Can Human Verb Associations help identify Salient Features for Semantic Verb Classification?

Sabine Schulte im Walde

Computational Linguistics

Saarland University, Saarbrücken, Germany

IGK Meeting 2006, Edinburgh/Firbush

September 6-15

Motivation

1. Verb associations model aspects of verb meaning.
2. Semantic verb classes are based on verb meaning.
3. For an automatic acquisition of semantic verb classes we need to model verb meaning by verb features.

Questions:

Can human verb associations help identify salient features for semantic verb classification?

Are the same types of features salient for different types of semantic verb classes?

Variety of Semantic Verb Classes

- Semantic verb classes with various instantiations of semantic similarity, e.g.
  - » syntactic similarity / alternations (Levin, 1993):
    buy, catch, earn, find, steal, ... (obtaining:get verbs)
  - » synonymy (WordNet):
    buy, purchase
  - » situation-based agreement (FrameNet):
    buy, purchase (commerce_buy) inherits from
    acquire, gain, get, obtain, procure, secure (getting)

Creation of Semantic Verb Classes (1)

- Resource-intensive vs. automatic methods
- Examples of automatic methods:
  - » Merlo & Stevenson (CL Journal, 2001):

    classify 60 English verbs which alternate between an intransitive and a transitive usage into three classes;

    features model syntactic frame alternation proportions and heuristics for semantic role assignment
  - » Stevenson & Joanis (CoNNL, 2003):

    classify into 13 Levin classes with 20 verbs each in a two-/three-class distinction;

    general feature space and manual/seed subsets for syntactic slots, tense, voice, aspect, animacy

Creation of Semantic Verb Classes (2)

- Examples of automatic methods (cont'd):
  - » Korhonen, Krymolowski & Marx (ACL, 2003):
    48+26+57 polysemous Levin/Dorr/Korhonen classes; subcategorisation info from Briscoe/Carroll parser
  - » Schulte im Walde (PhD, 2003; CL Journal, 2006):
    classifies 168 verbs into 43 FrameNet-style classes; subcategorisation frames, PPs, selectional preferences

- Classification and clustering parameters:
  verbs, classes, algorithm, features, etc.

Semantic Verb Classes: Features

- Model semantic similarity of interest
- Similarity at the syntax-semantics interface: *behaviour*
- Potentially salient features:
  - » syntactic frames
  - » prepositional phrases
  - » argument role fillers
  - » adverbial adjuncts, etc.
- Granularity of features

Associations: Guide to Feature Selection

- Goal: human associations to identify salient features
- Assumptions:
  - » associations model verb meaning aspects
  - » theory-independent
  - » variety of semantic verb relations
  - » guidance to feature selection
- Comparison with empirical models of verb features: window co-occurrence vs. syntactic frame fillers

Procedure

1. Collect human verb associations
2. Hierarchical clustering on experiment verbs; basis: verb associations (assoc-classes)
3. Validate clustering against GermaNet and FrameNet
4. Compare cluster features with corpus-based features
5. Hierarchical clustering on experiment verbs; basis: corpus-based features (corpus-classes)
6. Compare corpus-classes against assoc-classes
7. Evaluation of hypothesis

Human Verb Associations: Collection and Analysis

Joint work with Alissa Melinger and Katrin Erk.

Web Experiment: Material

- 330 German verbs
- Variety of semantic verb classes, possible ambiguity:
  - self-motion: *gehen* 'walk', *schwimmen* 'swim'
  - cause: *verbrennen* 'burn', *reduzieren* 'reduce'
  - experiencing: *lachen* 'laugh', *überraschen* 'surprise'
  - communication: *erzählen* 'tell', *klagen* 'complain'
  - body: *schlafen* 'sleep', *abnehmen* 'lose weight'
- Variety of frequency ranges (1 < freq < 71,604)
- Random distribution: 6 data sets à 55 verbs,
  balanced for class affiliation and frequency ranges

Web Experiment: Procedure

schneien

kalt

rodeln

Schneemann

weiß

dämmern

Web Experiment: Data

- 299 accepted data files
- Participants per data set: between 44 and 54
- Number of trials: 16,445
- Number of associations per target verb:
  range 0-16, average: 5.16
- Responses: 79,480 tokens for 39,254 types

Quantification over Association Types

klagen ‘complain, moan, sue’		
Gericht	‘court’	19
jammern	‘moan’	18
weinen	‘cry’	13
Anwalt	‘lawyer’	11
Richter	‘judge’	9
Klage	‘complaint, lawsuit’	7
Leid	‘suffering’	6
Trauer	‘mourning’	6
Klagemauer	‘Wailing Wall’	5
laut	‘noisy’	5

Linguistic Analyses of Experiment Data

- Preference for morpho-syntactic category of responses?
  → distinguish major parts-of-speech:
  nouns, verbs, adjectivs, adverbs

- Typical argument holders of verb valency?
  → investigate linguistic functions realised by nouns:
  empirical grammar model

- Common appearance in corpus data?
  → determine co-occurrence of target and reponse:
  German newspaper corpus, 200 million words

Morpho-Syntactic Distribution

		V	N	ADJ	ADV
TOKEN	Freq	19.863	48.905	8.510	1.268
	Prob	25	62	11	2
TYPES	Freq	9.317	23.524	4.983	802
	Prob	24	61	13	2

Morpho-Syntactic Correlations/Tests

- Correlation **+**

  target verb frequency ⇔ verb/adverb responses

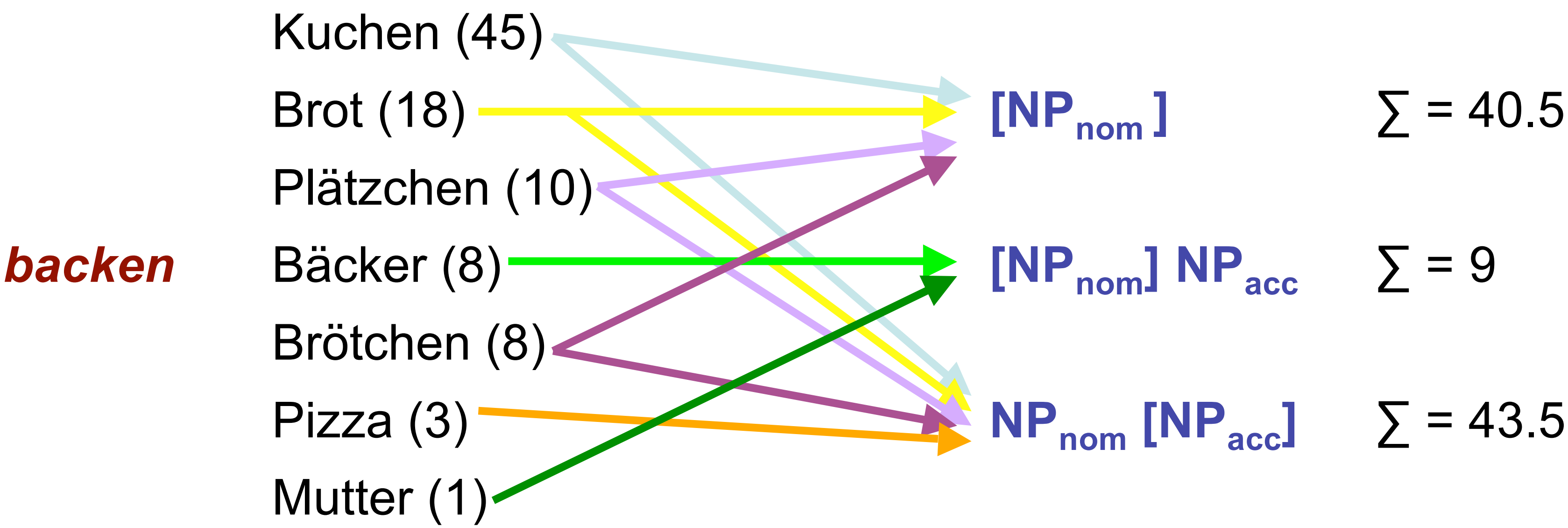
- Correlation **-**

  target verb frequency ⇔ noun responses

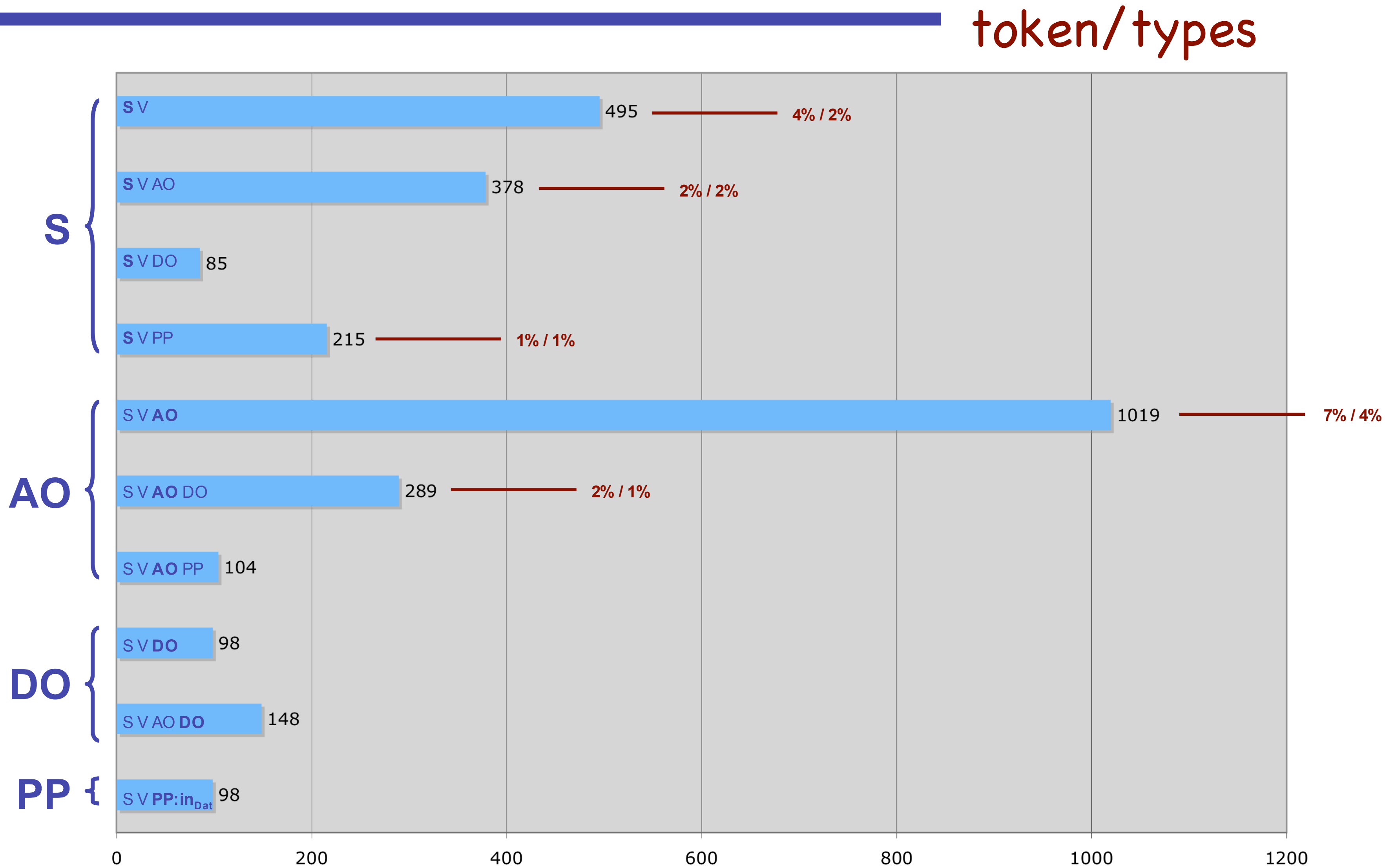
- Variation across verb classes

Syntax-Semantic Functions: Analysis

- Look up syntactic relationships between verb and nouns
- Typical conceptual roles which speakers have in mind
- Example:



Syntax-Semantic Frame Inspection



Window Co-Occurrence across POS

- Corpus data: 200 million word newspaper text
- Window (left+right): 5/20 words, excluding symbols
- Basis: association tokens
- Distinction with respect to window frequency

window	*1*	*2*	*3*	*5*	*10*	*20*	*50*
5	*66*	*56*	*50*	*42*	*33*	*23*	*14*
20	*77*	*70*	*66*	*59*	*50*	*40*	*27*

Window Co-Occurrence Verb-Noun

- Corpus data: 200 million word newspaper text
- Window (left+right): 5/20 words, excluding symbols
- Basis: association tokens
- Distinction with respect to window frequency

window	*1*	*2*	*3*	*5*	*10*	*20*	*50*
5	*68*	*59*	*53*	*45*	*36*	*25*	*15*
20	*78*	*72*	*68*	*61*	*52*	*42*	*28*

Missing Verb-Noun Co-Occurrence

» lemmatisation and tagging:

e.g. composita: *Übermacht, Zeitspanne, Autorennen*

» domain:

radeln ‘bike’ - *Oma* ‘grand-mom’ (1)

stoppen ‘stop’ - *Plosiv* ‘plosive’ (1)

» scene information/world knowledge:

trocknen ‘dry’ - *Trockner* ‘dryer’ (11)

auftauen ‘defrost’ - *Wasser* ‘water’ (14)

Association Analysis: Summary

- Morpho-syntactic distribution: nouns dominate
- Nouns represent argument roles of verbs
- Scene information in addition to subcategorisation
- Co-occurrence counts to supplement argument counts
- Results depend on verb frequency and semantic class
- Usage of roles and window-based nouns for distributional verb descriptions

Association-based Verb Classes:

Creation and Validation

Association Overlap

klagen / jammern 'moan'

overlap:
35 types

Association-based Clustering

- Agglomerative (bottom-up) hierarchical clustering
- Similarity measure: *skew divergence*
- Merging criterion: *Ward's method* (sum-of-squares)
- Hierarchy cut: 100 classes
- Cluster analysis informs about
  - » classes
  - » verbs
  - » class features, i.e. associations

Association-based Example Classes

Class	Features
bedauern \`regret´, heulen \`cry´, jammern \`moan´, klagen \`complain, moan, sue´, verzweifeln \`become desperate´, weinen \`cry´	Trauer \`mourning´, weinen \`cry´, traurig \`sad´, Tränen \`tears´, jammern \`moan´, Angst \`fear´, Mitleid \`pity´, Schmerz \`pain´, etc.
abnehmen \`lose weight´, abspecken \`lose weight´, zunehmen \`gain weight´	Diät \`diet´, Gewicht \`weight´, dick \`fat´, abnehmen \`lose weight´, Waage \`scale´, Essen \`food´, essen \`eat´, Sport \`sports´, dünn \`thin´, Fett \`fat´, etc.

Validation

- Claim: A clustering based on verb associations and a standard setup compares well with existing semantic classes.
- Lexical semantic resources:
  » GermaNet (Kunze, 2000)
  » Salsa / FrameNet (Erk *et al.*, 2003)
- Extraction of sub-classifications of resources:
  » GermaNet 33 classes with 56 verbs (71 senses)
  » FrameNet 49 classes with 104 verbs (220 senses)
- Hierarchical clustering of verb subsets;
  pair-wise evaluation (Hatzivassiloglou/McKeown, 1993):
  $\langle v1, v2 \rangle \in$ cluster $\rightarrow$ $\langle v1, v2 \rangle \in$ gold standard ?
  » GermaNet 62.69% (UB: 82.35%),
  » FrameNet 34.68% (UB: 49.90%)

Association-based Classes: Summary

- Considerable overlap between association-based classes and the lexical resources GermaNet and FrameNet
- Differences in validation for GermaNet vs. FrameNet:
  - » types of semantic similarity
  - » degrees of ambiguity
  - » clustering parameters: number of verbs, etc.

→ Potential use of association-based classes as gold standard for clustering experiments

→ Associations provide guidance to feature selection

Exploring

Semantic Class Features

Exploring Semantic Class Features

- Grammar-based relations from statistical grammar:

  verb-noun pairs with nominal heads of NPs and PPs,
  verb-adverb pairs from adverbial modifiers

- Co-occurrence window:

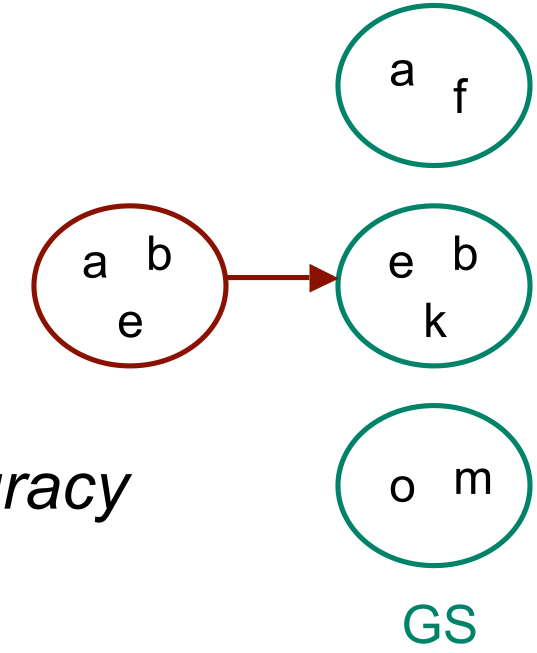
  200-million word newspaper corpus,
  20-word window (left and right)

Exploring Semantic Class Features

features	grammar relations						
	n	na	na	NP	PP	NP&PP	ADV
	12,635	14,458	13,416	20,792	14,513	22,366	10,080
cov. (%)	3.82	4.32	6.93	12.23	5.36	14.08	3.63

features	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
	934,783	100,305	96,178	5,688	660,403	34,095
cov. (%)	66.15	57.79	9.13	1.72	39.27	15.51

Corpus-based Clustering



- Experiment verbs:
  agglomerative hierarchical clustering,
  evaluation against assoc-classes: *accuracy*
- GermaNet:
  random selection of 100 synsets,
  random hard version with 233 verbs,
  clustering and evaluation as above
- FrameNet:
  pre-release version from May 2005,
  random hard version with 406 verbs in 77 classes,
  clustering and evaluation as above

Corpus-based Clustering: Questions

- Do the results of the clusterings with respect to the underlying feature types correspond to the overlap of associations and feature types?
- Do the corpus-based feature types which were identified on the basis of the associations outperform previous clustering results?
- Do the results generalise over the semantic class type?

Corpus-based Clustering: Results

	frames		grammar relations						
	f-pp	f-pref	n	na	na	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

	frames		**grammar relations**						
	f-pp	f-pref	<u>n</u>	<u>n</u>a	n<u>a</u>	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

	frames		grammar relations						
	f-pp	f-pref	n	na	na	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

no correlation!

	frames		grammar relations						
	f-pp	f-pref	n	na	na	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

no significant difference!

	frames		grammar relations						
	f-pp	f-pref	n	na	na	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

	frames		grammar relations						
	f-pp	f-pref	n	na	na	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

	frames		grammar relations						
	f-pp	f-pref	<u>n</u>	<u>n</u>a	n<u>a</u>	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

	frames		grammar relations						
	f-pp	f-pref	n	na	na	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Corpus-based Clustering: Results

	frames		grammar relations						
	f-pp	f-pref	n	na	na	NP	PP	NP&PP	ADV
Assoc	37.50	37.80	35.90	37.18	39.25	39.14	37.97	**41.28**	38.53
GN	46.98	49.14	**58.01**	53.37	51.90	53.10	54.21	51.77	51.82
FN	33.50	32.76	29.46	30.13	32.74	34.16	28.72	33.91	**35.24**

	co-occurrence: window-20					
	all	cut	ADJ	ADV	N	V
Assoc	39.33	**39.45**	37.31	36.89	39.33	38.84
GN	51.53	52.42	50.88	47.79	**52.86**	49.12
FN	32.01	32.84	31.08	31.00	**34.24**	31.75

Properties of Gold Standard Verb Classes

	verbs	average verb freq	no. of verbs with freq < 50/20/10		
Assoc	330	2,465	41	16	8
GN	233	1,040	98	65	40
FN	406	1,876	54	16	11

Summary of Results

- No correlation between overlap of associations / feature types and respective clustering results (Pearson's correlation, p>.1)
- Window-based features are not significantly worse than selected grammar-based functions;
  applying cut-offs has almost no impact
- Several cases of grammar-based and window-based features outperform frame-based features (i.e., previous work)
- Adverbs outperform frame-based features, even some nominals
- Most successful feature types vary for gold standards
- Significantly better results for GermaNet clusterings than for experiment-based and FrameNet clusterings

Outlook

- Which feature types are appropriate to model human associations?
- Which types of (semantic) verb classifications rely on which types of features?
- Which classification parameters (e.g., size of classes, ambiguity of verbs, empirical properties of verbs) influence the clustering result?
- How do the features and parameters differ with respect to a specific semantic verb class?