



## Introduction

- Semantic Role Labeling (SRL) determines ‘who’ did ‘what’ to ‘whom’ in sentences by identifying and associating predicates with their semantic arguments
- **Complex predicates (CPs):**
  - Multi-headed, composed of more than one grammatical element
  - Frequent phenomenon: 1 CP in every 5th sentence in the Wiki50 corpus
  - Most frequent types:
    - VPCs:** verb particle constructions, e.g. *watch out*, 50% of CPs in Wiki50
    - LVCs:** light verb constructions, e.g. *take a walk*, 40% of CPs in Wiki50
- **The PropBank corpus (PB):** [Palmer et al., 2005]
  - One of the most widely used resources of training data for SRL systems
  - Poor coverage and treatment of CPs (e.g. *take a hard line*, *take time* and other *take* CPs annotated with the most frequent sense of *take* denoting *acquire*, *come to have*, *chose*, *bring with you from somewhere*)

## Representing CPs for SRL

- The current PB representation treats CPs as lexical usages of the verb
 

WHO?      OF WHAT?

WHO?    WHAT?    OF WHAT?

Frank takes care of business

*take.01      care.01*

Current (top) and improved (bottom) PB annotation of take care [Duran et al., 2011]

WHO?      OF WHAT?

Frank takes care of business

*(take+care).01*
- (Manually) creating new rolesets for CPs is cumbersome
- **Aliasing:** map CPs to existing PB entries with the same meaning and argument structure (*take care* → *care.01* ‘to be concerned’) [Bonial et al., 2014]
- **Our approach:**  
**Speed up the mapping process with an automatic system**

## Human Annotation

- 4 annotators independently annotated 100 CPs (50 VPCs and 50 LVCs) from the Wiki50 corpus in their original contexts
  - Set was balanced for frequency
  - All occurrences (197) of the CPs are included → account for polysemy
  - Disagreements were discussed to obtain a consensus
- | IAA     | A & B | A & C | A & D | Micro Average |
|---------|-------|-------|-------|---------------|
| Strict  | 67%   | 51%   | 44%   | 53%           |
| Lenient | 85%   | 74%   | 67%   | 75%           |
- | Decision       | #  | CP example        |
|----------------|----|-------------------|
| aliased        | 96 | take part         |
| mw PB pred.    | 60 | open up           |
| compositional  | 18 | obtain permission |
| no alias found | 16 | go into politics  |
| discarded      | 7  | take control      |
- **‘aliased’** PB roleset as true alias
  - **‘mw. PB pred.’** PB roleset for this CP already exists

## Method

- Automatically finding aliases for CPs is related to finding (near-) synonymous predicates and their accompanying roleset for the CPs
- Apply the distributional hypothesis: assess the similarity of expressions by looking at their contexts
- Multilingual variant: multilingual contexts work better for synonym acquisition than monolingual syntactic contexts [Van der Plas and Tiedemann, 2006]

### Overview of the alias finder

**Parallel corpus**  
 + lemmas  
 + POS  
 + synt. dependencies  
 + SRL  
 + word alignments

Extract CPs and PB roleset groups

Populate matrix with translation counts (alignments of CPs and PB roleset groups)

For each CP vector calculate similarity with each PB roleset group vector

**Alias:** roleset with the highest similarity score

- The English portion of the Europarl corpus was processed with the MATE tools
- Word alignments (grow-diag-final-and) are taken from the OPUS project
- Pre-defined **syntactic patterns** used for finding occurrences of LVCs and VPCs in the corpus
- **PB roleset group:** all predicates that share a PB roleset (e.g. all verbs that were assigned to *care.01* by the SRL system belong to the PB roleset group of *care.01*)

	cuidar (es)	prendre soin (fr)	penser à (fr)	kümmern (de)
take care	89	71	0	40
care.01	143	36	0	81
think.01	0	0	14K	2

*Excerpt of the co-occurrence matrix*

- Use the cosine similarity to measure the similarity between CPs and PB roleset groups

**Alias PB roleset for the predicate *take care*:**  
**Roleset id:** *care.01* ‘to be concerned’  
**Arg0:** carer, agent  
**Arg1:** thing cared for/about

## Evaluation Setup

- Evaluate on **‘aliased’** and **‘mw. PB pred.’** of the annotation effort (total 154, *Wiki50 set*) + 70 *take* CPs from Bonial et al. [2014] (*take set*)
- **Strict accuracy:** predicted alias corresponds to one of the gold aliases → Synonymous rolesets are counted as incorrect
- **Lenient acc.:** predicted alias belongs to same VerbNet class as gold alias
- **Baseline** selects for every CP first roleset (if available) of the (LVCs) verb predicate corresponding to the noun (*take care* → *care.01*) (VPCs) PB multiword predicate (*open up* → *open.03*)
- Very strong and informed baseline, only fails in lack of coverage

## Results and Discussion

Set	Strict Cov	Strict Acc	Strict Hm	Len. Cov	Len. Acc	Len. Hm
Wiki50	<b>98.7</b> (65.6)	44.1 <b>(54.5)</b>	<b>60.9</b> (59.5)	<b>98.0</b> (59.5)	69.0 <b>(85.9)</b>	<b>81.0</b> (70.3)
alias	<b>98.9</b> (50.0)	<b>36.6</b> (34.0)	<b>53.4</b> (40.5)	<b>98.4</b> (40.5)	60.0 <b>(68.8)</b>	<b>74.5</b> (51.0)
mw PB pred.	<b>98.3</b> (86.7)	55.9 <b>(71.2)</b>	71.3 <b>(78.1)</b>	<b>97.6</b> (84.6)	82.5 <b>(97.7)</b>	89.4 <b>(90.7)</b>
take	67.1 <b>(71.4)</b>	25.5 <b>(32.0)</b>	37.0 <b>(44.2)</b>	56.6 <b>(64.9)</b>	<b>60.0</b> (45.0)	<b>58.3</b> (53.8)

*Percentage coverage (Cov), accuracy (Acc) and the harmonic mean (Hm) of coverage and accuracy of the predicted aliases in the Wiki50 set (+ its two subsets) and the take set; the results of the baseline are in brackets*

- The system always outperforms the baseline in terms of coverage
- Beats the baseline in terms of strict accuracy for the alias subset
- Performance is quite close to the IAA from human annotation

## Error Analysis

1. Domain specificity of the corpus → low frequency of some gold aliases (e.g. *melt down* → *fuse.01* ‘melt into lump’)
2. Errors in the automatic SRL annotation
3. Light verb structure of the CPs remains in other languages
4. No WSD for CPs → predicted alias fits only for CP’s predominant sense

## Conclusions

- The presented approach automatically links CPs to their PB roleset alias
- Annotation effort resulted in frequency-balanced, contextualized, more natural and larger evaluation set than the pilot set by Bonial et al. [2014]
- Method alleviates the manual annotation effort: 44% correct aliases, 69% when taking synonyms into account, not far from upper bound of human annotation

## Future Work

- More parallel data from different domains
- Apply frequency weights to the matrix
- Include more divergent languages
- Retrain SRL system on the new annotations

## More Details

Jagfeld, G. and van der Plas, L. (2015). Towards a better semantic role labeling of complex predicates. NAACL: Student Research Workshop

## Selected References

- [Bonial et al., 2014] Bonial, C., Green, M., Preciado, J., and Palmer, M. (2014). An approach to take multi-word expressions. In *Proc. of the 10th Workshop on Multiword Expressions*.
- [Duran et al., 2011] Duran, M. S., Ramisch, C., Aluisio, S. M., and Villavicencio, A. (2011). Identifying and analyzing brazilian portuguese complex predicates. In *Proceedings of the workshop on multiword expressions: from parsing and generation to the real world*.
- [Palmer et al., 2005] Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1).
- [Van der Plas and Tiedemann, 2006] Van der Plas, L. and Tiedemann, J. (2006). Finding synonyms using automatic word alignment and measures of distributional similarity. In *Proceedings of the COLING/ACL on Main conference poster sessions*.