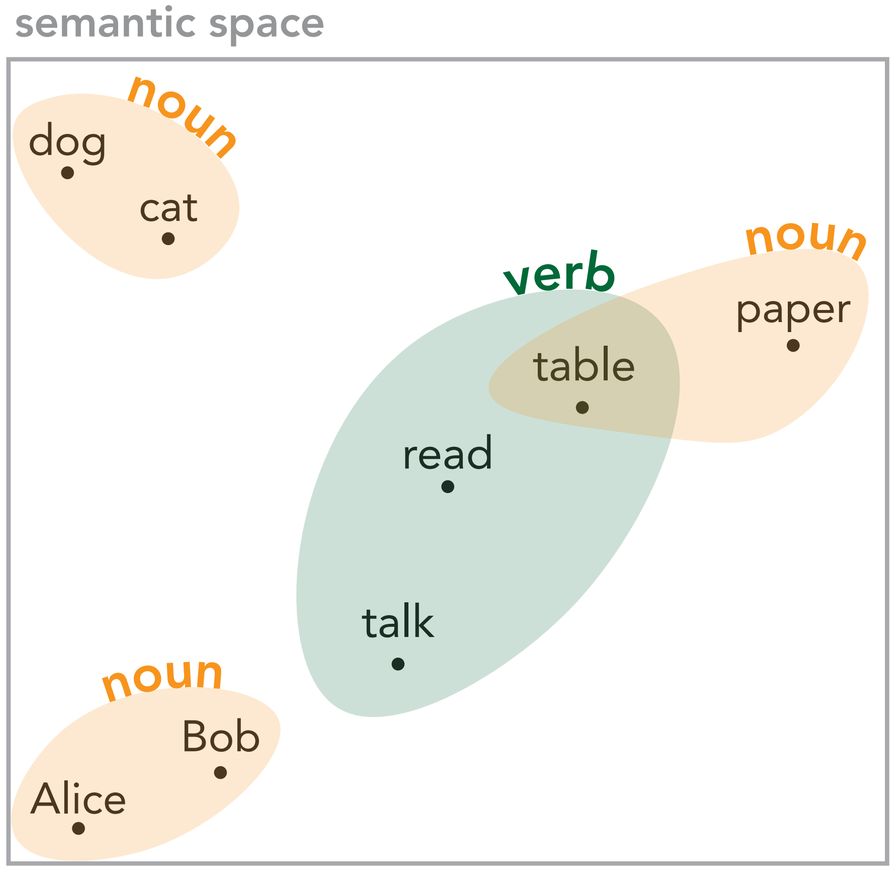
Alice read a paper.

Alice read a **table**.

Alice **tabled** a paper.

Alice talks.

The **table** talks.



We need a more fine-grained subcategorisation of word types

Selectional restrictions

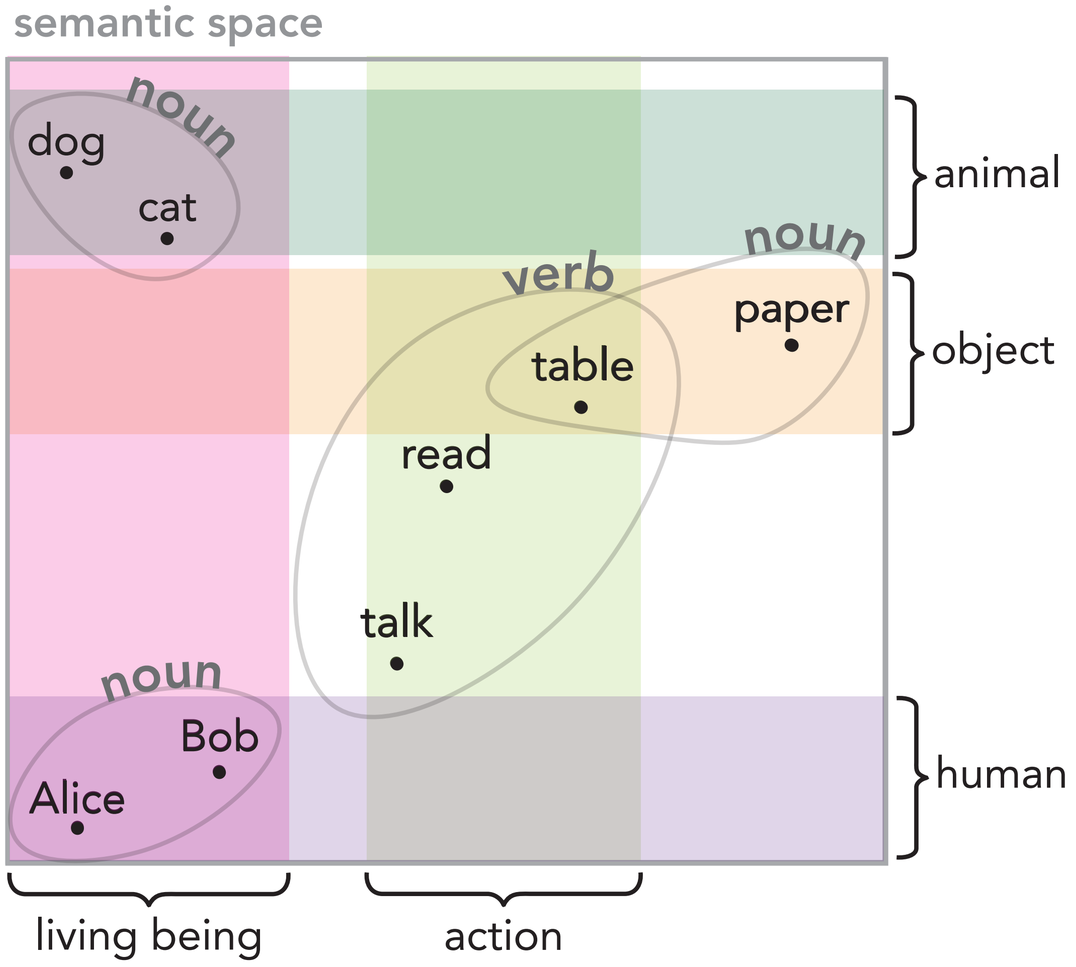
Alice read a paper.

Alice read a **table**.

Alice **tabled** a paper.

Alice talks.

The **table** talks.



Semantic subcategorisation for creative generation of light verb constructions

1. What are light verbs?

   semantically bleached verbs
   interesting syntactic & semantic properties

2. What is my goal?

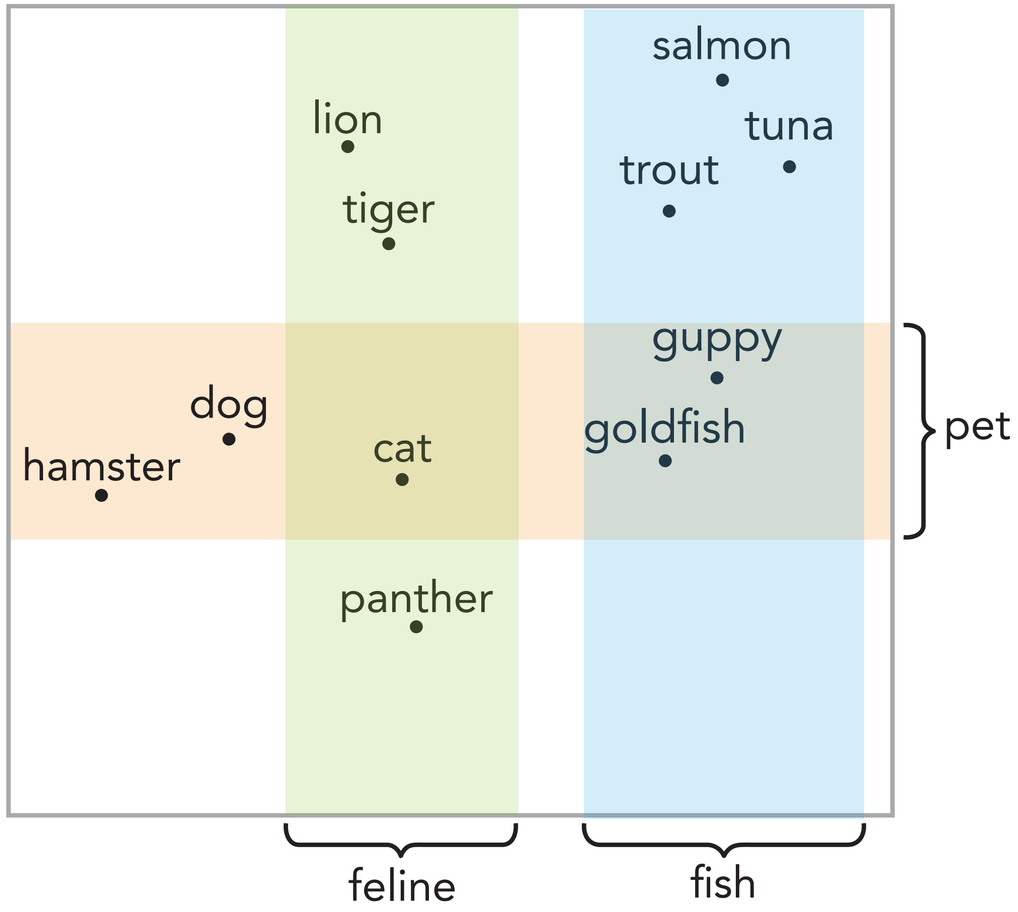
   interpret, generate & create new LVCs

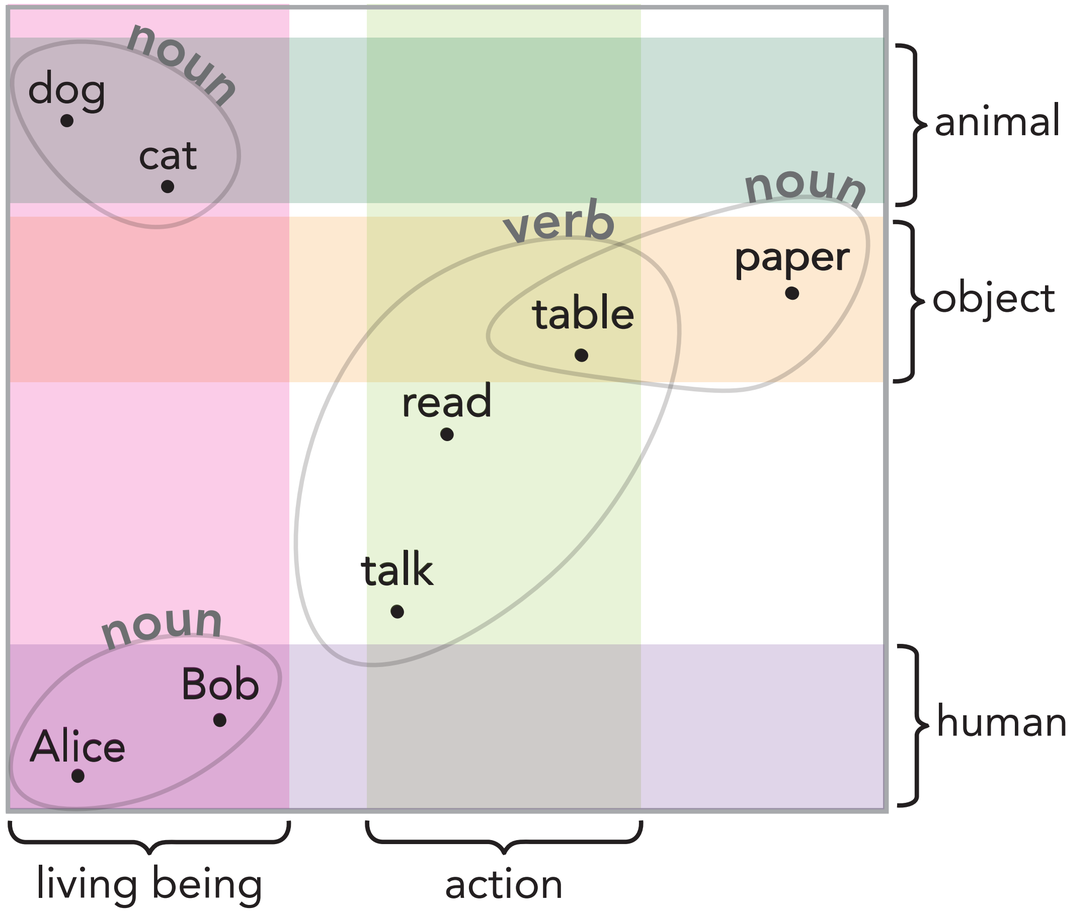
3. What is my approach?

   subcategorisations of semantic and syntactic types

This research will give us a deeper understanding
of how semantic representations interact to convey meaning in language

This will contribute towards more transparent and explainable technology





One last example of light verb constructions:

If this ***sparks an interest***,

come and ***take a look*** at my poster

so we can ***have a chat***!