

PD Dr. Sabine Schulte im Walde

Institut für Maschinelle Sprachverarbeitung (IMS)
Universität Stuttgart

Pfaffenwaldring 5B
70569 Stuttgart

schulte@ims.uni-stuttgart.de

Kognitive und Distributionelle Perspektiven auf deutsche Partikelverben

Deutsche Partikelverben repräsentieren eine äußerst produktive Klasse von komplexen Verben im Lexikon, die sich durch idiosynkratische Eigenschaften auf der Syntax-Semantik-Schnittstelle auszeichnet: Zum einen sind die abtrennbaren Partikeln extrem ambig. Zum anderen entstehen durch die Komposition von Partikel und Basisverb in Abhängigkeit von der semantischen Klasse des Basisverbs (reguläre) Verschiebungen bezüglich der Argumentstruktur und des Grades der Kompositionalität des Partikelverbs. In diesem Artikel stelle ich eine breite Auswahl von kognitiven und computerlinguistischen Studien vor, die verschiedene Perspektiven auf das Zusammenspiel von semantischen Verbklassen, Partikel-Bedeutungen sowie Argumentstruktur und Kompositionalität von Partikelverben ermöglichen.

1. Einleitung

Deutsche Partikelverben repräsentieren eine äußerst produktive Klasse von komplexen deutschen Verben, die aus Präfix-Partikeln und Basisverben zusammengesetzt werden. Die Partikeln können den Haupt-Wortklassen entnommen werden und ein- oder zweistellig sein. Der Fokus dieser Arbeit liegt auf Partikelverben mit einstelligen präpositionalen Partikeln. Um ein umfassendes Bild über deren Eigenschaften zu bekommen, befassen wir uns in den meisten perspektivischen Betrachtungen mit den 11 Partikeln *ab, an, auf, aus, durch, ein, hinter, über, um, unter, zu*. Ein besonderer Fokus liegt auf den drei Partikeln *ab, an* und *auf*, für die uns bereits detaillierte theoretische semantische Beschreibungen vorlagen (Kliche, 2009; Lechler & Roßdeutscher, 2009; Springorum, 2011).

Deutsche Partikelverben stellen aus empirischer Sicht aus mehreren Gründen eine Herausforderung dar. Im Folgenden stellen wir einen kurzen Überblick vor und beschränken uns dabei auf die Phänomene, die für diesen Artikel relevant sind. Ausgiebige theoretische Untersuchungen zu deutschen Partikelverben können zum Beispiel Stiebels & Wunderlich (1994), McIntyre (2007) oder Dehe (2015) entnommen werden.

Wie in den Beispielen (1a-c) vs. (1d) dargestellt, können Partikelverben sowohl zusammen als auch getrennt geschrieben werden. Die Zusammen-/Getrennschreibung hängt grundsätzlich vom Satzstrukturtyp ab und von der morphologischen Realisierung des Partikelverbs. Für Korpus-basierte Untersuchungen stellt die Getrennschreibung ein Problem dar, da eine abgetrennte Partikel ambig ist in Bezug auf ihre Wortart und daher selbst nach syntaktischer Annotation durch einen Parser keine sichere Wortart-Zuweisung erfolgen kann.

- (1) (a) Das Kind wird/will *ankommen*.
(b) Das Kind ist *angekommen*.
(c) Er bittet das Kind, pünktlich *anzukommen*.
(d) Das Kind *kommt* pünktlich *an*.

Ebenso eine empirische Herausforderung stellen die Verschiebungen der Argumentstruktur von Partikelverben in Bezug auf ihre Basisverben dar. Beispiel (2a) stellt eine Instanz von Argument-Sättigung vor: Die Partikel *an* des Partikelverbs „*anschrauben*“ sättigt die Präpositionalphrase „*auf die Öffnung*“, was beim Basisverb „*schrauben*“ nicht möglich ist. Beispiel (2b) stellt eine Instanz von Argument-Erweiterung vor: Das Partikelverb „*anlächeln*“ erlaubt ein direktes Objekt „*ihre Mutter*“, was beim Basisverb „*lächeln*“ nicht möglich ist. Obwohl diese Verschiebungen in den meisten Fällen regelmäßig sind, verlangen sie spezifische Korpus-basierte Vorverarbeitungen, denn die grundsätzlich empirisch wertvolle Annahme von Levin (1993), dass Verbverhalten einher geht mit Verbbedeutung, ist nur eingeschränkt einsetzbar.

- (2) (a) Der Chef *schraubt* den Deckel auf die Öffnung.
*Der Chef *schraubt* den Deckel.
Der Chef *schraubt* den Deckel *an*.
- (b) Sie *lächelt*.
*Sie *lächelt* ihre Mutter.
Sie *lächelt* ihre Mutter *an*.

Auf der Bedeutungsebene sind die Partikeln semantisch äußerst ambig. Zum Beispiel hat die Partikel *an* in „*anbeißen*“ eine partitive Bedeutung, in „*anhäufen*“ eine kumulative Bedeutung und in „*anbinden*“ eine topologische Bedeutung (Springorum, 2011). Außerdem weisen die Partikelverben als Ganzes unterschiedliche Grade von Kompositionalität in Bezug auf ihre Basisverben auf. Ohne diachrone Untersuchungen hinzuzuziehen würde man zum Beispiel intuitiv dem Verb „*abholen*“ eine relativ kompositionelle Bedeutung zuschreiben, dem Verb „*anfangen*“ eine nicht-kompositionelle Bedeutung, und bei dem Verb „*einsetzen*“ würde man mindestens eine eher kompositionelle Bedeutung (im Sinne von „*in etwas einsetzen/einstecken*“) und eine eher nicht-kompositionelle Bedeutung (im Sinne von „*beginnen*“) feststellen.

In diesem Artikel stelle ich eine breite Auswahl von Studien vor, die verschiedene Perspektiven auf das Zusammenspiel von semantischen Verbklassen, Partikel-Bedeutungen sowie Argumentstruktur und Kompositionalität von Partikelverben

ermöglichen: Für Kapitel 2 habe ich zwei kognitive Studien ausgewählt. Ein Produktions-Experiment untersucht die Interaktion von Partikel-Bedeutungen und semantischen Basisverb-Klassen bezüglich Partikelverb-Neologismen. Ein Priming-Experiment spezialisiert sich auf die spatialen Bedeutungen der Verb-Partikeln *an* und *auf* und erforscht aus spatialen Konflikten resultierende Bedeutungsverschiebungen. Für Kapitel 3 habe ich zwei komputationelle Studien ausgewählt. Nach einer kurzen Einleitung in die Thematik von distributionellen Modellen untersuchen Klassifikations-Experimente die Kompositionalität und (nicht-)wörtliche Sprachverwendung von deutschen Partikelverben auf der Basis von Korpus-Daten, Bild-Informationen und affektiven Normen, mit und ohne Einbezug von syntaktischen Alternationen.

2. Kognitive Perspektiven

2.1 Kompositionalitätsforschung anhand von Neologismen

Im Hinblick auf die Forschungsfragen, ob deutsche Partikelverben generell kompositionell sind, und ob es prototypische Bedeutungsbeiträge der Partikeln und der (semantischen Klassen der) Basisverben gibt, haben wir eine Produktionsstudie durchgeführt, die als Datenbasis frei verfügbar ist (Springorum, 2014): Wir haben für fünf nicht notwendigerweise disjunkte semantische Klassen jeweils fünf Basisverben ausgewählt, die mit den fünf Partikeln *ab*, *an*, *auf*, *aus* und *nach* zu 125 Partikelverben kombiniert wurden. Aufgrund der systematischen Kompositionsmethode entstanden sowohl existierende als auch nicht-existierende Partikelverben. Tabelle 1 gibt einen Überblick über die Basisverben und ihre Klassen.

<i>Semantische Klasse</i>	<i>Basisverben (Beispiele)</i>
De-Adjectival	<i>kürzen, röten, töten</i>
Achievement/Accomplishment	<i>finden, platzen, treffen</i>
Physical Process	<i>nageln, rühren, stricken</i>
Mental Process	<i>denken, meinen, sorgen</i>
State	<i>freuen, wissen, wundern</i>

Tabelle 1: Systematische Komposition von Partikelverben: semantische Klassen und Basisverben.

Die Partikelverben wurden als Typen (also ohne Kontext) Teilnehmern in einer Produktionsstudie vorgelegt. Die Aufgabe war, gezwungenermaßen zu jedem Partikelverb einen Satz zu generieren. Unsere Hypothesen zur Interpretation der Sätze waren wie folgt: (i) Wenn Partikelverben kompositionell und produktiv sind, sollten auch die Bedeutungen der Partikelverb-Neologismen in den generierten Sätzen interpretierbar sein. (ii) Sofern den Sätzen mit Partikelverb-Neologismen eine Interpretation zugewiesen werden kann, sollte diese regelbasiert und nicht idiosynkratisch sein.

Die generierten Sätze werden zur Zeit noch systematisch analysiert, sowohl im Hinblick auf lexikalisch-semantische Beschreibungen als auch im Hinblick auf Korpus-basierte Untersuchungen zu Bedeutungsindikatoren. Im Folgenden führe ich einige Beispielsätze auf, zusammen mit vorläufigen Ideen für die Interpretation. In den Beispielen (3) und (4) finden wir zwei Kontexte, die für den Neologismus „*anlauschen*“ generiert wurden. Satz (3) kann eine Interpretation bezüglich topologischen Kontakts zugewiesen werden, der Kontext in (4) bezieht sich auf eine partitive Lesart der Zeit-Dimension. Die meisten generierten Sätze mit „*anlauschen*“ können tatsächlich einer dieser beiden Interpretationen zugewiesen werden. In Satz (5) haben wir eine Interpretation von „*nachwundern*“, die die Fortführung eines Zustands charakterisiert, analog zu der Fortführung eines Prozesses wie zum Beispiel beim existierenden Verb „*nachreifen*“. In Satz (6) haben wir einen Satz zu einer eigentlich implausiblen Partikelverb-Kombination „*antöten*“ von partitiver Partikelbedeutung mit einem Basisverb der Zustandsänderung. Satz (7) zeigt, dass auch metaphorische Interpretationen generiert wurden, wie hier für „*abnageln*“ analog zu „*abschminken*“. In beiden Fällen findet die Separation eines abstrakten Kontakt-Basisverbs statt.

- (3) Er hatte an der Wand *angelauscht* und wusste Bescheid.
- (4) Ich muss das Gespräch *anlauschen*. Dann weiß ich mehr.
- (5) Ich musste mich noch lange Zeit *nachwundern*.
- (6) Ich werde den Zombie schon mal *antöten*, damit du ihn erledigen kannst.
- (7) Das Gespräch lief nicht gut, den Job kann ich mir *abnageln*.

Die Teilnehmer in der Studie haben zusätzlich zur Generierung des Satzes eine Angabe dazu gemacht, ob ihnen das Partikelverb bekannt war oder nicht, und ob es schwierig zu verstehen war oder nicht (jeweils auf einer 4-Punkt-Skala). Diese Information können wir nutzen, um Aussagen über das Vorliegen eines Partikelverb-Neologismus zu machen.

Zusammenfassend kann zu diesem Zeitpunkt berichtet werden, dass Teilnehmer in einer Produktionsstudie sowohl literale als auch nicht-literale Sätze für Partikelverb-Neologismen bilden konnten. Die generierten Sätze ermöglichen uns die Erkennung von Konstruktionsmustern, und Analogien zu existierenden Partikelverben werden offensichtlich. Kontext-Untersuchungen und -Vergleiche werden die Gemeinsamkeiten von Partikel-Basisverb-Kombinationen vertiefen, und ein Priming-Experiment wird die (Nicht-)Plausibilität von Partikelbedeutungen und semantischen Basisverb-Klassen weiter untersuchen.

2.2 Spatiale Bedeutungen der Partikeln "an" und "auf"

Eine zweite kognitive Studie hatte als Ziel, in die Bedeutungen von zwei spezifischen Partikeln zu zoomen (Frassinelli et al., 2017). Unsere Hypothese war, dass die Partikel *an* eine starke horizontale Richtungsbedeutung hat, während die Partikel *auf* eine starke vertikale Richtungsbedeutung hat (ohne dabei gänzlich andere Bedeutungen auszuschließen). Auf Basis dieser Hypothese haben wir bei der Kombination von *an* und Basisverben mit horizontaler Richtungsbedeutung (zum Beispiel „*schieben*“) einen Match zwischen den Bedeutungen von Partikel- und Basisverb-Bedeutungen (mit tendenziell literaler Partikelverb-Lesart) erwartet, während wir bei der Kombination von *an* und Basisverben mit vertikaler Richtungsbedeutung einen Mismatch zwischen den Bedeutungen von Partikel- und Basisverb-Bedeutungen (mit tendenziell nicht-literaler Partikelverb-Lesart) erwartet haben. Entsprechend für *auf*. Anhand eines Priming-Experiments haben wir getestet, ob Verben mit konzeptueller Übereinstimmung (Match) schneller verarbeitet werden („*facilitation*“) als Verben mit konzeptueller Nicht-Übereinstimmung (Mismatch: „*inhibition*“).

Für die Auswahl der Experiment-Verben haben deutsche Muttersprachler 230 Basisverben bezüglich ihrer prominentesten Richtungsbedeutung (horizontal, vertikal) klassifiziert. Pro Basisverb gab es 15 Bewertungen. Anhand der Klassifikationen wurden die Basisverben hinsichtlich ihrer präferierten Richtungsbedeutung kategorisiert. Zum Beispiel wurde das Basisverb „schieben“ zu 80% als horizontales Basisverb klassifiziert und somit der Kategorie „horizontal“ zugewiesen. Insgesamt haben wir 11 horizontale und 11 vertikale Basisverben ausgewählt, die in Kombination mit den beiden Partikeln *an* und *auf* vorkommen. Die horizontalen Basisverben bilden zusammen mit der Partikel *an* einen Partikelverb-Match und zusammen mit der Partikel *auf* einen Mismatch. Entsprechend für die vertikalen Basisverben.

Die Experiment-Durchführung beinhaltete eine Lexical Decision Task, mit den Partikeln als Primes und den Basisverben als Targets. 66 deutsche Muttersprachler haben an dem Experiment teilgenommen. Abbildung 1 zeigt den Unterschied in der Reaktionszeit zwischen Match- und Mismatch-Kombinationen, $\beta_{\text{mismatch}} = 0,05$, $p < 0,001$. Die Teilnehmer verarbeiten ein Verb in einer Match-Bedingung signifikant schneller ($729 \pm 10\text{ms}$ Standardabweichung) als in einem Mismatch ($768 \pm 8\text{ms}$).

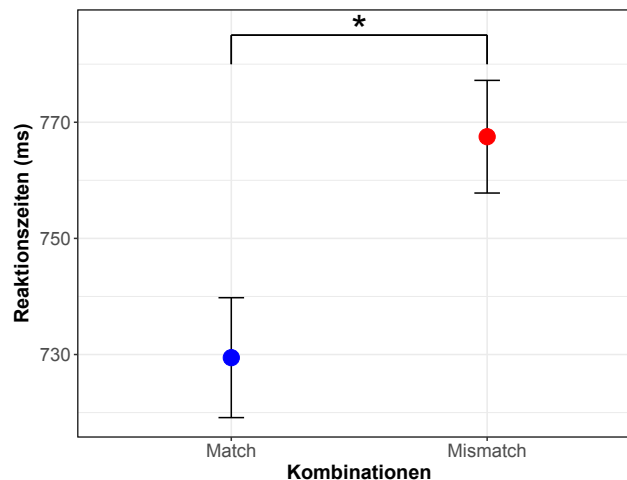


Abbildung 1: Reaktionszeiten bei Match- vs. Mismatch-Bedingungen.

Eine qualitative Analyse hat für die einzelnen Partikelverben untersucht, ob die Partikelverben in der Mismatch-Bedingung abstrakter sind als in der Match-Bedingung, weil der Mismatch eventuell mit einer Bedeutungsverschiebung einher geht. Der Grad der Abstraktheit wurde dabei in einer semi-automatisch erstellten Datenbasis abgelesen (Köper & Schulte im Walde, 2016a). Insgesamt wurde unsere Vermutung bei 13 von 22 Partikelverben bestätigt.

Abbildung 2 zeigt sowohl die Unterschiede in Match/Mismatch-Reaktionszeiten für alle horizontale Basisverben als auch die Unterschiede bezüglich des Grades der Abstraktheit. In der horizontalen Ebene wird der Unterschied in Reaktionszeiten dargestellt. Bei den positiven Fällen (rechts der Null-Achse) war die Reaktionszeit bei einem Match geringer als bei einem Mismatch. Die grünen Balken zeigen an, dass der Grad der Abstraktheit bei einem Mismatch größer war (und somit auf eine Bedeutungsverschiebung hindeutet); die roten Balken zeigen an, wann der Grad der Abstraktheit entgegen unserer Hypothese bei einem Match größer war. Bei den Verben in den roten Balken liegen interessanterweise oft mehrere Bedeutungen vor, von denen mindestens eine einer Bedeutungsverschiebung unterliegt und mindestens eine synonym ist zu der des entsprechenden Match-Partikelverbs. Zum Beispiel ist „*aufdrehen*“ im Sinne von „*den Hahn aufdrehen*“ synonym zu einer Bedeutung von „*andrehen*“, aber das Verb hat auch eine nicht-literale Bedeutung im Sinne von „*in Stimmung kommen*“. Zur Zeit untersuchen wir diese neue Hypothese anhand von Partikelverb-Sätzen mit Bedeutungsannotationen.

Zusammenfassend können wir sagen, dass Partikel-Basisverb-Paare mit spatialer Bedeutungs-Übereinstimmung signifikant schneller verarbeitet werden als Paare mit spatialer Nichtübereinstimmung. Wir folgern daraus, dass die Partikeln *an* und *auf* starke horizontale bzw. vertikale Richtungsbedeutungen haben. Auf Basis einer qualitativen Analyse vermuten wir, dass dieser Effekt außerdem mit einer Bedeutungsverschiebung der ambigen Partikelverben einhergeht. Dieser Vermutung wird aktuell durch eine Bedeutungsannotation literaler und nicht-literaler Partikelverb-Kontexte nachgegangen.

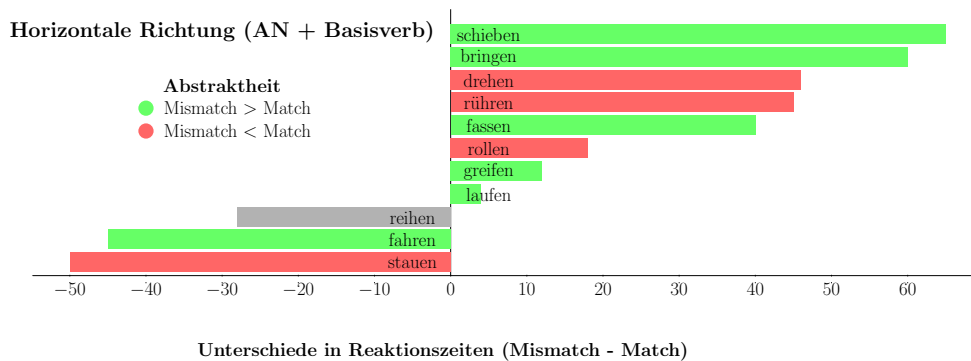


Abbildung 2: Unterschiede in Reaktionszeiten und Abstraktheit.

3. Distributionelle Perspektiven

3.1 Distributionelle Modelle

Distributionelle Modelle gehen zurück auf mindestens zwei grundlegende Zitate.

„Each language can be described in terms of a distributional structure, i.e., in terms of the occurrence of parts relative to other parts.“ (Harris, 1954)

„You shall know a word by the company it keeps.“ (Firth, 1957)

Beide Zitate sagen im Kern aus, dass der Kontext einer linguistischen Einheit Indikatoren für deren Verwendung und Bedeutung enthält. Die Summe von Kontext-Beschreibungen einer linguistischen Einheit würde in einer idealisierten Vorstellung dann der lexikalischen Beschreibung dieser Einheit entsprechen. Korpus-basierte quantitative Untersuchungen machen sich diese Idee zu Nutze, indem sie Korpus-Daten für die Induktion und Definition von lexikalischen Eigenschaften linguistischer Einheiten verwenden. Ein Vergleich der Korpus-basierten Eigenschaften von zwei linguistischen Einheiten bietet entsprechend die Möglichkeit, Aussagen zu dem Grad der Ähnlichkeit der linguistischen Einheiten zu machen.

Die Repräsentation einer linguistischen Einheit auf Basis distributioneller Merkmale nutzt typischerweise eine mathematische Formulierung im Vektorraum (siehe Turney & Pantel (2010) für eine detaillierte Beschreibung): Eine linguistische Einheit (in unserem Fall: ein Wort) wird als Vektor bzw. als Punkt im Vektorraum dargestellt. Ein fünf-dimensionaler Vektor wie zum Beispiel $\xrightarrow{\text{Maus}} = \langle 23, 116, 0, 0, 346 \rangle$ stellt die Bedeutung des Wortes „Maus“ dar. In diesem Fall bezieht sich der Vektor auf fünf Dimensionen, die je nach Forschungsziel unterschiedliche relevante Korpus-Merkmale quantifizieren. Typischerweise basieren die Werte der Merkmale auf Frequenzen, die durch Assoziationsmaße normalisiert werden können (Evert, 2005). A priori können die Wort-Vektoren nicht zwischen verschiedenen Wortbedeutungen unterscheiden, d.h. sie subsumieren die Merkmale aller Wortvorkommen im Korpus.

Die Merkmale, die als Vektor-Dimensionen ausgewählt werden, sollten möglichst salient sein in Bezug auf die Aufgabenstellung. Typische Kriterien für die Auswahl sind (a) Kookkurrenzen in einem bestimmten Wortfenster, dem Satz, Paragraphen oder Dokument; (b) dito, aber zusätzlich unter Einbezug bestimmter syntaktischer Abhängigkeiten (zum Beispiel könnten nur direkte Objekte von Verben betrachtet werden); (c) lineare Sequenzen von Kontext-Wörtern. Der Vergleich von zwei Wort-Vektoren soll Aufschluss geben über die (semantische) Ähnlichkeit der entsprechenden beiden Wörter.

Im Folgenden nutzen wir distributionelle Vektoren (Modelle) für die Repräsentation und den Vergleich von deutschen Partikelverben und ihren Basisverben. Die unterliegenden Korpus-Daten entstammen je nach Aufgabenstellung und zeitlicher Einordnung den Web-Korpora *deWaC* (Baroni et al., 2009), *SdeWaC* (Faaß & Eckart, 2013) oder Varianten von *DECOW* (Schäfer & Bildhauer, 2012). Für die morphologische und syntaktische Aufbereitung und Darstellung der Korpus-Daten haben wir *SMOR* (Faaß et al., 2010), *MarMoT* (Müller et al., 2013), den *MATE*-Parser (Bohnet, 2010) sowie die Subkategorisierungs-Datenbank (Scheible et al., 2013) genutzt.

3.2 Modellierung von Kompositionalität

Partikelverben sind zu unterschiedlichen Graden kompositionell in Bezug auf ihre Konstituenten: die Partikeln und die Basisverben. Eine traditionelle Aufgabe im Bereich der Computerlinguistik ist es, diesen Grad der Kompositionalität vorherzusagen, denn in der automatischen Textverarbeitung ist es wichtig zu wissen, inwieweit ein komplexer Ausdruck wörtlich oder nicht-wörtlich in Bezug auf seine Konstituenten zu interpretieren ist. In den letzten zwei Dekaden wurden –hauptsächlich fürs Englische und fürs Deutsche– Bewertungen gesammelt, um einen Einblick in die wahrgenommenen Grade der Kompositionalität zu bekommen, und auch um Goldstandards für die Evaluation der automatischen Vorhersagen zu erstellen (Baldwin et al., 2003; Bannard et al., 2003; Bott et al., 2016; McCarthy et al., 2003; Cook & Stevenson, 2006; Hartmann, 2008). Abgesehen von wenigen Ausnahmen (Baldwin et al., 2003; Bannard et al., 2003; Cook & Stevenson, 2006) beschränken sich die Kompositionalitätsbewertungen auf den Grad der Kompositionalität des Partikelverbs in Bezug auf das Basisverb. Der semantische Beitrag der typischerweise sehr ambigen Partikel wird dabei weitestgehend ignoriert.

Unseren Arbeiten liegen zwei Goldstandards für deutsche Partikelverben zugrunde: *GS-99* enthält 99 zufällig ausgewählte Partikelverben mit 11 verschiedenen präpositionalen Partikeln und ist ausbalanciert über 8 Frequenzbereiche (Hartmann et al., 2008). *GhoSt-PV* enthält eine Zufallsauswahl von 400 Partikelverben mit 9 verschiedenen präpositionalen Partikeln und ist ausbalanciert über drei Partikeltyp-spezifische Frequenzbereiche basierend auf den drei Korpora *DECOW12*, *SdeWaC* und *HGC* (Bott et al., 2016). Tabelle 2 enthält Beispiele der Kompositionalitätsbewertungen aus *Ghost-PV*. Neben der durchschnittlichen Kompositionalitätsbewertung auf einer Skala von 1 (sehr opak) bis 7 (sehr kompositionell), der Anzahl der Bewertungsteilnehmer (Teiln.) und der Standardabweichung (σ) enthält der Goldstandard Informationen zu der harmonischen Frequenz (Harm. Freq.) der Partikelverben, zu einer Kategorisierung in hoch-, mittel- und niedrig-frequente Partikel-spezifische Frequenzbereiche (H/M/N), und die Anzahl der Bedeutungen laut Duden (Amb.).

Verb	Komp.	Teiln.	σ	Harm. Freq.		Amb.
<i>abkratzen</i>	5,29	14	2,52	39,8	M	>3
<i>abwinken</i>	3,31	16	2,06	157,0	H	2
<i>anfressen</i>	5,47	15	1,41	35,0	H	3
<i>anleuchten</i>	5,95	20	1,50	6,4	N	1
<i>aufhorchen</i>	4,55	29	1,97	74,6	H	1
<i>ausreizen</i>	3,62	29	2,13	19,4	M	2
<i>nachkochen</i>	4,83	23	4,83	19,8	M	1
<i>nachtragen</i>	4,47	15	2,03	3,9	N	2
<i>zulegen</i>	3,86	14	2,07	4,0	N	>3

Table 2: Beispiele aus dem Kompositionalitäts-Goldstandard.

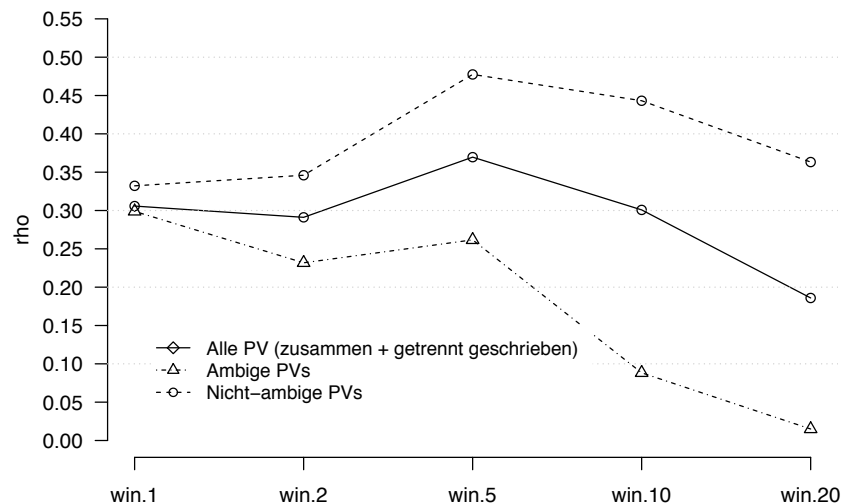
Im Folgenden vergleichen wir drei verschiedene Vektorräume bezüglich ihrer automatischen Vorhersage der Kompositionalitätsbewertungen. Dabei wurden sowohl für die Partikel- als auch für die Basisverben Vektoren erzeugt, und ihre Ähnlichkeit wurde durch den Kosinus der beiden Vektoren als Ähnlichkeitsmaß berechnet: Je kleiner der Winkel der beiden Vektoren im Raum ist, umso stärker wird der Grad der Kompositionalität vorhergesagt. Die Vorhersagen wurden anschließend mit den Goldstandard-Bewertungen anhand des Spearman-Rangkorrelationskoeffizienten ρ evaluiert (Siegel & Castellan, 1988).

Der erste Vektorraum ist ein Text-basiertes distributionelles Modell und beruht auf Kookkurrenzen zwischen Verben und Kontext-Wörtern in einem bestimmten Kontext-Fenster des SdeWaC (Bott & Schulte im Walde, 2014a). Die Größe des Kontext-Fensters wurde variiert: 1, 2, 5, 10, und 20 Wörter (jeweils links und rechts). Tabelle 3 zeigt die Ergebnisse für GS-99 über die verschiedenen Fenstergrößen hinweg. Außerdem wurde verglichen, welchen Unterschied es gibt, wenn nur die Partikelverben-Vorkommen betrachtet werden, die im Korpus zusammengeschrieben werden („zusammen“), im Vergleich dazu, dass auf Basis von syntaktischen Parses auch getrennt geschriebene Partikelverben betrachtet werden („zusammen+getrennt“). Die Tabelle zeigt, dass eine Fenstergröße von 5 Wörtern optimal ist für die Vorhersage, und dass das Modell deutlich besser ist, wenn auch getrennt geschriebene Partikelverb-Vorkommen beachtet werden.

	1	2	5	10	20
<i>zusammen</i>	0.2102	0.2507	0.2308	0.2416	0.2668
<i>zusammen+ getrennt</i>	0.3058	0.2910	0.3696	0.3008	0.1859

*Tabelle 3: Vorhersage von Kompositionalität
(zusammen- vs. getrenntgeschriebene Partikelverben).*

Abbildung 3 vergleicht die Vorhersagen für ambige vs. nicht-ambige Partikelverben in GS-99 und macht deutlich, dass Kompositionalitäts-Vorhersagen für nicht-ambige Partikelverben deutlich besser sind. Tabelle 4 vergleicht die Vorhersagen für unterschiedliche Frequenz-Bereiche von Partikelverben. Die Kompositionalitäts-Vorhersagen für niedrig-frequente und hoch-frequente Partikelverben sind schlechter als die für mittel-frequente Partikelverben. Dies ist nicht überraschend, da hoch-frequente Wörter typischerweise ambiger sind als mittel-frequente Wörter, und da niedrig-frequente Wörter typischerweise an spärlichen Datenbelegen im Korpus krankten.



*Abbildung 3: Vorhersage von Kompositionalität
(ambige vs. nicht-ambige Partikelverben).*

<i>Frequenz</i>	ρ
(2,5]	0,16
(5,10]	0,27
(10,18]	0,26
(18,55]	0,59
(55,110]	0,25
(110,300]	0,06
(300,6000]	0,13

Tabelle 4: Vorhersage von Kompositionalität über Frequenzbereiche der Partikelverben.

Der zweite Vektorraum ist ebenfalls ein Text-basiertes distributionelles Modell, bezieht aber syntaktische Abhängigkeiten mit ein (Bott & Schulte im Walde, 2015). Die Hypothese ist, dass Kompositionalitäts-Vorhersagen durch syntaktische Komplement-Vorauswahl verbessert werden können. Mit anderen Worten, ein distributionelles Modell, das nicht einfach alle (Inhalts-)Wörter der Kontexte von Partikel- und Basisverben miteinander vergleicht, sondern stattdessen deren direkte Objekte, bestimmte Präpositional-Objekte etc. sollte gezielter den Grad der Kompositionalität bestimmen können. Wie bereits erwähnt, kommen hier Verschiebungen der Argumentstruktur als empirisches Problem ins Spiel. Während zwei Basisverben problemlos auf Basis ihrer Komplement-Realisierungen hin verglichen werden können, müssen bei dem Vergleich von Partikel- und Basisverben Unterschiede in der Realisierung der Komplemente mit einbezogen werden.

Bott & Schulte im Walde (2014b) haben eine Methode vorgestellt, die für die wahrscheinlichsten Argumentrahmen von Partikel- und Basisverben sämtliche Komplement-Realisierungen miteinander vergleicht und dadurch die stärksten Korrespondenzen zwischen Komplement-Typen identifiziert. Abbildung 4 vergleicht die Modelle des ersten Fenster-Vektorraums (win.20) mit den Syntax-Modellen dieses Vektorraums. Die Schwellenwerte geben an, wie hoch eine Kosinus-Ähnlichkeit zwischen Komplement-Realisierungen mindestens sein muss, damit eine Korrespondenz zwischen den entsprechenden Komplement-Typen überhaupt

instanziiert wird: Je höher der Schwellenwert, umso weniger Korrespondenzen werden angenommen.

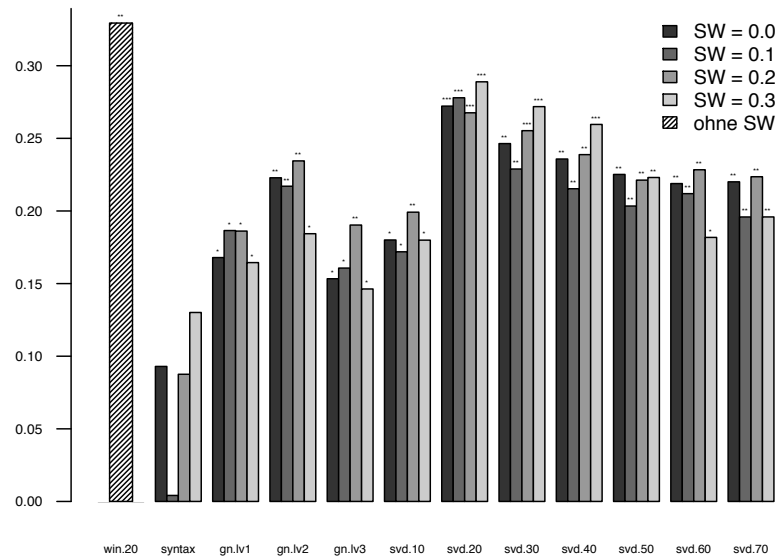


Abbildung 4: Vorhersage von Kompositionalität unter Einbezug syntaktischer Abhängigkeiten.

Grundsätzlich erreichen die Syntax-Modelle nur einen ρ -Wert von 0.13 (bei einem Schwellenwert von 0.3) und liegen somit weit unter den Durchschnittsergebnissen für die Fenster-Modelle (0.37, vgl. auch Tabelle 3). Durch Generalisierungen der Komplement-Realisierungen auf Basis der GermaNet-Hierarchie (Hamp & Feldweg, 1997; Kunze, 2000) bzw. nach Anwendung von Singular Value Decomposition (SVD) erreichen die Syntax-Modelle ρ -Werte von 0.23 (GermaNet Hierarchie-Level 2: gn.lv2, Schwellenwert: 0.2) bzw. 0.28 (SVD mit 20 Dimensionen: svd.20, Schwellenwert: 0.3). Die Syntax-Modelle sind somit signifikant schlechter als die Fenster-Modelle. Wir vermuten, dass die deutlich spärlicheren Daten der Grund für die schlechteren Modelle sind. Diese Vermutung wird dadurch erhärtet, dass die Generalisierungen durch GermaNet bzw. SVD, die die Spärlichkeit vermindern, zu besseren Vorhersagen führen.

Der dritte Vektorraum untersucht für Ghost-PV, inwieweit Text-basierte distributionelle Modelle durch Bildinformationen verbessert werden können (Köper & Schulte im Walde, 2017). Die Idee beruht auf der Basis von „*Grounded Cognition*“ (Barsalou, 2008), dass Kognition nicht nur auf sprachlichen Daten beruht, sondern Situationen und Perzeption mit einbezieht. Entsprechend haben wir Text- und Bild-Vektoren miteinander verknüpft. Die Bildinformation wurde wie bei Modellierungen vorhergehender Arbeiten (Kiela et al., 2014) entweder komplett für alle Partikel- und Basisverben hinzugefügt oder nur für eine gefilterte Auswahl. Wir haben drei verschiedene Filter angewandt: (i) Basierend auf der visuellen Vorstellbarkeit von Partikelverben (deren Einschätzung wir aus der semi-automatisch erstellten Datenbasis von Köper & Schulte im Walde (2016a) ablesen konnten) haben wir nur Bilder hinzugefügt für Partikelverben, deren Vorstellbarkeit einen bestimmten Schwellenwert erreicht hat. (ii) Entsprechend der paarweisen Dispersion der Partikelverb-Bilder haben wir nur die Bilder genommen für Partikelverben, die eine hohe durchschnittliche Ähnlichkeit aufwiesen. (iii) Wir haben die Bild-Vektoren geclustert und nur die Vektoren der Bilder genommen, die in dem größten Cluster waren.

Abbildung 5 zeigt die Ergebnisse der multi-modalen Modellierungen. Die Text-basierten Vorhersagen erreichen einen ρ -Wert von 0.22. Dieser Wert ist anders als bei den ursprünglichen Fenster-Modellen, da er auf „*Embeddings*“ beruht, distributionellen Vektoren auf Basis von Neuronalen Netzen. Das Hinzufügen von ungefilterter Bild-Information kann die Vorhersagen nicht verbessern. Die multi-modalen Modellierungen werden erst dann besser als die Text-Modelle, wenn nur die Bilder der visuellsten Partikelverben verwendet wurden bzw. wenn nur die Bilder verwendet wurden, die dem größten Cluster entstammen. Bei *Embeddings*, die bessere Text-basierte Vorhersagen machen (nicht Teil der Abbildung), hilft Bild-Information überhaupt nicht.

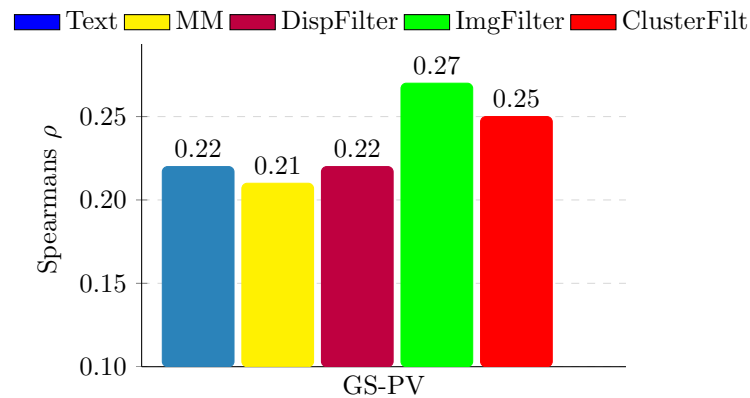


Abbildung 5: Vorhersage der Kompositionalität in einem multi-modalen Modell mit Embeddings.

Zusammenfassend können wir sagen, dass die Vorhersage des Grades von Partikelverb-Kompositionalität mit einem maximalen ρ -Wert von 0.37 (siehe Tabelle 3) einigermaßen erfolgreich ist, aber deutlich schlechter als zum Beispiel bei Nomen-Komposita mit 0.65 (Schulte im Walde et al., 2013; Köper & Schulte im Walde, 2017). Wir führen den Unterschied darauf zurück, dass erstens die Vorhersage der Kompositionalität auf den Bedeutungsbeitrag des Basisverbs reduziert wird (was aus theoretischer Sicht unzureichend ist), und dass zweitens Partikelverben aufgrund ihrer stärkeren Neigung zu Ambiguitäten, ihrer syntaktischen Idiosynkrasien und schlechterer Visualisierung eine größere Herausforderung darstellen als Nomen-Komposita.

3.3 Nicht-literale Bedeutungen von Partikelverben

Die Vorhersage des Grades der Kompositionalität wird offensichtlich der lexikalischen Ambiguität und syntaktischen Vielfalt von Partikelverben nicht gerecht. Im Folgenden präsentiere ich daher ein Klassifikationsexperiment, das sich auf der Token- statt auf der Typen-Ebene bewegt (Köper & Schulte im Walde, 2016b). Die Aufgabe des Experiments war es, die (Nicht-)Wörtlichkeit von Partikelverben im Kontext vorherzusagen. Beispiel (3) zeigt zwei Beispielsätze mit dem Verb „*abschminken*“.

In (3a) ist das Verb in wörtlichem Sprachgebrauch zu finden, in (3b) in nicht-literalem Sprachgebrauch.

- (3) (a) Den Lippenstift kannst du dir *abschminken*.
 (b) Den Job kannst du dir *abschminken*.

Für die automatische Klassifikation haben wir 159 Partikelverben über 10 präpositionale Partikeln hinweg aus einer Vorstudie (Springorum et al., 2013) ausgewählt. Für diese Partikelverben haben wir jeweils 50 Sätze zufällig aus dem DECOW14-Korpus gesucht und von jeweils drei deutschen Muttersprachlern auf einer Skala von 0 (sehr wörtlich) bis 5 (sehr nicht-literal) annotieren lassen. Die Übereinstimmung der Annotatoren lag bei 43%, Fleiss' $\kappa = 0.35$. Bei einer binären, disjunktiven Unterteilung in $[0,2]$ für wörtliche und $[3,5]$ für nicht-literale Sprachbedeutungen ergab sich eine Übereinstimmung von 79%, Fleiss' $\kappa = 0.70$. Aus der Gesamtmasse der annotierten Sätze haben wir diejenigen genommen, bei denen es eine hundertprozentige Übereinstimmung der drei Annotatoren bezüglich der binären Klassifikation gab, insgesamt 6436 Sätze (4174 literale und 2262 nicht-literale Sätze). Abbildung 6 zeigt, wie sich die Proportionen von literalen vs. nicht-literalen Sätzen pro Partikeltyp unterscheiden. Abbildung 7 zeigt dasselbe über alle Partikelverben hinweg. Die Abbildungen machen deutlich, dass es pro Verb und pro Partikeltyp unterschiedliche Präferenzen gibt.

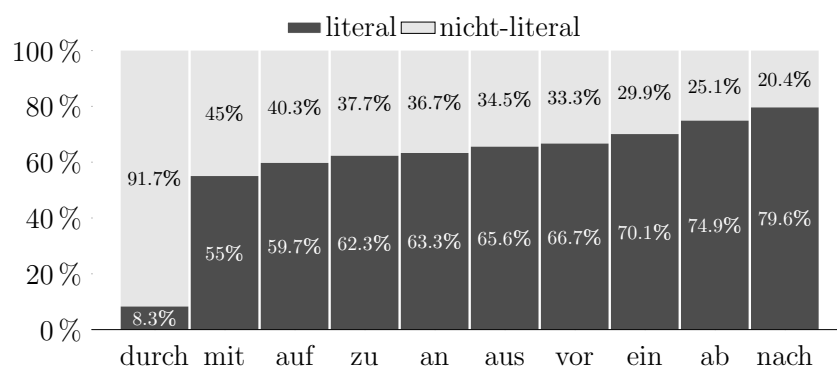


Abbildung 6: Literale vs. nicht-literale Sätze über Partikeltypen.

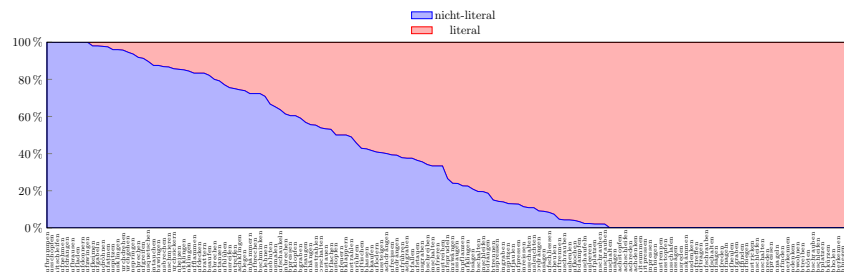


Abbildung 7: Literale vs. nicht-literale Sätze über Partikelverben.

Die automatische Klassifikation wurde anhand eines Random-Forest-Klassifikators durchgeführt. Ein Hauptinteresse lag auf der Identifikation von salienten Korpus-basierten Merkmale für die Unterscheidung von literalen und nicht-literalen Partikelverb-Sätzen. Dafür haben wir folgende Merkmale verglichen und kombiniert: (i) alle Inhaltswörter des entsprechenden Satzes („Unigramme“) und deren Generalisierung auf Basis von k-Means-Clustering; (ii) die Grade der Abstraktheit und der Vorstellbarkeit der Inhaltswörter, auf Basis der semi-automatisch erstellten Datenbasis von Köper & Schulte im Walde (2016a); (iii) die kontextuelle Passung von Basisverben in Partikelverb-Kontexten, auf Basis von distributioneller Ähnlichkeit; (iv) wahlweise die Information zu dem Partikel-Typen. Tabelle 5 zeigt die Ergebnisse der Klassifikation in Accuracy-Werten, mit und ohne Information zur Partikel. Die Tabelle zeigt deutlich, dass die Kontextwörter (vor allem in generalisierter Form), die Information zu ihrer Abstraktheit sowie die Passung von Basisverben in Partikelverb-Kontexten die salientesten Korpus-Merkmale sind bei einer Unterscheidung in literalen und nicht-literalen Sprachgebrauch. Die Vorgabe des Partikel-Typs ist ebenfalls relevant. Insgesamt treffen wir bei einer Kombination dieser Merkmale zu 86.8% richtige binäre Entscheidungen.

	<i>Accuracy (ohne P)</i>	<i>Accuracy (mit P)</i>
1 Baseline	64,9	-
2 Unigramme	75,6	76,5
3 Unigramme + Cluster	76,3	79,3
4 Abstraktheit	74,7	76,3

5 visuelle Vorstellbarkeit	68,6	71,6
6 PV/BV-Passung	76,5	80,2
Kombination 2+4+6	84,8	86,6
Kombination 3+4+6	85,0	86,8

Tabelle 5: Vorhersage von (nicht-)literalem Sprachgebrauch.

In einer qualitativen Analyse haben wir untersucht, inwieweit sich Partikelverben, die gegenseitig die Wörtlichkeit ihrer Kontexte vorhersagen können, in Bedeutungsaspekten ähneln. Dafür haben wir drei automatische Klassifikationen von Teilmengen von Partikelverben mit dem erfolgreichsten Klassifikator durchgeführt. In jeder dieser zusätzlichen Klassifikationen wurde der Klassifikator auf den Kontexten eines Partikelverbs trainiert und auf den Kontexten eines anderen Partikelverbs evaluiert. Die relevanten Paarungen von Partikelverben waren wie folgt: (i) „BVSyn“: Die Basisverben von zwei Partikelverben mit identischen Partikeln sind synonym (Quelle: Duden), zum Beispiel „aufreißen“ vs. „aufplatzen“. (ii) „PVSyn“: Die Partikelverben sind in mindestens einer Bedeutung synonym zueinander (Quelle: Duden), zum Beispiel „auftragen“ vs. „auftischen“. (iii) „SameBV“: Das Trainings- und das Test-Verb haben dasselbe Basisverb, zum Beispiel „abgraben“ vs. „aufgraben“.

Abbildung 8 stellt die Ergebnisse der drei zusätzlichen Klassifikationen dar, im Vergleich zum Gesamtexperiment. Die linken Boxen zeigen die Vorhersagen von literalen, die rechten Boxen die Vorhersagen von nicht-literalen Kontexten. In BVSyn verbessern sich die Vorhersagen sowohl für literale als auch für nicht-literale Kontexte. Wir interpretieren dieses Ergebnis als ein Indiz für reguläre Bedeutungsverschiebung, d.h. Basisverben, die aus einer gemeinsamen semantischen Klasse stammen und mit derselben Partikel kombiniert werden, resultieren in ähnliche literale und nicht-literale Bedeutungen der Partikelverben. Bei PVSyn verschlechtern sich die Vorhersagen für literale Kontexte, aber die Vorhersagen für nicht-literale Kontexte verbessern sich. Unabhängig von den Bedeutungen der Basisverben entsprechen sie offensichtlich den nicht-literalen Bedeutungen der Partikelverben, aber da sich die Basisverben in ihrer Bedeutung unterscheiden können (wie in unserem Beispiel bei „tragen“ und „tischen“), sind Vorhersagen für literale Kontexte ungenau. Bei

SameBV sehen wir sehr starke Vorhersagen für literale Kontexte, da literale Kontexte mit den identischen Bedeutungen der Basisverben zusammenhängen. Nicht-literale Kontexte werden sehr schlecht vorhergesagt, da die unterschiedlichen Partikeln unterschiedliche Bedeutungsbeiträge leisten.

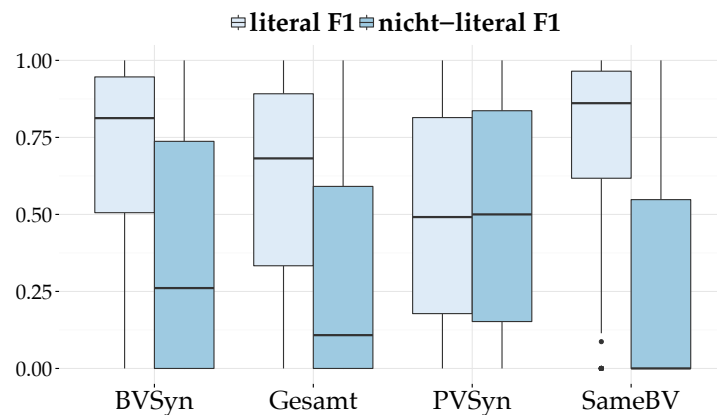


Abbildung 8: Vorhersage von (nicht-)literalem Sprachgebrauch bei Unter-Klassifikationen.

Zusammenfassend können wir sagen, dass die Vorhersage der (Nicht-)Wörtlichkeit mit einem Standard-Klassifikator und Partikelverb-salienten Merkmalen sehr erfolgreich ist: Wir können literale und nicht-literale Kontexte zu 86.8% richtig klassifizieren. Die salientesten Merkmale schließen Kontext-Wörter, die Abstraktheit des Kontexts sowie Information zur Partikel ein. Eine qualitative Studie zeigte zudem, dass die semantische Klasse der Basisverben, die Bedeutungen der Partikeln sowie regelmäßige Bedeutungsverschiebungen einen Beitrag zu (nicht-)literalem Sprachgebrauch bei Partikelverben leisten.

4. Zusammenfassung

Ich habe in diesem Artikel viele Puzzlesteine vorgestellt, die für die Komposition von deutschen Partikelverben relevant sind: (i) Unsere Datensammlung zur Generierung von Kontexten für Partikelverb-Neologismen hat gezeigt, dass es bei der Komposition von Partikelverben eine Systematik gibt bezüglich der Kombination von Partikelbedeutungen und semantischen Basisverb-Klassen. (ii) Ein Priming-Experiment konnte bezüglich der Partikeln *an* und *auf* spatiale Bedeutungsaspekte nachweisen. (iii) Distributionelle und multi-modale Modelle können empirische Bedeutungsaspekte für die Kompositionalität von Partikelverben aufdecken und den Grad der Kompositionalität modellieren. (iv) Distributionelle Klassifikationsmodelle geben Hinweise auf saliente Korpus-basierte und affektive Merkmale von literalen vs. nicht-literalen Kontexten von Partikelverben, und können die (Nicht-)Wörtlichkeit bestimmen.

Diese Puzzlesteine erzeugen offensichtlich noch kein vollständiges Bild von den (regulären) Mechanismen, die bei der Komposition von deutschen Partikelverben zusammenwirken, aber sie zeigen erste methodische Wege sowie kognitive und empirische Einsichten auf. Unsere zukünftige Arbeit zu Partikelverben wird die Modellierung auf der Token-Ebene vertiefen und mit Partikelverb-Kontexten arbeiten. Nach und nach werden wir hierbei Bedeutungsunterscheidungen bei Partikeln und Basisverbklassen integrieren. Ein mittelfristiges Ziel dabei sind Datensammlungen und Modellierungen von Analogien bei Bedeutungsverschiebungen.

5. Anerkennungen

Die Arbeiten in diesem Jahrbuch-Artikel sind in Zusammenarbeit mit meinen MitarbeiterInnen und KollaborationspartnerInnen Dr. Stefan Bott, Dr. Diego Frassinelli, Nana Khvtisavrivili, Maximilian Köper, Dr. Silke Scheible, Sylvia Springorum und Jason Utt entstanden.

6. Literaturverzeichnis

Timothy Baldwin, Colin Bannard, Takaaki Tanaka, & Dominic Widdows (2003). An Empirical Model of Multiword Expression Decomposability. In: *Proceedings of the ACL Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, pp. 89–96, Sapporo, Japan.

Colin Bannard, Timothy Baldwin, & Alex Lascarides (2003). A Statistical Approach to the Semantics of Verb-Particles. In: *Proceedings of the ACL Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, pp. 65–72, Sapporo, Japan.

Marco Baroni, Silvia Bernardini, Adriano Ferraresi & Eros Zanchetta (2009). The WaCky Wide Web: A Collection of Very Large Linguistically Processed Web-Crawled Corpora. *Language Resources and Evaluation* 43(3):209-226.

Lawrence W. Barsalou (2008). Grounded Cognition. *Annual Review of Psychology* 59:617-645.

Bernd Bohnet (2010). Top Accuracy and Fast Dependency Parsing is not a Contradiction. In: *Proceedings of the 23rd International Conference on Computational Linguistics*, pp. 89-97. Beijing, China.

Stefan Bott, Nana Khvtisavrishvili, Max Kisselew, & Sabine Schulte im Walde (2016). Ghost-PV: A Representative Gold Standard of German Particle Verbs. In: *Proceedings of the 5th Workshop on Cognitive Aspects of the Lexicon*, pp. 125-133. Osaka, Japan.

Stefan Bott & Sabine Schulte im Walde (2014a). Optimizing a Distributional Semantic Model for the Prediction of German Particle Verb Compositionality. In: *Proceedings of the 9th Conference on Language Resources and Evaluation*, pp. 509-516. Reykjavik, Iceland.

Stefan Bott & Sabine Schulte im Walde (2014b). Modelling Regular Subcategorization Changes in German Particle Verbs. In: *Proceedings of the 1st Workshop on Computational Approaches to Compound Analysis*, pp. 1-10. Dublin, Ireland.

Stefan Bott & Sabine Schulte im Walde (2015). Exploiting Fine-grained Syntactic Transfer Features to Predict the Compositionality of German Particle Verbs. In: *Proceedings of the 11th Conference on Computational Semantics*, pp. 34-39. London, UK.

Paul Cook & Suzanne Stevenson. 2006. Classifying Particle Semantics in English Verb-Particle Constructions. In *Proceedings of the Workshop*

on Multiword Expressions: Identifying and Exploiting Underlying Properties, pp. 45–53, Sydney, Australia.

Nicole Dehé (2015). *Particle Verbs in Germanic*. In: Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (Hrsg.), *Word Formation, An International Handbook of the Languages of Europe*, pp. 611-626. Berlin: De Gruyter.

Stefan Evert (2005). *The Statistics of Word Co-Occurrences: Word Pairs and Collocations*. Dissertation. Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Gertrud Faaß & Kerstin Eckart (2013). SdeWaC – A Corpus of Parsable Sentences from the Web. In: *Proceedings of the International Conference of the German Society for Computational Linguistics and Language Technology*, pp. 61-68. Darmstadt, Germany.

Gertrud Faaß, Ulrich Heid & Helmut Schmid (2010). Design and Application of a Gold Standard for Morphological Analysis: SMOR in Validation. In: *Proceedings of the 7th International Conference on Language Resources and Evaluation*, pp. 803-810. Valletta, Malta.

John R. Firth (1957). *Papers in Linguistics 1934-51*. Longmans, London, UK.

Diego Frassinelli, Alla Abrosimova, Sylvia Springorum & Sabine Schulte im Walde (2017). Meaning (Mis-)Match in the Directionality of German Particle Verbs. Poster at the *30th Annual CUNY Conference on Human Sentence Processing*, MIT, Cambridge, MA.

Birgit Hamp & Helmut Feldweg (1997). GermaNet – A Lexical-Semantic Net for German. In: *Proceedings of the ACL Workshop on Automatic Information Extraction and Building Lexical Semantic Resources for NLP Applications*, pp. 9-15. Madrid, Spain.

Zellig Harris (1954). *Distributional Structure*. *Word* 10(23):146-162.

Silvana Hartmann (2008). *Einfluss syntaktischer und semantischer Subkategorisierung auf die Kompositionalität von Partikelverben*. Studienarbeit. Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Douwe Kiela, Felix Hill, Anna Korhonen & Stephen Clark (2014). Improving Multi-Modal Representations Using Image Dispersion: Why Less is Sometimes More. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 835-841. Baltimore, MA.

Fritz Kliche (2011). Semantic Variants of German Particle Verbs with „ab“. *Leuvense Bijdragen*, 97:3–27.

Maximilian Köper & Sabine Schulte im Walde (2016a). Automatically Generated Norms of Abstractness, Arousal, Imageability and Valence for 350 000 German Lemmas. In: *Proceedings of the 10th Conference on Language Resources and Evaluation*, pp. 2595-2598. Portoroz, Slovenia.

Maximilian Köper & Sabine Schulte im Walde (2016b). Distinguishing Literal and Non-Literal Usage of German Particle Verbs. In: *Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 353-362. San Diego, CA.

Maximilian Köper, Sabine Schulte im Walde (2017). Complex Verbs are Different: Exploring the Visual Modality in Multi-Modal Models to Predict Compositionality. In: *Proceedings of the 13th Workshop on Multi-Word Expressions*, pp. 200-206. Valencia, Spain.

Claudia Kunze (2000). Extension and Use of GermaNet, a Lexical-Semantic Database. In: *Proceedings of the 2nd International Conference on Language Resources and Evaluation*, pp. 999-1002. Athens, Greece.

Andrea Lechler and Antje Roßdeutscher (2009). German Particle Verbs with „auf“. Reconstructing their Composition in a DRT-based Framework. *Linguistische Berichte* 220:439-478.

Beth Levin (1993). *English Verb Classes and Alternations*. The University of Chicago Press.

Diana McCarthy, Bill Keller & John Carroll (2003). Detecting a Continuum of Compositionality in Phrasal Verbs. In: *Proceedings of the ACL-SIGLEX Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, pp. 73-80. Sapporo, Japan.

Andrew McIntyre (2007). Particle Verbs and Argument Structure. In: *Language and Linguistics Compass* 1(4):350-397.

Thomas Müller, Helmut Schmid & Hinrich Schütze (2013). Efficient Higher-Order CRFs for Morphological Tagging. In: *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 322-332. Seattle, Washington, USA.

Roland Schäfer & Felix Bildhauer (2012). Building Large Corpora from the Web Using a New Efficient Tool Chain. In: *Proceedings of the 8th*

International Conference on Language Resources and Evaluation, pp. 486-493. Istanbul, Turkey.

Silke Scheible, Sabine Schulte im Walde, Marion Weller & Max Kisselew (2013). A Compact but Linguistically Detailed Database for German Verb Subcategorisation relying on Dependency Parses from a Web Corpus: Tool, Guidelines and Resource. In: *Proceedings of the 8th Web as Corpus Workshop*, pp. 63-72. Lancaster, UK.

Sabine Schulte im Walde, Stefan Müller & Stephen Roller (2013). Exploring Vector Space Models to Predict the Compositionality of German Noun-Noun Compounds. In: *Proceedings of the 2nd Joint Conference on Lexical and Computational Semantics*, pp. 255-265. Atlanta, GA.

Signey Siegel & N. John Castellan (1988). *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, Boston, MA.

Sylvia Springorum. 2011. DRT-based Analysis of the German Verb Particle „an“. *Leuvense Bijdragen* 97:80-105.

Sylvia Springorum, Jason Utt, Sabine Schulte im Walde (2013). Regular Meaning Shifts in German Particle Verbs: A Case Study. In: *Proceedings of the 10th International Conference on Computational Semantics*, pp. 228-239. Potsdam, Germany.

Springorum (2014). (Re)constructing German Verb Particle Meanings for Familiar and Novel Verbs. *Vortrag auf der "International Conference on the Cross-linguistic Comparison of Indo-Germanic and Semitic Languages"*, Konstanz, Germany.

Barbara Stiebels & Dieter Wunderlich (1994). Morphology feeds Syntax: The Case of Particle Verbs. *Linguistics* 32:913-968.

Peter D. Turney and Patrick Pantel (2010). From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research* 37:141-188.