VPS-GradeUp – Documentation

Version 2.0, *October 10, 2016*

Vít Baisa, Silvie Cinková, Ema Krejčová, Anna Vernerová

Differences between versions

The first version was missing annotations in 10 sentences due to an undiscovered data import error. There were NA 's in these places:

```
##        Lemma SentID
## 2085  adjust   33.1
## 12930   pack    2.1
## 12962   pack    3.1
## 12994   pack    4.1
## 13026   pack    5.1
## 13058   pack    6.1
## 13090   pack    7.1
## 13122   pack    8.1
## 13154   pack    9.1
## 13186   pack   10.1
```

In Line 2085 it was the LikSC and WSDNumSC, all other NA's were in LikAV and WSDNumAV, along with KWIC.

Please cite:

Baisa Vít, Cinková Silvie, Krejčová Ema, Vernerová Anna: VPS-GradeUp: Graded Decisions on Usage Patterns. In: *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, Copyright © European Language Resources Association, Paris, France, ISBN 978-2-9517408-9-1, pp. 823-827, 2016

@inproceedings{ biblio:BaCiVPSGradeUpGraded2016,
booktitle = {Proceedings of the 10th International Conference on Language Resources and Evaluation ({LREC} 2016)},
title = {{VPS}-GradeUp: Graded Decisions on Usage Patterns},
editor = {Nicoletta Calzolari and Khalid Choukri and Thierry Declerck and Marko Grobelnik and Bente Maegaard and Joseph Mariani and Asunci{\'{o}}n Moreno and Jan Odijk and Stelios Piperidis},
author = {V{\'{i}}t Baisa and Silvie Cinkov{\'{a}} and Ema Krej{\v{c}}ov{\'{a}} and Anna Vernerov{\'{a}}},
year = {2016},
publisher = {European Language Resources Association},
address = {Paris, France},
venue = {Grand Hotel Bernardin Conference Center},
pages = {823--827},
isbn = {978-2-9517408-9-1},
}

Introduction

VPS-GradeUp is a collection of triple manual annotations of 29 English verbs[1] based on the Pattern Dictionary of English Verbs (PDEV)[1]. It contains results from two different tasks:

1. Graded decisions
2. Best-fit pattern (WSD) .

In both tasks, the annotators were matching verb senses defined by the PDEV patterns with 50 actual uses of each verb (using concordances from the BNC [2]). The verbs were randomly selected from a list of completed PDEV lemmas with at least 3 patterns and at least 100 BNC concordances not previously annotated by PDEV's own annotators. Also, the selection excluded verbs contained in VPS-30-En[3].

The annotators were all trained linguists familiar with PDEV, but they were not English native speakers.

The Graded Decisions task

In this task, the annotators were each given a PDEV entry concerning the given verb lemma and a BNC concordance with this lemma being the keyword in context (KWIC). For each PDEV pattern, the annotator answered the question "How well does this pattern illustrate the use of the target verb in this context?" using a 7-point Likert scale. The Likert scale had the following anchors:

1 = Irrelevant 2 = Somewhat relevant but poor match 3 = Literal match but there is a meaning shift (in idioms)/there is a substantial meaning shift (other cases) 4 = Certainly a match, but quite many things wrong 5 = Partial match - different domain/granularity/atypical arguments 6 = Good match – just extend pattern definition 7 = Exact match

There was a consensus before the annotation started that, whenever it was possible to say whether the reason for a mismatch was rather syntactic or semantic, the semantic mismatches should be considered graver than the syntactic ones.

The annotators should also judge their own comprehension of the concordance (1 = understood, 0 = comprehension issues).

For each verb, there were 50 concordances to be judged against each pattern in the corresponding PDEV entry (that is, the number of graded decisions varies per verb according to the number of patterns it has in PDEV). Each verb was treated in a separate online survey form, where each concordance was dedicated a separate page and an additional optional page with exactly the same questions. The alternative page was there to capture an alternative reading of the concordance in case the concordance had been distinctly ambiguous. The alternative reading was strictly dedicated

[1] The following verbs: *abolish, act, adjust, advance, answer, approve, bid, cancel, conceive, cultivate, cure, distinguish, embrace, execute, hire, last, manage, murder, need, pack, plan, point, praise, prescribe, sail, seal, see, talk, urge*

to evident ambiguity cases with no other comprehension issues. There are only a few cases in the data. Without any alternative readings, the data set contains 11,400 graded decisions.

The Best-fit Pattern (WSD) task

This task was the classical Word-Sense-Disambiguation task: for each KWIC the annotators assigned the ID of the best-fitting PDEV pattern, with the options to say that no pattern was really fitting (*unclassified*) or that the keyword was not a verb (*not verb*) and therefore the verbal patterns were all irrelevant. In addition, the annotators could specify their reservations with respect to the syntactic definition of the PDEV pattern, the Semantic Types, or diagnose a meaning shift.

Data format

VPS-GradeUp comes as a single .csv file separated with semicolons and each cell enclosed by double quotes, encoded in UTF 8. It contains 22,800 rows and 43 columns with a header.

Rows

Each row primarily represents one observation in the Graded-Decision experiment; i.e. one score on a 7-point Likert scale rendering how well a given PDEV pattern (e.g. [*Pattern*] *3*) of a given verb lemma (e.g. *abolish*) illustrates a given KWIC identified by an index (e.g. *2.1*).

Graded decisions are filled in all rows containing `.1` at the end of the KWIC index (Column `SentID`). Most rows with KWICs indexed with `.2` do not contain any graded decisions (NA filled in). These rows contain the unused alternative readings.

Each row also contains all WSD (best-fit) annotation related to the given KWIC; i.e. the WSD information repeats for each KWIC as many times as the given verb has PDEV patterns. To explore the WSD results independently of the graded decisions, mind to eliminate duplicate rows.

Columns

Column name	Description	Example
`JointID`	Unique ID for each row containing lemma, KWIC ID, and pattern number	`abolish:Sent_1.1:Pattern_1`
`PatternID`	NB: only unique in combination with the Lemma column; when working with all lemmas, use JointID!	`1`
`Lemma`		`abolish`
`SentID`	NB: only unique in combination with the Lemma column; when working with all lemmas, use JointID!	`1.1`

`LikAV`
`LikEK`
`LikSC`
	Score on the 7-point Likert scale saying how well the given PDEV pattern illustrates the given KWIC according to the	`7`
		`5`
		`7`

	annotator identified by their initials. 1 = Irrelevant, 7 = Perfect match.	
WSDNumAV WSDNumEK WSDNumSC	For each annotator separately: ID of the best-fitting pattern in a classical WSD setup, when the annotator is forced to select only one pattern, or claim that the given KWIC is not a verb (value *not verb*) or that no pattern is really suitable (*unclassified*).	2 3 3
UnderstandAV UnderstandEK UnderstandSC	For each annotator, the options are 1 and 0. (1 = the annotator is confident that they understand the KWIC well, 0 indicates comprehension problems)	1 1 1
KWIC	The annotated BNC KWIC – the largest span allowed by BNC. The key word is capitalized and surrounded by three spaces on both sides. Apostrophes and double quotes are escaped. Horizontal ellipsis is rendered by the corresponding HTML entity … (as copied from the BNC).	Anna Tomforde and Michael Farr PRESIDENT Franois Mitterrand , the first head of state of the wartime Allies to visit East Germany , said yesterday that the existence of two sovereign German states could not be `  ABOLISHED  at a stroke \' . Reflecting French anxiety over German reunification , Mr Mitterrand said the two Germanys were jointly responsible for stability in Europe . ` German unity depends first of all on the German people …
BNCdocID	The document code the KWIC was associated with in the BNC	AAK/1
Number of Patterns	How many patterns (senses) the given verb lemma has in PDEV (Pattern Dictionary of English Verbs)	3
CommentsAV CommentsEK CommentsSC	Annotators' comments. Most of them are in English, but some are in Czech.	NA
WSDExploitAV_ coercion_agent WSDExploitEK_ coercion_agent WSDExploitSC_ coercion_agent	Exploitation markup. Binary values. 1 = the agent of the keyword was coerced into a different PDEV Semantic Type, although it actually corresponds to the Semantic Type listed in the pattern definition. 0 = no markup	0 (1 would occur e.g. if the pattern definition contained the Semantic Type Liquid for agent and the KWIC said: *The second cup poured on the floor*. Although, strictly speaking, *cup* corresponds to Container, Liquid is evidently meant at the same time.
WSDExploitAV_ coercion_object WSDExploitEK_ coercion_object WSDExploitSC_ coercion_object	Cf. coercion agent above. Applies to direct object. Binary (0,1).	

WSDExploitAV_coercion_other WSDExploitEK_coercion_other WSDExploitSC_coercion_other	Cf. coercion agent above. Typically applies to indirect object and adverbials, but it can apply to any clause element except agent and object. Binary (0,1).	
WSDExploitAV_meaning_shift WSDExploitEK_meaning_shift WSDExploitSC_meaning_shift	Exploitation markup indicating any type of meaning shift between the implicature of the selected pattern (sense) and the KWIC; e.g., metaphor or any rhetorical figure. Binary (0,1).	
WSDExploitAV_unexpected_agent WSDExploitEK_unexpected_agent WSDExploitEK_unexpected_agent	Exploitation markup indicating that the agent of the given KWIC does not conform to the Semantic Type prescribed by PDEV. A more general markup than coercion. Binary (0,1).	
WSDExploitAV_unexpected_object WSDExploitEK_unexpected_object WSDExploitEK_unexpected_object	Cf. unexpected agent and coercion object above, applies to direct object. Binary (0,1).	
WSDExploitAV_unexpected_other WSDExploitEK_unexpected_other WSDExploitEK_unexpected_other	Cf. unexpected agent and coercion other above, applies to indirect object and all other clause elements except agent and direct object. Binary (0,1).	

References

[1] P. Hanks and J. Pustejovsky, "A Pattern Dictionary for Natural Language Processing," *Rev. Francaise Linguist. Appliquée*, vol. 10, no. 2, 2005.

[2] "British National Corpus, version 3 (BNC XML edition)." British National Corpus Consortium, 2007.

[3] S. Cinková, M. Holub, A. Rambousek, and L. Smejkalová, "A database of semantic clusters of verb usages," in *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012)*, \.Istanbul, Turkey, 2012, pp. 3176–3183.

Papers on VPS-GradeUp

Cinková Silvie, Krejčová Ema, Vernerová Anna, Baisa Vít: What Do Graded Decisions Tell Us about Verb Uses. In: *Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity*, Copyright © Tbilisi University Press, Tbilisi, Georgia, ISBN 978-9941-13-542-2, pp. 318-328, 2016 – url: http://euralex2016.tsu.ge/proceedings-p10.pdf

Cinková Silvie, Krejčová Ema, Vernerová Anna, Baisa Vít: Graded and Word-Sense-Disambiguation Decisions in Corpus Pattern Analysis: a Pilot Study. In: *Proceedings of the 10th International Conference on Language*

Resources and Evaluation (LREC 2016), Copyright © European Language Resources Association, Paris, France, ISBN 978-2-9517408-9-1, pp. 848-854, 2016. url: http://www.lrec-conf.org/proceedings/lrec2016/index.html
Baisa Vít, Cinková Silvie, Krejčová Ema, Vernerová Anna: VPS-GradeUp: Graded Decisions on Usage Patterns. In: *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, Copyright © European Language Resources Association, Paris, France, ISBN 978-2-9517408-9-1, pp. 823-827, 2016 url: http://www.lrec-conf.org/proceedings/lrec2016/index.html

Bibtex citations

```
@inproceedings{ biblio:CiKrWhatDo2016,
booktitle = {Proceedings of the {XVII} {EURALEX} International Congress: Lexicography and Linguistic Diversity},
title = {What Do Graded Decisions Tell Us about Verb Uses},
editor = {Tinatin Margalitadze and George Meladze},
author = {Silvie Cinkov{\'{a}} and Ema Krej{\v{c}}ov{\'{a}} and Anna Vernerov{\'{a}} and V{\'{i}}t Baisa},
year = {2016},
publisher = {Tbilisi University Press},
organization = {Ivane Javakhishvili Tbilisi State University},
address = {Tbilisi, Georgia},
pages = {318--328},
isbn = {978-9941-13-542-2},
}

@inproceedings{ biblio:CiKrGradedand2016,
booktitle = {Proceedings of the 10th International Conference on Language Resources and Evaluation ({LREC} 2016)},
title = {Graded and Word-Sense-Disambiguation Decisions in Corpus Pattern Analysis: a Pilot Study},
editor = {Nicoletta Calzolari and Khalid Choukri and Thierry Declerck and Marko Grobelnik and Bente Maegaard and Joseph Mariani and Asunci{\'{o}}n Moreno and Jan Odijk and Stelios Piperidis},
author = {Silvie Cinkov{\'{a}} and Ema Krej{\v{c}}ov{\'{a}} and Anna Vernerov{\'{a}} and V{\'{i}}t Baisa},
year = {2016},
publisher = {European Language Resources Association},
address = {Paris, France},
venue = {Grand Hotel Bernardin Conference Center},
pages = {848--854},
isbn = {978-2-9517408-9-1},
}

@inproceedings{ biblio:BaCiVPSGradeUpGraded2016,
booktitle = {Proceedings of the 10th International Conference on Language Resources and Evaluation ({LREC} 2016)},
title = {{VPS}-GradeUp: Graded Decisions on Usage Patterns},
editor = {Nicoletta Calzolari and Khalid Choukri and Thierry Declerck and Marko Grobelnik and Bente Maegaard and Joseph Mariani and Asunci{\'{o}}n Moreno and Jan Odijk and Stelios Piperidis},
author = {V{\'{i}}t Baisa and Silvie Cinkov{\'{a}} and Ema Krej{\v{c}}ov{\'{a}} and Anna Vernerov{\'{a}}},
year = {2016},
publisher = {European Language Resources Association},
address = {Paris, France},
venue = {Grand Hotel Bernardin Conference Center},
pages = {823--827},
isbn = {978-2-9517408-9-1},
}
```