# The Syntax-Semantics Interface of German Particle Verbs

## Sabine Schulte im Walde

Computational Linguistics

Saarland University

Saarbrücken, Germany

# German Particle Verbs (PVs)

- Morphological composition:
  particle verb (PV) = particle + base verb (BV)

- Focus on preposition particles: *ab, aus, bei, für, zu, ...*

- Compositionality: transparent vs. opaque PV senses
  - » *ab-holen* `fetch´  ↔  *holen* `fetch´
  - » *an-fangen* `begin´  ↔  *fangen* `catch´
  - » *ein-setzen* `insert, begin´  ↔  *setzen* `put/sit (down)´

- Syntax-Semantics Interface:
  - » PVs may change behaviour of base verbs
  - » changes quite regular (Stiebels, 1996; Aldinger, 2004)

# Syntactic Change (Addition)

Sie *lächelt*.

`She smiles.´

\* Sie *lächelt* [NP$_{acc}$ ihre Mutter].

`She smiles her mother.´

Sie *lächelt* [NP$_{acc}$ ihre Mutter] *an*.

`She smiles at her mother.´

# Syntactic Change (Saturation)

Er *stellt* [NP$_{acc}$ das Glas] [PP$_{dir}$ auf den Tisch].
`He puts the glass on the table.´

\* Er *stellt* [NP$_{acc}$ das Glas].
`He puts the glass.´

Er *stellt* [NP$_{acc}$ das Glas] *ab.*
`He puts the glass down.´

# Aktionsart: Ingressive PVs

Er *schrie* [durative drei Stunden lang].

`He cried for three hours.´

* Er *schrie* [durative drei Stunden lang] *auf*.

`He cried out for three hours.´

Er *schrie* [punctual plötzlich] *auf*.

`He cried out suddenly.´

# Questions about German Particle Verbs

- **Word senses**: Which are the PV senses? Which are the factors that determine the senses (semantic class of BV, behaviour of BV, transfer pattern, particle type, etc.)?

- **Compositionality**: Is a PV (sense) transparent or opaque (compositional or non-compositional)?

- **Transfer patterns BV → PV**: How many and which patterns exist, and are they regular (with respect to what)?

- **Meaning of particle**: What's the meaning contribution of the particle, and to what extent is it regular?

- **Semantic class**: Which are the semantic classes of a PV?

# Semantic Classes of German PVs

Lexical meaning of a verb determines its behaviour;
verb behaviour as indicator of semantic classes

**einsetzen** $\longrightarrow$ beginnen, anfangen
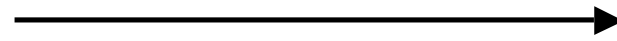
*opaque,
syn-sem applies*

**einsetzen** $\longrightarrow$ setzen

*transparent,
syn-sem does not apply a priori,
syn-sem applies after pattern transfer (?)*

# German Particle Verb Exploration:

# Identification and Quantitative Description

# Identification of German PVs

- Lexicalised probabilistic context-free grammar (Charniak, 1997; Carroll and Rooth, 1998)

- Unsupervised training by *EM-Algorithm* (Baum, 1972), robust statistical parser *LoPar* (Schmid, 2000)

- 35 million words of German newspaper corpora

- 17,851 verbs in grammar with *freq(verb)>=1;* 12,404 potential PVs with 2,215 BVs

- Subcategorisation behaviour of verbs

# Frame Comparison of PVs and BVs

| *ankommen* (freq: 1,831) | | | |
|---|---|---|---|
| n | 38.82 | np:Dat.bei | 1.50 |
| x | 16.12 | nap:Dat.in | 1.40 |
| na | 10.56 | np:Akk.auf | 1.02 |
| ns-w | 5.76 | | |
| ns-2 | 4.63 | | |
| *einsetzen* (freq: 3,390) | | | |
| na | 40.29 | nap:Dat.in | 4.33 |
| nap | 16.44 | nap:Dat.bei | 2.81 |
| nr | 15.87 | npr:Akk.für | 2.63 |
| n | 10.86 | nap:Dat.zu | 1.76 |
| nad | 4.71 | nap:Akk.für | 1.35 |
| *umbringen* (freq: 683) | | | |
| na | 53.60 | nap:Dat.in | 5.43 |
| nr | 19.36 | nap:Dat.nach | 1.54 |
| nap | 12.23 | npr:Dat.in | 1.23 |
| nad | 3.20 | | |
| nas-2 | 1.97 | | |

# Selectional Noun Preferences

- Grammar model provides selectional preference information on a fine-grained level

- Argument heads with respect to a specific verb-frame-slot combination

- Example:   *einsetzen* `insert, begin, apply´

  **n**     →     *Run* `run´, *Regen* `rain´, *Prozess* `process´,
                  *Welle* `wave´, *Kampf* `fight´, *Kritik* `criticism´,  ...

  **n**a    →     *Polizei* `police´, *Regierung* `government´,
                  *Wehr* `army´, *Bahn* `railway services´, ...

  n**a**    →     *Gas* `gas´, *Mittel* `means´,
                  *Kommission* `commission´, *Waffe* `weapon´, ...

# Selectional Noun Preferences

| ankommen - _n_ | | | kommen - _n_ |
|---|---|---|---|
| _Botschaft_ | 'message' | 17.17 | 10.22 |
| _Zug_ | 'train' | 11.14 | 14.78 |
| _Flüchtling_ | 'refugee' | 7.13 | 28.07 |
| _Film_ | 'movie' | 5.18 | 13.32 |
| _Spende_ | 'donation' | 5.00 | 6.61 |
| _Brief_ | 'letter' | 4.72 | 22.18 |
| _Buch_ | 'book' | 4.60 | 13.78 |

| umbringen - _na_ | | | bringen - _na_ | | |
|---|---|---|---|---|---|
| _Mensch_ | 'human' | 13.63 | Ergebnis | 'result' | 111.86 |
| _Frau_ | 'woman' | 10.24 | Erfolg | 'success' | 73.39 |
| _Kind_ | 'child' | 9.54 | Geld | 'money' | 54.29 |
| _Mann_ | 'man' | 7.46 | Problem | 'problem' | 52.44 |
| _Vater_ | 'father' | 5.97 | Vorteil | 'advantage' | 47.58 |
| _Million_ | 'million' | 4.91 | Opfer | 'victim' | 44.34 |
| _Leute_ | 'people' | 4.87 | Entscheidung | 'decision' | 39.92 |

# German Particle Verb Exploration:

# Similarity at the Syntax-Semantics Interface

# Features at the Syntax-Semantics Interface

- Distributions for verb description:

  » *syntax:* syntactic frame types

  » *syntax-pp:* syntactic frame types + PPs

  » *pref:frame-noun*: selectional preferences;
      nouns <u>with</u> reference to frame type and slot

  » *pref:noun*: selectional preferences;
      nouns <u>without</u> reference to frame type and slot

- Cut-offs for noun corpus frequencies

# Semantic Nearest Neighbours (NNs)

1. Instantiate verbs and NN candidates by probability distributions over feature descriptions

2. Determine the nearest neighbour by distance measure
   *skew divergence*

$$D(x \parallel y) = \sum_{i=1}^{n} x_i * \log \frac{x_i}{y_i}$$

$$Skew(x, y) = D(x \parallel w * y + (1 - w) * x)$$

3. Evaluate nearest neighbours against gold standards (existence and type) $\rightarrow$ precision

# Gold Standards

| | pv | cand | avg rel | baseline |
|---|---:|---:|---:|---:|
| GermaNet | 1,856 | 2,338 | 10 | 0.43% |
| GermaNet-red | 95 | 613 | 12 | 1.93% |
| DIC (syn, ant) | 63 | 1,645 | 47 | 2.84% |
| Association-1 | 36 | 623 | 25 | 4.01% |
| Association-2 | 76 | 1,040 | 19 | 1.84% |

# Semantic Nearest Neighbour Precision

| | syn | syn-pp | pref:frame-noun | | | | pref:noun | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 10 | 100 | 500 | 1000 | 10 | 100 | 500 | 1000 |
| GN | 2 | 3 | 8 | 9 | 8 | 7 | **10** | **10** | 9 | 8 |
| GN-red | 6 | 9 | 18 | 18 | 16 | 13 | **20** | **20** | **20** | 16 |
| DIC | 6 | 13 | 33 | 35 | **37** | 35 | 32 | 32 | 33 | 32 |
| Assoc1 | 17 | 22 | 50 | 50 | 50 | 47 | 50 | 53 | **56** | 50 |
| Assoc2 | 9 | 12 | **21** | **21** | **21** | 20 | 18 | 16 | 16 | 20 |

# Insights on Feature Distributions

- Low results for syntax

- Only slight improvement by syntax-pp
  → specific for particle verbs

- Most successful distributions are nominal preferences; significant differences *syntax* vs. *pref* for DIC and Assoc1

- Minimal differences: pref:frame-noun and pref:noun!

→ Relevant information in the distributions are nouns, references to argument structure (functions) are minor

→ Hypothesis does not apply to PVs as to verbs in general

# Outlook

1. Learn PV-BV patterns.

2. Apply pattern change to PVs.

3. Determine similarity of PVs with other verbs.

4. Distinguish semantic classes of BVs, PVs, particles.

5. Determine compositionality of PVs.