

Semi-Automatic Extension of GermaNet with Sense Definitions from Wiktionary

Verena Henrich, Erhard Hinrichs, and Tatiana Vodolazova
University of Tübingen
Department of Linguistics

LTC 2011



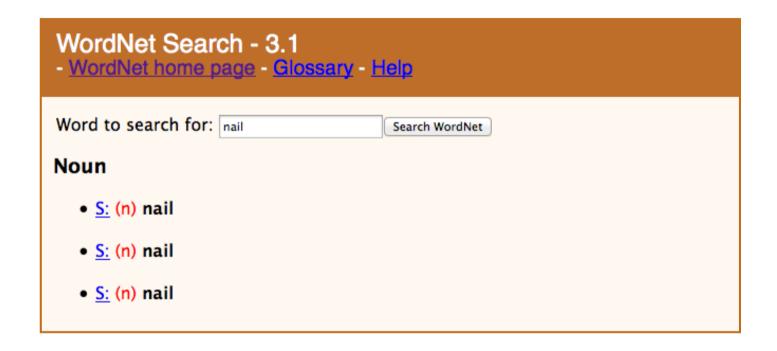
Introduction: The Necessity of Sense Descriptions



- Descriptions illustrate individual word senses in dictionaries
- For example: Princeton WordNet contains 3 senses for nail



Introduction: The Necessity of Sense Descriptions



- Descriptions illustrate individual word senses in dictionaries
- For example: Princeton WordNet contains 3 senses for nail
- Without definitions it is not easy to distinguish senses



Extend GermaNet with Sense Definitions

GermaNet

Anzeige

```
1. (n) [Anzeige] 'advertisement'
```

synonyms: [Annonce, Inserat]

hypernyms: [Ausschreibung]

hyponyms: [Versandanzeige] [Kaufgesuch]

[Verkaufsangebot] [Familienanzeige]

[Partnergesuch, Kontaktanzeige] [Stellenanzeige, Stellenangebot, Stellenannonce] [Stellengesuch] [Kleinanzeige] [Großanzeige] [Zeitungsanzeige]

is part of: [Zeitung, Blatt, Gazette]

2. (n) [Anzeige] 'complaint'

synonyms: [Strafanzeige]

hypernyms: [juristischer Text] [Klage, Anklage]

hyponyms: [Denunziation]

3. (n) [Anzeige] 'display'

synonyms: [Display]

hypernyms: [Device, Gerät]

hyponyms: [Handydisplay] [Bildschirmseite]



Extend GermaNet with Sense Definitions

Kurze Mitteilungen in den Medien, die der Bekanntmachung oder Werbung dienen

Short notices in the media for making announcements

Recht: Bekanntgabe einer Straftat bei einer Behörde

Law: report of a crime at an authority

Technik: eine Vorrichtung zur Signalisierung von Zuständen und Werten

Technical device for signaling visual information

GermaNet

Anzeige 'advertisement' 1. (n) [Anzeige] synonyms: [Annonce, Inserat] hypernyms: [Ausschreibung] hyponyms: [Versandanzeige] [Kaufgesuch] [Verkaufsangebot] [Familienanzeige] [Partnergesuch, Kontaktanzeige] [Stellenanzeige, Stellenangebot, Stellenannonce] [Stellengesuch] [Kleinanzeige] [Großanzeige] [Zeitungsanzeige] is part of: [Zeitung, Blatt, Gazette] 'complaint' 2. (n) [Anzeige] [Strafanzeige] synonyms: [juristischer Text] [Klage, Anklage] hypernyms: [Denunziation] hyponyms: 'display' 3. (n) [Anzeige] [Display] synonyms: [Device, Gerät] hypernyms: [Handydisplay] [Bildschirmseite] hyponyms:



GermaNet-Wiktionary Mapping

Wiktionary Anzeige Anzeige (Deutsch) [Bearbeiten] Substantiv, f [Bearbeiten] Bedeutungen: [1] kurze Mitteilungen in den Medien, die der Bekanntmachung oder Werbung dienen [2] Recht: Bekanntgabe einer Straftat bei einer Behörde [3] Technik: eine Vorrichtung zur Signalisierung von Zuständen und Werten [4] veraltend: ein sichtbarer Hinweis auf etwas Zukünftiges Synonyme: [1] Annonce; Inserat [4] Anzeichen Sinnverwandte Wörter: [3] Display Unterbegriffe: [1] Familienanzeige, Geburtsanzeige, Heiratsanzeige, Hochzeitsanzeige, Kontaktanzeige, Todesanzeige, Traueranzeige, Verlobungsanzeige Werbeanzeige [1] Kleinanzeige [2] Selbstanzeige, Strafanzeige

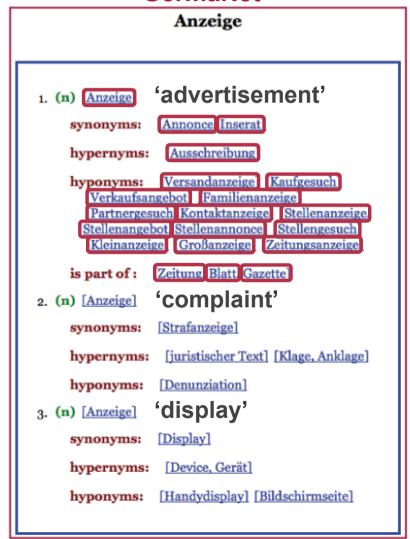
[3] Temperaturanzeige, Wasserstandsanzeige

GermaNet Anzeige 'advertisement' 1. (n) [Anzeige] synonyms: [Annonce, Inserat] hypernyms: [Ausschreibung] hyponyms: [Versandanzeige] [Kaufgesuch] [Verkaufsangebot] [Familienanzeige] [Partnergesuch, Kontaktanzeige] [Stellenanzeige, Stellenangebot, Stellenannonce] [Stellengesuch] [Kleinanzeige] [Großanzeige] [Zeitungsanzeige] is part of: [Zeitung, Blatt, Gazette] 'complaint' 2. (n) [Anzeige] synonyms: [Strafanzeige] [juristischer Text] [Klage, Anklage] hypernyms: [Denunziation] hyponyms: 3. (n) [Anzeige] 'display' [Display] synonyms: hypernyms: [Device, Gerät] [Handydisplay] [Bildschirmseite] hyponyms:



Bag of Words from GermaNet Sense Anzeige 'advert.'

GermaNet





Bag of Words from GermaNet Sense Anzeige 'advert.'

Anzeige Annonce Inserat Ausschreibung Versandanzeige Kaufgesuch Verkaufsangebot Familienanzeige Partnergesuch Kontaktanzeige Stellenanzeige Stellenangebot Stellenannonce Stellengesuch Kleinanzeige Großanzeige Zeitungsanzeige Zeitung Blatt Gazette

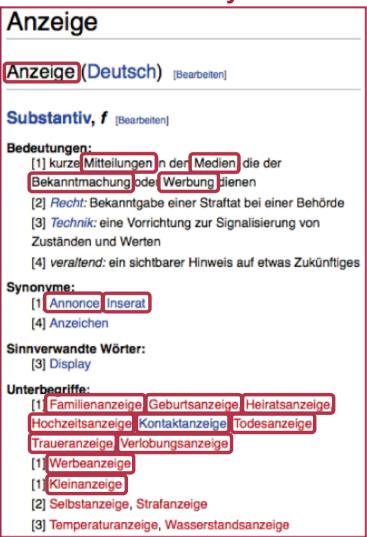
GermaNet Anzeige





Bag of Words from Wiktionary Sense Anzeige 'advert.'

Wiktionary





Bag of Words from Wiktionary Sense Anzeige 'advert.'

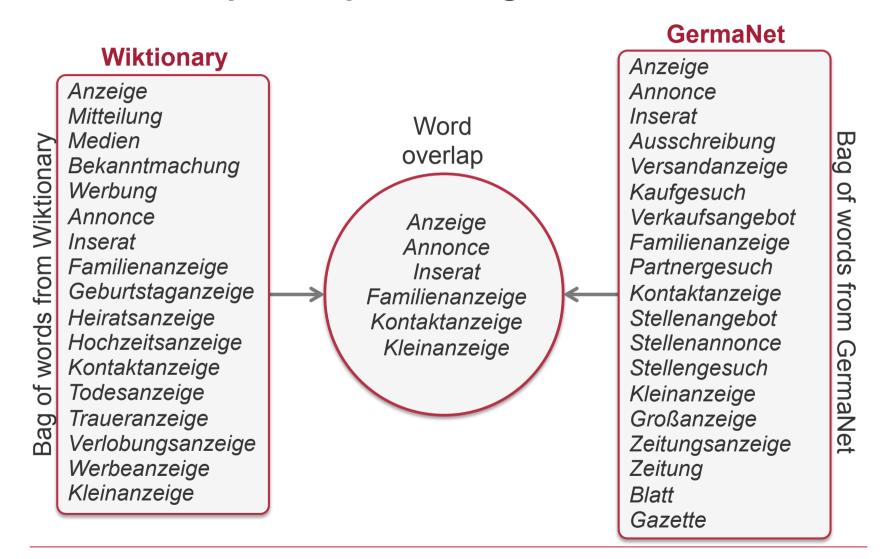
Wiktionary



Anzeige Mitteilung Medien Bekanntmachung Werbung Annonce Inserat Familienanzeige Geburtstaganzeige Heiratsanzeige Hochzeitsanzeige Kontaktanzeige Todesanzeige Traueranzeige Verlobungsanzeige Werbeanzeige Kleinanzeige



Word Overlap Example: Anzeige 'advertisement'





Coordinated Relations Example: Anzeige 'advert.'

Wiktionary

Anzeige Anzeige (Deutsch) [Bearbeiten] Substantiv, f [Bearbeiten] Bedeutungen: [1] kurze Mitteilungen in den Medien, die der Bekanntmachung oder Werbung dienen [2] Recht: Bekanntgabe einer Straftat bei einer Behörde [3] Technik: eine Vorrichtung zur Signalisierung von Zuständen und Werten [4] veraltend: ein sichtbarer Hinweis auf etwas Zukünftiges Synonyme [1] Annonce Inserat [4] Anzeichen Sinnverwandte Wörter: [3] Display Unterbegriffe 1 Familienanzeige Geburtsanzeige, Heiratsanzeige, Hochzeitsanzeige Kontaktanzeige Todesanzeige, Traueranzeige, Verlobungsanzeige Werbeanzeige [1] Kleinanzeige [2] Selbstanzeige, Strafanzeige [3] Temperaturanzeige, Wasserstandsanzeige

GermaNet

```
Anzeige
                 'advertisement'
1. (n) [Anzeige]
     synonyms:
                 (Annonce Inserat
     hypernyms:
                  [Ausschreibung]
     hyponyms: [Versandanzeige] [Kaufgesuch]
       [Verkaufsangebot] Familienanzeige
       Partnergesuck, Kontaktanzeige, Stellenanzeige,
       Stellenangebot, Stellenannonce [Stellengesuch]
       [Kleinanzeige] [Großanzeige] [Zeitungsanzeige]
     is part of: [Zeitung, Blatt, Gazette]
                 'complaint'
2. (n) [Anzeige]
     svnonvms:
                  [Strafanzeige]
                  [juristischer Text] [Klage, Anklage]
     hypernyms:
                  [Denunziation]
     hyponyms:
                 'display'
3. (n) [Anzeige]
                 [Display]
     synonyms:
     hypernyms: [Device, Gerät]
     hyponyms: [Handydisplay] [Bildschirmseite]
```



Coordinated Relations Example: Anzeige 'advert.'

Wiktionary

Anzeige

Anzeige (Deutsch) [Bearbeiten]

Substantiv, f [Bearbeiten]

Bedeutungen:

 kurze Mitteilungen in den Medien, die der Bekanntmachung oder Werbung dienen

[2] Recht: Bekanntgabe einer Straftat bei einer Behörde

[3] Technik: eine Vorrichtung zur Signalisierung von

Zuständen und Werten

[4] veraltend: ein sichtbarer Hisweis auf etwas Zukünftiges

Synonyme:

[1] Annonce Inserat

[4] Anzeichen

Sinnverwandte Wörter:

[3] Display

Unterbeariffe

[1 Familienanzeige Geburtsanzeige, Heiratsanzeige, Hochzeitsanzeige Kontaktanzeige Tedesanzeige,

Hochzeitsanzeige Kontaktanzeige Hodesanzeig

Traueranzeige, Verlobungsanzeige

- [1] Werbeanzeige
- [1] Kleinanzeige
- [2] Selbstanzeige, Strafanzeige
- [3] Temperaturanzeige, Wasserstandsanzeige

Synonyms in common

Hyponyms in common

GermaNet

Anzeige

1. (n) [Anzeige] 'advertisement'

synonyms: Annonce Inserat

hypernyms: [Ausschreibung]

hyponyms: [Versandanzeige] [Kaufgesuch]

[Verkaufsangebot] Familienanzeige

[Partnergesuck, Kontaktanzeige | Stellenanzeige, Stellenangebot, Stellenannonce] [Stellengesuch]

[Kleinan eige] [Großanzeige] [Zeitungsanzeige]

is part of: [Zeitung, Blatt, Gazette]

2. (n) [Anzeige] 'complaint'

synonyms: [Strafanzeige]

hypernyms: [juristischer Text] [Klage, Anklage]

hyponyms: [Denunziation]

3. (n) [Anzeige] 'display'

synonyms: [Display]

hypernyms: [Device, Gerät]

hyponyms: [Handydisplay] [Bildschirmseite]



Different Sense Granularities

Wiktionary

Sammlung historisch oder aus anderen Gründen bedeutsamer Dokumente

'Collection of documents that are historically or for other reasons important'

Einrichtung, Institution zur Aufbewahrung und Pflege historisch oder aus anderen Gründen bedeutsamer Dokumente

'Institution for storing and maintenance of historically or for other reasons important documents'

Gebäude oder Gebäudeteil, der eine Institution zur Aufbewahrung von Dokumenten enthält

'Building or part of the building containing an institution for storing documents'

GermaNet

Archiv 'data repository'

Archiv 'archive'

Archiv 'archived file'



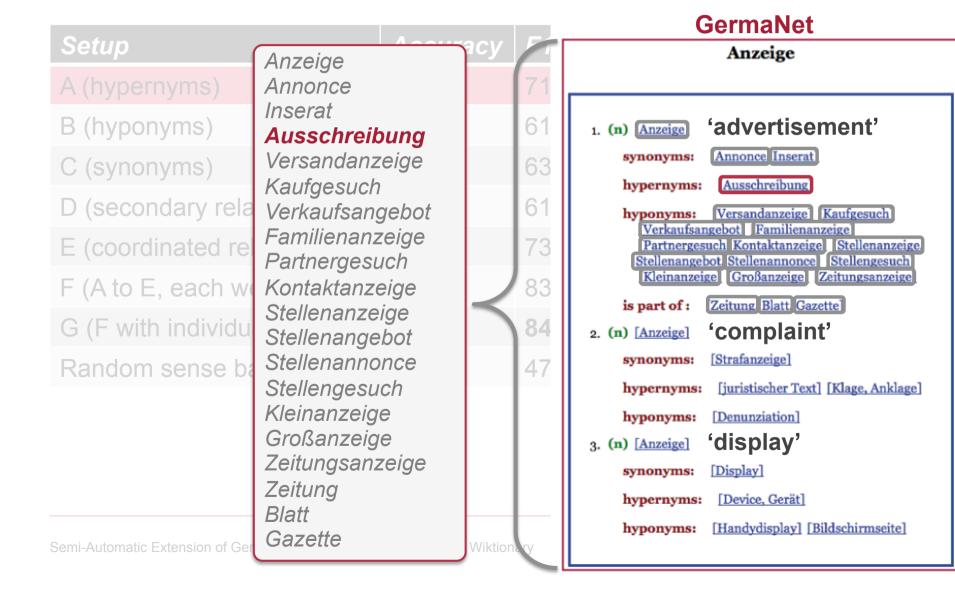
Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2

- Evaluation is based on the alignment of 20997 distinct words
- Accuracy: ratio of correctly classified mappings compared to all possible mappings



Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2

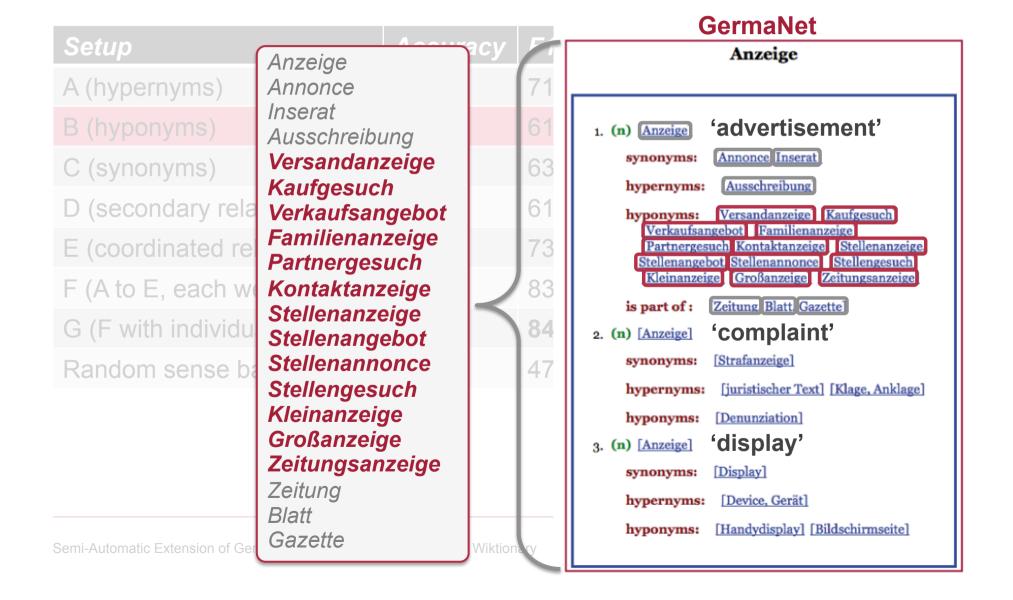






Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2

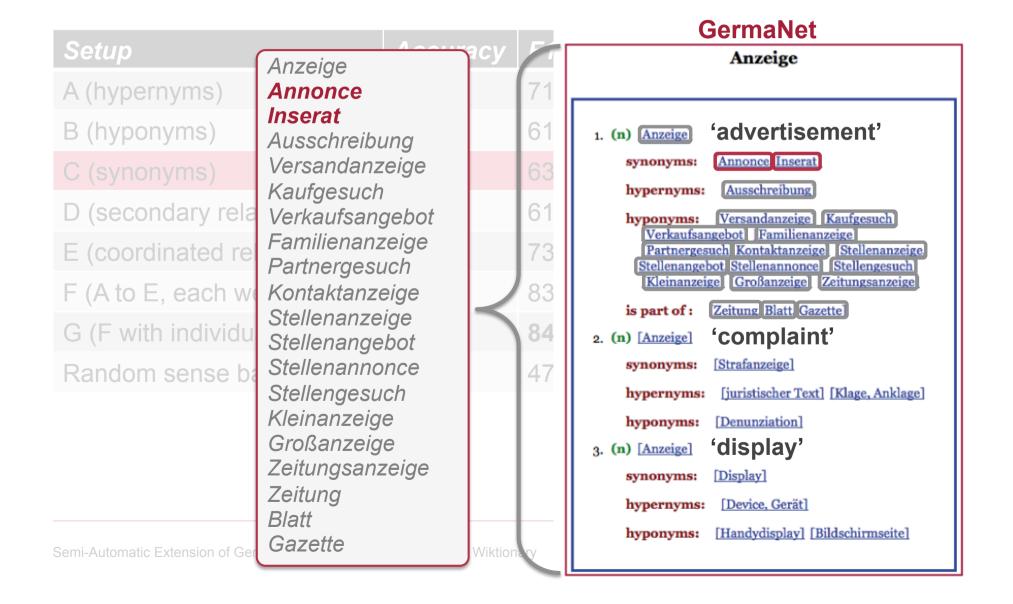






Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2







Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2

• Secondary relations: association, causation, entailment, holonymy, meronymy, and pertainymy



Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2



Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2



Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2

• Individual weights: hypernyms 2; hyponyms 0.5; synonyms 3; secondary relations 0.5; coordinated relations 3



Setup	Accuracy	F1
A (hypernyms)	93.3%	71.7
B (hyponyms)	93.1%	61.0
C (synonyms)	93.8%	63.6
D (secondary relations)	93.2%	61.0
E (coordinated relations)	93.2%	73.5
F (A to E, each weight 1)	92.3%	83.3
G (F with individual weights)	91.9%	84.3
Random sense baseline	53.7%	47.2



Conclusion

- Sense definitions are a crucial component for wordnets
- We have presented a method for semi-automatically enriching GermaNet with sense definitions from Wiktionary
- Evaluation results underscore the feasibility of the approach
- Until now, 22296 synsets (32%) in GermaNet have sense definitions from Wiktionary
- Extension of GermaNet will be made freely available
- Future work: use the sense-mapping between GermaNet and Wiktionary to increase GermaNet's coverage



Thank you.

Verena Henrich, Erhard Hinrichs, and Tatiata Vodolazova

Department of Linguistics University of Tübingen

Wilhelmstr. 19 72074 Tübingen Germany

verena.henrich@uni-tuebingen.de http://www.verenahenrich.de