

Ad Hoc Compounds for Stance Detection

Qi Yu¹, Fabian Schlotterbeck², Hening Wang², Naomi Reichmann¹,
Britta Stolterfoht², Regine Eckardt¹, Miriam Butt¹

¹ University of Konstanz ² University of Tübingen

MWE-UD @ LREC-COLING 2024

Funded by

DFG Deutsche
Forschungsgemeinschaft

German Research Foundation

EXC-2035/1 – 390681379

In a Nutshell

This work is about a subclass of multi-word expressions:
ad hoc compounds in German

Examples from social media and newspapers:

(1) Flüchtlinge wollen Österreich meiden und lieber in **Merkel-Land** einreisen.

‘Refugees want to avoid Austria and instead enter **Merkel Country.**’

cf. ‘Refugees want to avoid Austria and instead enter **Germany.**’

(2) Jede 5. **China-Maske** ist unbrauchbar

‘Every fifth **China-mask** is unusable’

cf. ‘Every fifth **Chinese mask** is unusable’

(3) Neue Stelle für **Kopftuch-Praktikantin**

‘New position for **hijab-intern**’

cf. ‘New position for **hijab-wearing intern**’

Source: (1) – Facebook; (2), (3) – BILD

**DONAUDAMPFSCHIFFFAHRTS
ELEKTRIZITÄTENHAUPTBET
RIEBSWERKBAUUNTERBEAM
TENGESELLSCHAFT**

YAY GERMAN!

Bad news:

German has crazy compounds...



Good news:

you are free to invent new ones.

Roadmap

BACKGROUND

- ❖ Compound meaning in theoretical linguistics
- ❖ Automated detection of compounds in NLP

CORPUS STUDY

- ❖ Evidence of a systematic use of ad hoc compounds to express negative stance in newspapers from a dataset of >8 million tokens

PSYCHOLINGUISTIC EXPERIMENT

- ❖ Validating the stance-triggering effect of ad hoc compounds with ratings by 212 German native speakers on items extracted from corpora

SIMULATIONS WITH LARGE LANGUAGE MODELS

- ❖ Experiments with GPT-4 and GPT-3.5-turbo on their ability of identifying attitudinal meanings conveyed by ad hoc compounds

CONCLUSIONS AND OUTLOOK

- ❖ Recommendations for directions of future work

BACKGROUND



CORPUS STUDY



PSYCHOLINGUISTIC EXPERIMENT



SIMULATIONS WITH LARGE LANGUAGE MODELS



CONCLUSIONS AND OUTLOOK

- ❖ Compound meaning in theoretical linguistics
- ❖ Automated detection of compounds in NLP
- ❖ Evidence of a systematic use of ad hoc compounds to express negative stance in newspapers from a dataset of >8 million tokens
- ❖ Validating the stance-triggering effect of ad hoc compounds with ratings by 212 German native speakers on items extracted from corpora
- ❖ Experiments with GPT-4 and GPT-3.5-turbo on their ability of identifying attitudinal meanings conveyed by ad hoc compounds
- ❖ Recommendations for directions of future work

Attitudinal Meaning From Ad Hoc Compounds

A closer examination of the aforementioned examples:

(1) a. Flüchtlinge wollen Österreich meiden und lieber in
Merkel Land einreisen.
'Refugees want to avoid Austria and instead enter
Merkel Country.'



Attitudinal meaning:
the German refugee crisis is Merkel's fault

b. 'Refugees want to avoid Austria and instead enter Germany.' ➔ Attitudinal meaning present? 🤔
(Empirical validation follows later)

(2) a. Jede 5. China-Maske ist unbrauchbar
'Every fifth China-mask is unusable'



Attitudinal meaning:
China is notorious for low-quality products

b. 'Every fifth Chinese mask is unusable' ➔ Attitudinal meaning present?

Compound Meaning in Theoretical Linguistics

Compounds have a range of interpretational possibilities as their meanings are not determined compositionally.

Different types of ad hoc compounds:

– Purely referential ad hoc compounds:

- Serve as abbreviations of phrases for the aim of concise presentation
- Refer neutrally to the referent

(4) *Karajan-Schüler* ‘Karajan Student’ vs. *Schüler von Karajan* ‘student of Karajan’

– Enigmatic compounds (Wildgen 1981):

- Refer to the referent in a metaphoric / an expressive manner (“enigmatic” = referring in an obscure manner)
- Mostly have neutral alternatives → the more “marked” compounding form have appraisive or manipulative effects

(5) *Allesbestimmerpartei* ‘decide-everything party’: referring to the ruling party of East Germany

Focus of our work: **enigmatic compounds** – we explore their potential for automatized stance detection

Research Gaps and Challenges in NLP

- Analysis of subjective language is an important part of stance detection (Wiebe et al., 2004). However, there are to date no work on stance detection focusing on compounds
- Existing tools for the detection of German subjective language (El-Assady et al., 2016, 2019) do not include the detection of (enigmatic) compounds

Challenges:

- One possible clue for compounds: hyphens (e.g., *Merkel-Land*, *China-Maske*, *Kopftuch-Praktikantin*)
- However, most (established) German compounds do not include a hyphen
e.g., *Flüchtlingsorganisation* 'refugee organization', *Migrantenschreck* 'migrant scare'
- Even hyphenated compounds cannot be identified by dependency parsers (e.g., Mate Tools, spaCy, ParZu, SMOR). Most of them label the compounds simply as common nouns

(6) *Politik: Flüchtlingsorganisation gegen Asyl-Verschärfungen.*
'Politics: Refugee organization against tightening of asylum laws'



BACKGROUND



CORPUS STUDY



PSYCHOLINGUISTIC EXPERIMENT



SIMULATIONS WITH LARGE LANGUAGE MODELS



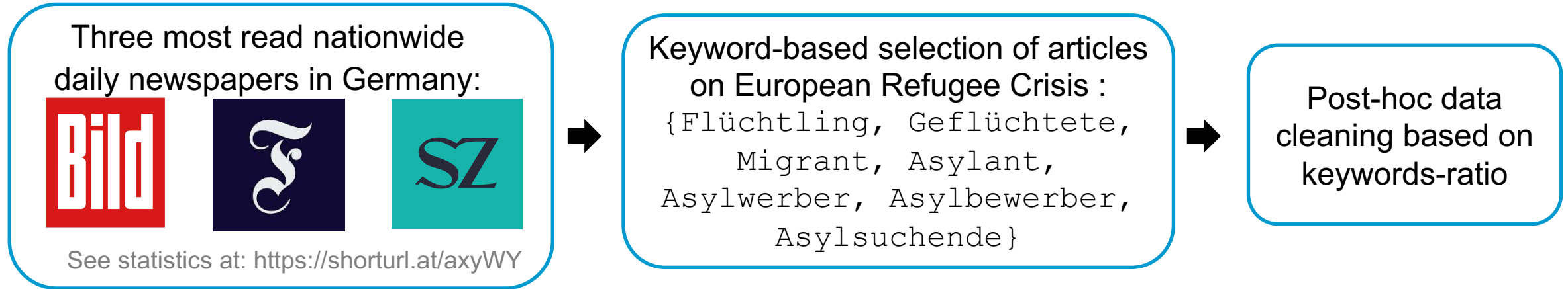
CONCLUSIONS AND OUTLOOK

- ❖ Compound meaning in theoretical linguistics
- ❖ Automated detection of compounds in NLP
- ❖ Evidence of a systematic use of ad hoc compounds to express negative stance in newspapers from a dataset of >8 million tokens
- ❖ Validating the stance-triggering effect of ad hoc compounds with ratings by 212 German native speakers on items extracted from corpora
- ❖ Experiments with GPT-4 and GPT-3.5-turbo on their ability of identifying attitudinal meanings conveyed by ad hoc compounds
- ❖ Recommendations for directions of future work




Dataset

Study case: framing in the discourse of the event “European Refugee Crisis” (2014-2018)

Data collection:



Summary of final dataset:

newspaper	category	# articles	# tokens
 BILD	tabloid, right-leaning	12,107	3,065,065
 Frankfurter Allgemeine Zeitung (FAZ)	broadsheet, right-leaning	6,686	3,332,444.
 Süddeutsche Zeitung (SZ)	broadsheet, left-leaning	4,536	1,757,929

Manual Annotation of Compounds

Material: headlines of all news articles in our corpus

(Using headlines only: compounds are more likely to be found there)

Criteria for a compound to be labeled as enigmatic:

- The compound carries an attitudinal meaning
- The compound is an ad hoc formation, thus not established in a recognized German dictionary/lexicon

(7) a. *Flüchtlings-Tsunami* ‘refugee tsunami’

✓ (carries an attitudinal meaning that the refugee influx is a threat)

b. *Karajan-Schüler* ‘Karajan student’

✗ (is an ad hoc formation, but does not carry attitudinal meaning)

Corpus Study

Statistical summary:

Newspaper	#Enigmatic	#Purely Referential
BILD	726	10,059
FAZ	58	5,525
SZ	44	3769

Qualitative investigation:

- Enigmatic compounds in BILD show a focus on issues of criminality
e.g., *Asylprügler* ‘asylum beater’, *Migrantenschreck* ‘migrant scare’, *Amok-Afrikaner* ‘amok African’,
- Enigmatic compounds in FAZ and SZ focus rather on...
 - problems of capacity: e.g., *Flüchtlingsandrang* (FAZ) *Flüchtlingsansturm* (SZ) ‘refugee onrush’,
 - rights of individual refugees: e.g., *Flüchtlingskind* / *Flüchtlingsjunge* / *Flüchtlingsmädchen* (SZ/ FAZ) ‘refugee child / refugee boy / refugee girl’

BACKGROUND



CORPUS STUDY



PSYCHOLINGUISTIC EXPERIMENT



SIMULATIONS WITH LARGE LANGUAGE MODELS



CONCLUSIONS AND OUTLOOK

- ❖ Compound meaning in theoretical linguistics
- ❖ Automated detection of compounds in NLP
- ❖ Evidence of a systematic use of ad hoc compounds to express negative stance in newspapers from a dataset of >8 million tokens
- ❖ Validating the stance-triggering effect of ad hoc compounds with ratings by 212 German native speakers on items extracted from corpora
- ❖ Experiments with GPT-4 and GPT-3.5-turbo on their ability of identifying attitudinal meanings conveyed by ad hoc compounds
- ❖ Recommendations for directions of future work

Experimental Setup

Data: 21 text snippets containing enigmatic compounds that intuitively trigger negative attitudinal meanings collected from newspapers and social media; 24 fillers

(Only considering enigmatic compounds with negative attitudinal meanings: these were more prevalent in the corpus study)

3 conditions created from each snippet:

Condition	Example
COMPOUND	'Every 5th China-mask is unusable'
PHRASAL	'Every 5th Chinese mask is unusable'
NEUTRAL	'Every 5th mask is unusable'

Null-Hypotheses:

COMPOUND vs. PHRASAL (same information load):

Compounds do not amplify the perceived attitudinal strength.

COMPOUND / PHRASAL vs. NEUTRAL (different information load):

The perceived attitudinal strength is not amplified by extra information unnecessary for reference resolution.

Participants' task: *How does the author talk about <THE_REFERENT>? (e.g., the masks)*

POSITIVE 1 2 3 4 5 6 7 NEGATIVE

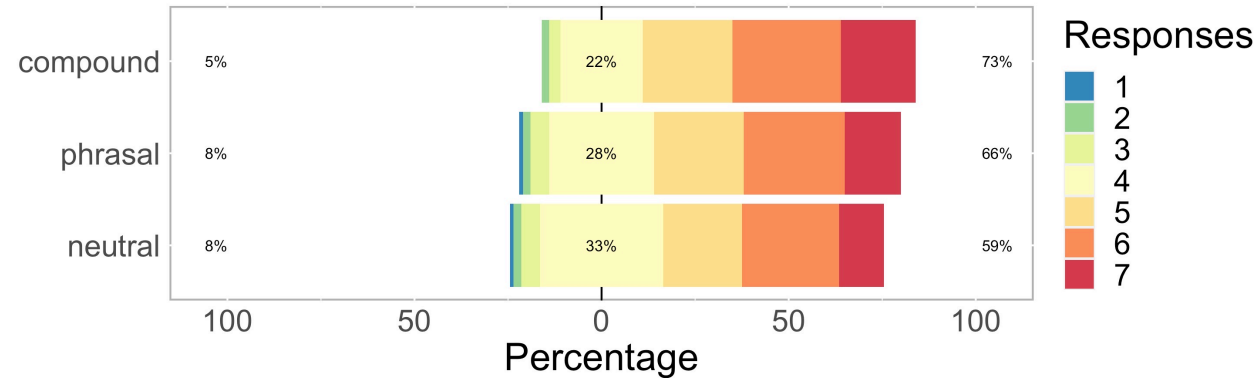
Meta-data collection: *Where would you place your own political leaning?*

LEFT 1 2 3 4 5 6 7 RIGHT

Results and Discussion

Participants: 212 German native speakers (103 female, 102 male, 7 other; mean age = 26.52 years, $SD = 8.10$ years)

Rating distribution:



Regression analysis with cumulative link model:

- Cumulative link model: a variant of logistic regression generalized to multinomial ordinal predictors

$$\text{logit}(P(Y \leq j)) = \theta_j - \mathbf{x}^T \boldsymbol{\beta}$$

Model 1: `rating ~ condition + (condition|item) + (condition|participant)`

Model 2: `rating ~ condition + political_leaning + (condition|item) + (condition|participant)`

Model 3: `rating ~ condition * political_leaning + (condition|item) + (condition|participant)`

PHRASAL as
reference level

Mapped into three levels:
1-3=left, 4=neutral, 5-7=right

Random intercepts and random slopes for
text snippets and participants

Results and Discussion

Model 1: `rating ~ condition + (condition|item) + (condition|participant)`

Model 2: `rating ~ condition + political_leaning + (condition|item) + (condition|participant)`

Model 3: `rating ~ condition * political_leaning + (condition|item) + (condition|participant)`

Regression results:

- **COMPOUND** vs. **PHRASAL**: the compounds significantly decrease the logit of ratings in positive categories ($\beta = 0.526, SE = 0.152, p < 0.001$)
→ Authors' negative attitudes are perceived as more pronounced when enigmatic compounds are used
- **NEUTRAL** and **PHRASAL**: no significant difference ($\beta = -0.272, SE = 0.176, p = 0.123$)
- **Likelihood ratio tests**: Model 2 and 3 did not significantly improve the goodness of fit (Model 2 vs. 1: $\chi^2(2) = 0.384, p = 0.826$; Model 3 vs. 1: $\chi^2(6) = 2.004, p = 0.919$)
→ The increased perception of attitudinal meaning in enigmatic compounds is **systematic part of language use**, rather than being specific to subpopulations with different political leanings

BACKGROUND



CORPUS STUDY



PSYCHOLINGUISTIC EXPERIMENT



SIMULATIONS WITH LARGE LANGUAGE MODELS



CONCLUSIONS AND OUTLOOK

- ❖ Compound meaning in theoretical linguistics
- ❖ Automated detection of compounds in NLP

- ❖ Evidence of a systematic use of ad hoc compounds to express negative stance in newspapers from a dataset of >8 million tokens

- ❖ Validating the stance-triggering effect of ad hoc compounds with ratings by 212 German native speakers on items extracted from corpora

- ❖ Experiments with GPT-4 and GPT-3.5-turbo on their ability of identifying attitudinal meanings conveyed by ad hoc compounds

- ❖ Recommendations for directions of future work

Simulations With Large Language Models

Given the groundbreaking results of large language models (LLMs) in a wide range of NLP tasks, we are interested in whether LLMs can capture the attitudinal meaning conveyed by enigmatic compounds.

Models: GPT-4 and GPT-3.5-turbo (various temperature settings are employed)

Prompt:

You will read short text excerpts. Then, address the question, rating how certain people or events (marked by <start of statement to assess> and <end of statement to assess>) are depicted in the text, or your personal stance on a specific statement in the text. Provide your response using only numbers from 1 to 7, where 1 is very positive and 7 is very negative.

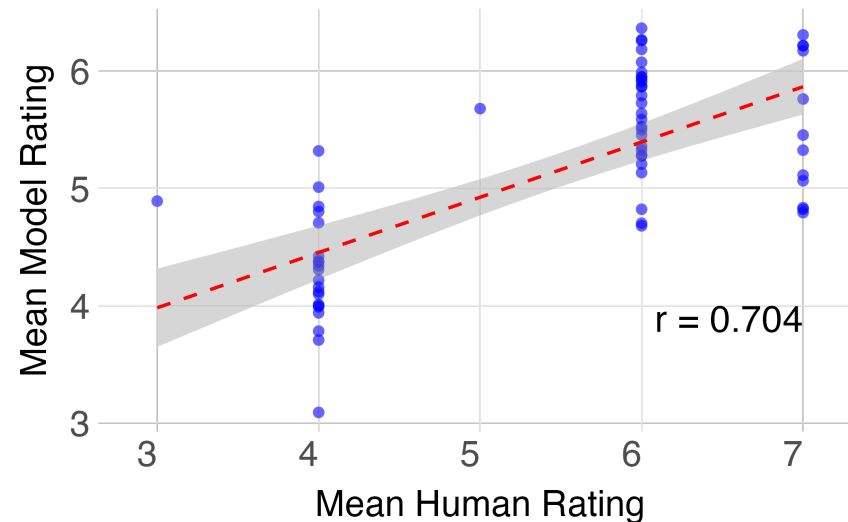
Corona numbers are exploding in Europe. And how is the situation in <start of statement to assess> the corona-hometown China <end of statement to assess>)? For a long time it seemed as if the country had the virus under control. People who tested positive had mainly been infected abroad. This trend is now changing.

How does the author speak about China?

Results and Discussion

Results from the best-performing model (GPT-4 with a temperature set to 0):

- **By-item alignment with human data:** a significant portion of the variance is captured ($R^2 = 0.48, p < 0.001$)



- **Within-item (condition-level) alignment with human data:** no significant effects ($R^2 = 0.43, p = 0.55$)
→ Indication: the LLM have difficulty recognizing enigmatic compounds as cues for attitudinal meanings

BACKGROUND



CORPUS STUDY



PSYCHOLINGUISTIC EXPERIMENT



SIMULATIONS WITH LARGE LANGUAGE MODELS



CONCLUSIONS AND OUTLOOK

- ❖ Compound meaning in theoretical linguistics
- ❖ Automated detection of compounds in NLP
- ❖ Evidence of a systematic use of ad hoc compounds to express negative stance in newspapers from a dataset of >8 million tokens
- ❖ Validating the stance-triggering effect of ad hoc compounds with ratings by 212 German native speakers on items extracted from corpora
- ❖ Experiments with GPT-4 and GPT-3.5-turbo on their ability of identifying attitudinal meanings conveyed by ad hoc compounds
- ❖ Recommendations for directions of future work

Conclusion

EMPIRICAL EVIDENCE

CORPUS STUDY



German newspapers systematically employ enigmatic compounds to express negative stances

PSYCHOLINGUISTIC EXPERIMENT



Enigmatic compounds systematically convey attitudinal meaning that are crucial for stance detection

vs.

REMAINING RESEARCH GAPS

DEPENDENCY PARSERS



State-of-the-art dependency parsers and UD treebanks are insufficient in identifying compounds

LARGE LANGUAGE MODELS



LLMs struggle with recognizing enigmatic compounds as triggers of attitudinal meanings



RECOMMENDATION:

A uniform approach for the annotation of compounds

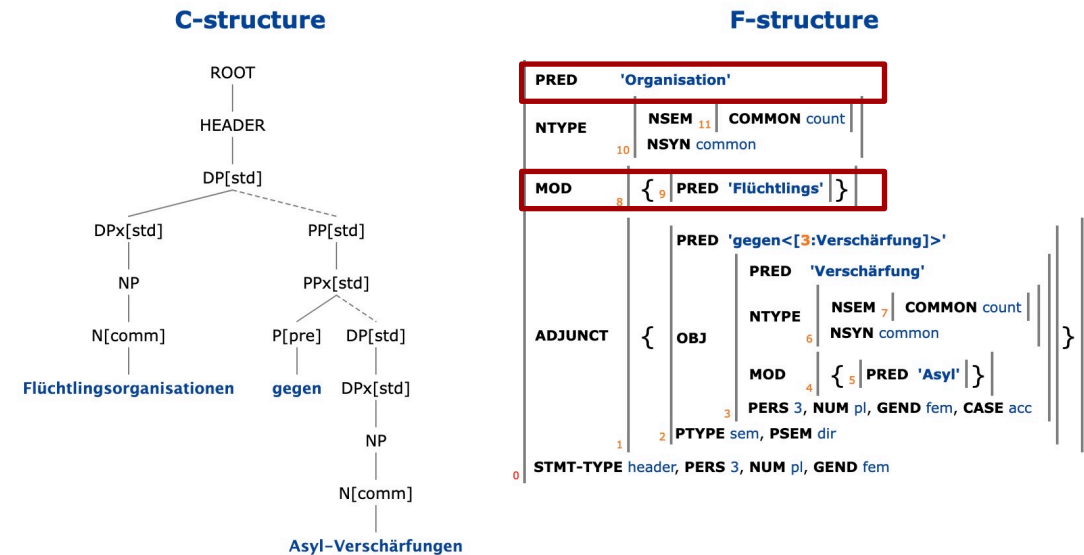
(ParGram Grammar as a potential solution)

(Source of icons: <https://www.flaticon.com>)

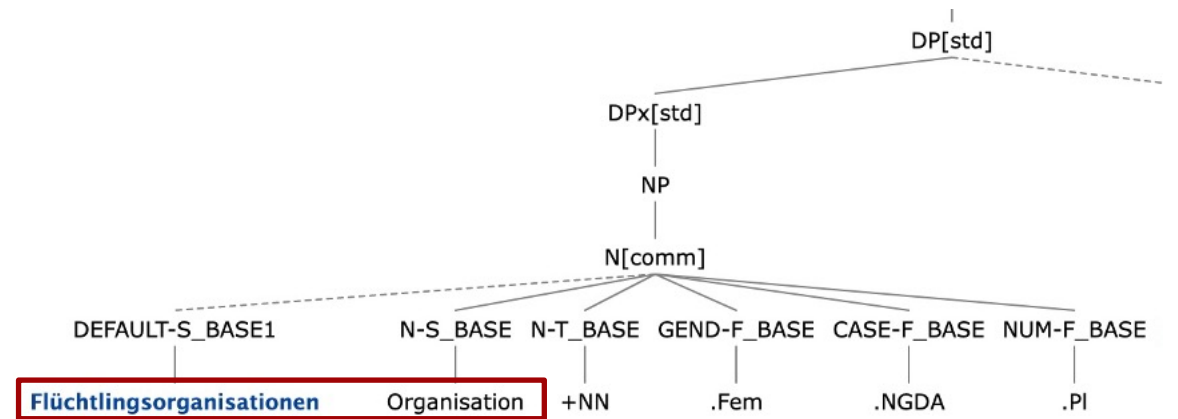
German ParGram Grammar (Butt et al., 1999; Sulger et al., 2013; Dipper, 2003)

1. Parsing based on Lexical Functional Grammar (LFG; Dalrymple, 2001):

- Context-free phrase structure (*c-structure*)
- Dependency (*f-structure*)



2. A built-in finite-state morphological analyzer separates the compound into the head noun and the modifier



Our recommendation: a morphology-level UD annotation that consistently separates the head noun from the modifier, with the modifier being identified clearly as such in the dependency analysis (As done in the LFG f-structure)

Direction of Future Work on Ad Hoc Compounds Detection

Potential directions with a morphology-level UD annotation:

- **Direction 1:** combining morphology parsing with consultation of existing dictionaries
- **Direction 2:**
 - Compiling an initial seed list of compounds for any given domain, with the heads and modifiers of the compounds identified by a morphological analyzer
 - Feeding the seed list into models calculating clusters of lexically similar words to identify further ad hoc compounds

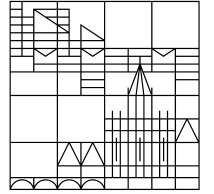
CODE & RESOURCES



OR VISIT

<https://github.com/qi-yu/enigmatic-compounds>

Universität
Konstanz



Thank you!
Questions & comments?