

German Light Verb Constructions in Business Process Models

Kristin Kutzner • Kristin.Kutzner@fh-zwickau.de
Ralf Laue • Ralf.Laue@fh-zwickau.de

Background of our Research

- **Business process models (BPMs)** are visual representations of processes in an organization. Such models can be created using the visual modelling language Event-Driven Process Chain (EPC). A simple EPC consists of functions and events which contain textual labels (Figure 1).
- The textual labels of models often contain errors, such as tasks described too vaguely or inconsistent naming. To identify and counteract these errors, **approaches for the automatic model analysis** have been developed.
- In particular, **light verb constructions (LVCs)** constitute a major challenge in automatic model analysis.

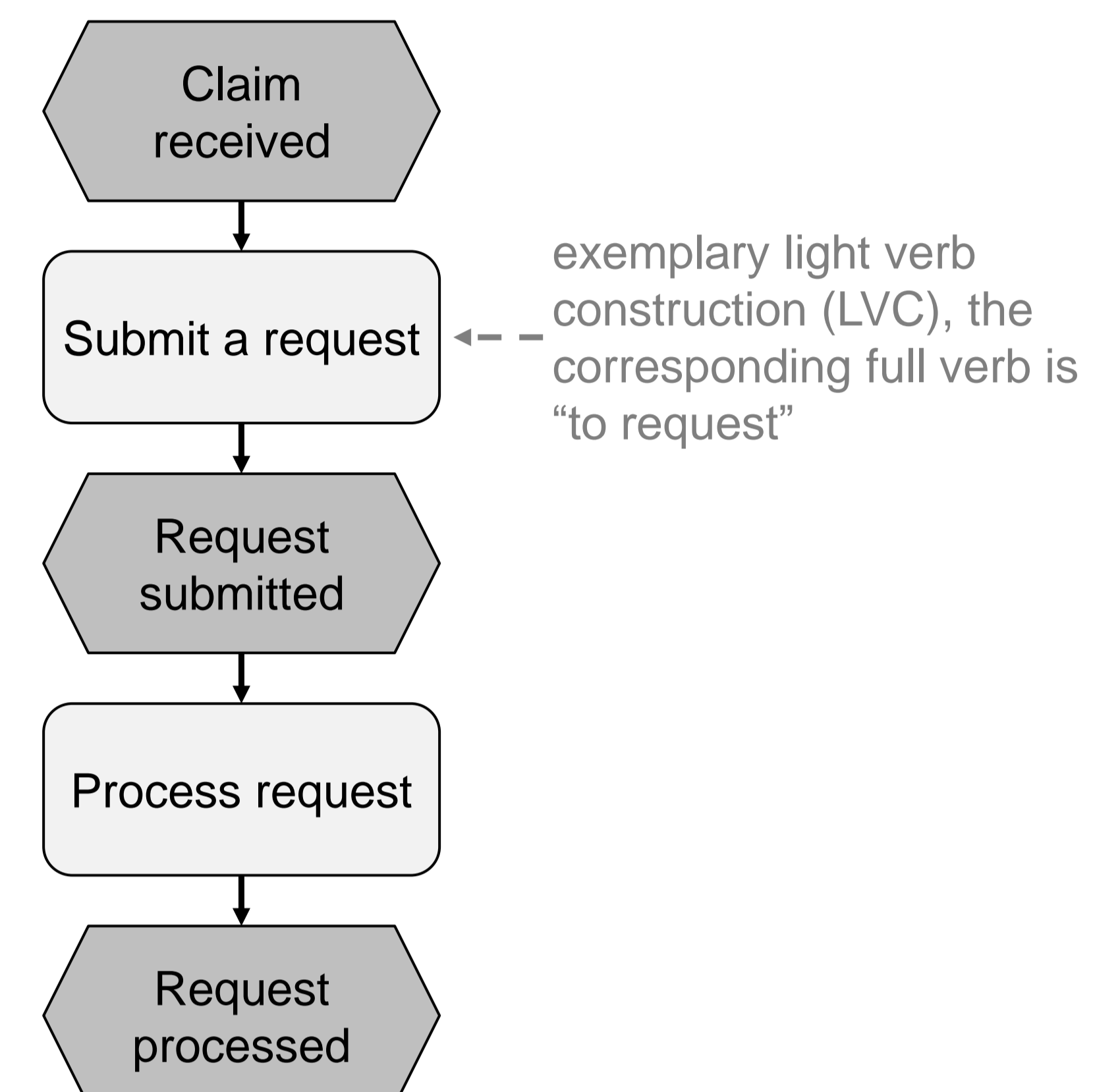


Figure 1: EPC model fragment

Research Design

- To identify which LVCs are used in BPMs, we used a **three-staged research design** (Figure 2).
- We collected and extracted BPMs (Stage 1), prepared a text corpus (Stage 2) and manually identified LVCs (Stage 3).

	Inputs	Steps	Outputs
Stage 1: Collecting and extracting BPMs	German EPC models from public repositories, textbooks, scientific papers, student papers and real-world projects	Extraction of models from the sources and conversion into one uniform file format	2,301 extracted and unified German EPC models
Stage 2: Preparing a text corpus	2,301 extracted and unified German EPC models	Preparation of a text corpus by implementing python scripts and using NLTK	Linguistic annotated BPM corpus
Stage 3: Investigating the text corpus and identifying LVCs	Linguistic annotated BPM corpus	Listing of all verbs with the associated text passages and manual identification of LVCs	Resource of German LVCs in BPMs

Figure 2: Research design

Results (Excerpt)

events	31,867
functions	21,096
in the labels of those events and functions:	
tokens	183,611
words	176,821
among those tokens were:	
nouns (NN & NE)	77,583
adjectives (ADJA & ADJD)	12,934
numbers (CARD)	2,318
verbs (V*)	57,178
articles (ART)	5,704
pronouns (P*)	1,441
adverbs (ADV)	996
conjunctions (K*)	2,924
pre- and postpositions (A*)	12,321
particles (PTK*)	5,663
others	9,259

Table 1: Results of POS-tagging

Antrag stellen (to file an application)	159
Plan erstellen (to draw up a plan)	148
Planung durchführen (to plan)	129
Zeit festlegen (to set time)	77
Prüfung durchführen (to conduct an examination)	74
Bewertung durchführen (to assess)	63
aktenkundig machen (to record something)	60
Rechnung durchführen (to compute)	59
Nummer vergeben (to assign a number)	55
Prüfung ablegen (to take an examination)	51
Analyse durchführen (to carry out an analysis)	50

Table 2: LVCs with at least 50 occurrences

Future Steps

- We will analyze further BPMs, also in other modelling languages (BPMN, OEPC) to identify additional LVCs.
- We will derive synonym sets of LVCs with the corresponding full verbs. Each synonym set will be assigned to a concept. Then, we will be able to identify more abstract concepts from textual labels.