

Quantifying similarities in the interaction between prefix, base class, and individual base in scalar *out*-prefixation

Sven Kotowski & Martin Schäfer

Morfologiedagen 2020, Utrecht/NL

April 23, 2021

Universities of Düsseldorf & Tübingen

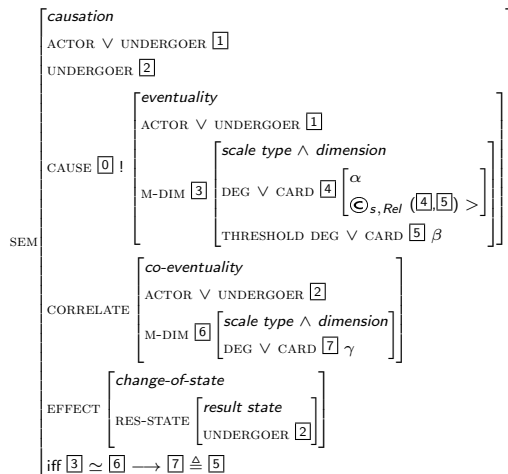


EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



- Highly productive verbal prefix
 - Introduces own argument structure and argument restrictions
 - Overall constructional semantics contested, scalar-comparative meaning component undisputed
- (1) We try to **outdrink our friends** and end up as alcoholics. (COCA)
 - (2) ??We try to **outdrink the beer**.

(see Ahn 2019; Kotowski 2020, Talmy 2000)



- Scalar dimensions basis for comparison (e.g. SPEED or HEIGHT)
 - One lemma \neq one dimension; often underspecified
 - Allows for assymmetric comparison
- (3) ...a backing singer must not try to **'out-sing' the lead vocalist** [...] producers will adjust vocal volumes... (iWeb)
- (4) ...you can't deny their vocal abilities [...] nobody out there can **outsing them**... (iWeb)
- (5) "I figured I'm not going to **outrun an eagle**, so we might as well just see what happens." (forbes.com)

(see Kotowski 2020)

- Division of labor between base, base class, and constructional semantics?
- Do verb classes steer interpretation wrt which dimensions are accessed for comparisons?
- *out*-derivatives highly similar in general? Possibly more similar to each other than base-derivative pairs?
- Different from other prefixes in this regard?

- 3 studies that look at base-base, base-derivative, and derivative-derivative similarities
- Study 1: *out*-lemmas and the dimensions they encode
- Study 2 (distributional measures): does *out*- lead to higher degrees of similarity in derivatives?
- Study 3 (distributional measures): findings general features of prefixation? Comparison with other prefixes

Study 1: verb classes, tokens, and dimensions

- iWeb (Davies 2018)
- 3 classes (PERFORMANCE, RUN, EXIST) from VerbNet (Kipper et al. 2008)
- 100 items (if available) each of 4 verbs per class
- Hypothesis: dimension profiles of lemmas cluster per VN class

Class	Verbs of	Base verbs study 1
PERFORMANCE	Creation and Transformation	<i>dance, rap, sing, write</i>
RUN	Motion	<i>fly, run, sprint, swim</i>
EXIST	Existence	<i>live, survive, stay, wait</i>

Study 1: results

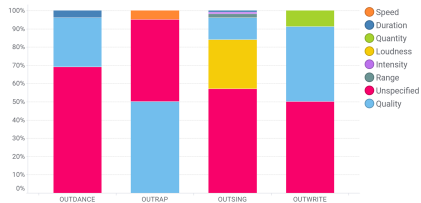


Figure: PERFORMANCE-lemmas

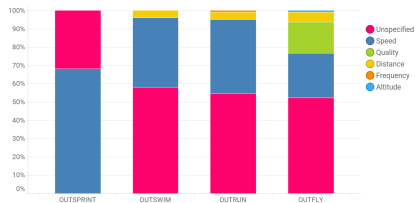


Figure: RUN-lemmas

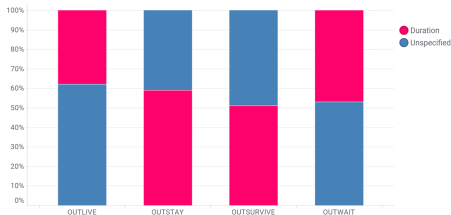


Figure: EXIST-lemmas

Study 1: discussion

- All classes: relative homogeneity wrt majority dimensions
 - Attestations with unspecified dimensions follow majority dimension
 - Each class: distinct dimension profile (items with less clear profile: cross-listing in other productive *out*-base class)
 - Unclear: apparently, similar bases → similar semantics (i.e. the dimension profiles)
 - Not reflected: hypothesized effect of causative semantics of *out*-
- **distributional similarity measures**

Study 2: *out-*, distributional similarities, and VN-classes

- Distributional measures (Sahlgren 2006), calculated from ukWaC (Ferraresi et al. 2008) and iWeb
- Cosine similarity measures within and across VN classes
- Comparing base-base, base-derived, and derived-derived similarities
- Same classes and lemmas as in study 1
- Expectations
 - Derivatives relatively similar to each other
 - Effect even more pronounced within classes (dimension profiles)
 - Expected similarities: base-base < derived-derived; base-derived unclear

Study 2: results & discussion

pairings	RUN		EXIST		PERFORM		all items	
	SIM	SD	SIM	SD	SIM	SD	SIM	SD
base-derived pairwise	0.08	0.02	0.05	0.00	0.05	0.02	0.06	0.02
base-derived pair-less	0.05	0.01	0.04	0.01	0.03	0.02	0.04	0.01
base-base	0.04	0.01	0.04	0.01	0.05	0.03	0.04	0.01
derived-derived	0.10	0.02	0.06	0.01	0.06	0.02	0.07	0.03

- **base-derived (pairwise)** more similar than **base-derived (pair-less)**: significant for all items + EXIST-class

→ Base semantics still relevant in derivative semantics

- **base-derived (pair-less)** less similar than **derived-derived**: significant for all but PERFORM-class

- **base-derived (pairwise)** less similar than **derived-derived**: not significant

→ Uniform semantics of derivatives

Study 3: distributional similarities *out-* v other prefixes

- General question: high degree of similarity for derived-derived feature of *out-* or general prefixation effect?
- Similarity measures for three further prefixes: spatial *over-* (+ RUN-class) as well *un-* and *re-* (both + TAPE-class)
- Different argument structural behavior than *out-*
 - Spatial *over-*: preposition incorporation (*fly over the field* \simeq *overfly the field*)
 - *un-* and *re-*: no universal effects (*seal a gasket* - *unseal a gasket* - *reseal a gasket*)
- Expectation for all three prefixes: lower (relative) derived-derived similarities than for *out-*

Study 3: results

overfly, overrun, overdrive, overstep

	RUN (+ <i>out</i>)		RUN (+ <i>over</i>)	
	SIM	SD	SIM	SD
base-derived pairwise	0.08	0.02	0.06	0.02
base-derived pair-less	0.05	0.01	0.04	0.01
base-base	0.04	0.01	0.05	0.01
derived-derived	0.10	0.02	0.04	0.02

reseal, relock, rewind, refasten

unseal, unlock, unwind, unfasten

	TAPE (+ <i>un</i>)		TAPE (+ <i>re</i>)	
	SIM	SD	SIM	SD
base-derived pairwise	0.13	0.07	0.14	0.07
base-derived pair-less	0.06	0.02	0.07	0.02
base-base	0.07	0.02	0.07	0.02
derived-derived	0.05	0.01	0.09	0.03

- *out*-prefixation comes with verb-specific, class-specific, and construction-specific effects
- Study 1: dimension-based similarity for derivatives with same base class
- Study 2: base still semantically visible in derivative; however, derivatives semantically even more coherent
- Study 3: distributional behavior based on *out*-, not a general feature of prefixation (possibly down to *out*'s applicative nature)

→ Similarity measures speak in favor of **word-formation process as most pronounced semantic contributor to out-derivatives**

THANK YOU

Partly supported by Deutsche Forschungsgemeinschaft (Grants SFB 991/2-C08 and PL 151/11-1 'Semantics of derivational morphology' to Ingo Plag)

References

- Ahn, Byron. Submitted. *out*-PRED: Generalizations and Derivation.
- Davies, M. 2018-. *The 14 Billion Word iWeb Corpus*. Available online at <https://corpus.byu.edu/iWeb/>.
- Ferraresi, A., E. Zanchetta, M. Baroni, and S. Bernardini. 2008. Introducing and evaluating ukWaC, a very large web-derived corpus of English. In *Proceedings of the WAC4 Workshop at LREC 2008*, Marrakech. ELRA.
- Kipper, K., A. Korhonen, N. Ryant & M. Palmer. 2008. A large-scale classification of English verbs. *Language Resources and Evaluation* 42(1). 21–40.
- Kotowski, Sven. 2020. The semantics of English *out*-prefixation: A corpus-based investigation. *English Language and Linguistics*, 1-29.
- Sahlgren, Magnus. 2006. The Word-Space Model. Using distributional analysis to represent syntagmatic and paradigmatic relations between words in high dimensional vector spaces. Ph.D. thesis, Stockholm University.
- Talmy, Leonard. 2000. *Toward a cognitive semantics, vol. II: Typology and process in concept structuring* (Language, Speech, and Communication). Cambridge, MA: MIT Press.

VerbNet classes: dimension distribution

