

Tagungsbericht / Conference Report:

“Word formation and discourse structure” May 5–6, 2022; Leipzig, Germany (Maximilian Frankowsky)

Am 5. und 6. Mai 2022 fand am Institut für Germanistik der Universität Leipzig der von Adele Baltuttis, Anna Bliß, Barbara Schlücker und dem Autor dieses Berichts organisierte internationale Workshop “Word Formation and Discourse Structure” statt. Gegenstand des Workshops war ein bisher in der Forschung weitgehend unbeachtetes Thema, nämlich die Rolle der Wortbildung für die Struktur und Verständlichkeit von Texten: Inwieweit tragen komplexe Wörter zum Strukturaufbau und zur inhaltlichen Verknüpfung über Satzgrenzen hinweg bei? Vor dem Hintergrund dieses Desiderats sollte der Workshop ein Forum bieten, in dem zu diesem Thema Perspektiven interdisziplinärer Forschung erarbeitet werden. Zwar sind bereits in den 1970er bis 1990er Jahren erste Arbeiten zur Interaktion von Wortbildung und Text- bzw. Diskurslinguistik entstanden, im Anschluss ist das Thema jedoch kaum noch verfolgt worden. Dabei fanden in der Text-/Diskurslinguistik enorme Entwicklungen statt, etwa bei der Einbeziehung neuer Methoden, insbesondere aus der Psycho- und Computerlinguistik oder hinsichtlich elaborierter theoretischer Modellierungen. Der Aspekt der Wortbildung blieb hierbei aber weitgehend außen vor. Der Workshops sollte deshalb neue und bekannte Fragen und Probleme (wieder) aufgreifen und vor dem Hintergrund der aktuellen Erkenntnisse und Methoden neu diskutieren. Das Thema Wortbildung und Diskursstruktur ist insofern ein neues und interdisziplinäres Thema, das das Zusammenspiel unterschiedlicher linguistischer Bereiche und Traditionen erfordert. Dies zeigt sich auch im Programm des Workshops, das internationale Wissenschaftler:innen aus verschiedenen linguistischen Teilbereichen (Wortbildung, historische Sprachwissenschaft, Diskurs-/Textlinguistik, Fachsprachen, Korpus- und Computerlinguistik) zusammengebracht hat.

Das Programm beinhaltete elf Vorträge, die von 20 Beitragenden erarbeitet wurden. Den Auftakt machte **Manfred Stede** (Universität Potsdam), der als eingeladener Sprecher unter dem Titel “Anaphoric connectives: Discourse processing at its best!” einen Überblick zum Workshopthema gab. Er ging dabei exemplarisch auf die Identifikation von Koreferenz und die Bestimmung von Kohärenzrelationen ein und zeigte theoretische und methodische

Probleme der Schnittstelle Computerlinguistik/Diskurslinguistik auf. Anhand zweier Fallstudien zu morphologisch komplexen Verbindungswörtern wie *demzufolge* und *trotzdem* behandelte Stede etwa die Frage, unter welchen Voraussetzungen große Distanzen zwischen diesen Verbindungswörtern und den entsprechenden koreferierenden Argumenten akzeptiert werden. In der Diskussion wurden zudem die Aspekte Register und Transparenz von Wortbildungen wie *trotzdem* diskutiert.

Im Vortrag “Word-formation of light verb-constructions and its role in German texts. A corpus-based study” von **Martina Werner** und **Sophia M. Beiter** (Universität Wien) ging es um die Lexikalisierung / Desemantisierung von Funktionsverbgefügen im Deutschen (*zur Aufführung bringen*) sowie ihre Interaktion mit dem Kontext. Zwei auf dem Austrian Media Corpus basierende Korpusstudien beleuchteten hier die Rolle, die Funktionsverbgefügen und ihrer Modifikation für die Entstehung von Textkohärenz und -kohäsion zukommen. Die Ergebnisse der Studien sprechen dafür, dass die Mehrheit der Funktionsverbgefüge Modifizierer enthält, die Möglichkeit zur Modifikation des nominalen Bestandteils (*zur erfolgreichen Aufführung bringen*) aber mit zunehmender Lexikalisierung der Funktionsverbgefüge abnimmt. In der Diskussion wurde zudem der Einfluss der Negation auf die Interpretation von Funktionsverbgefügen eruiert (*nicht zur Aufführung kommen*). Der Vortrag zeigte anschaulich, wie Kohäsion und Kohärenz auch durch die morphologische Struktur der in den Funktionsverbgefügen enthaltenen Elemente beeinflusst werden und bot somit Evidenz für die These des Workshops, dass Wortbildung mit Textstruktur interagiert.

Zi Huang (Universität Pompeu Fabra, Barcelona) behandelte in ihrem Vortrag “The discourse status of verbal gerunds” zwei Arten des Gerundiums im Englischen (POSS-*ing*: *Clay’s winning the game* und ACC-*ing*: *Clay winning the game*). Häufig wird angenommen, dass diese zwei Arten des Gerundiums sich darin unterscheiden, dass der Referent des nominalen Teils (*the game*) nur beim POSS-*ing* im Kontext gegeben sein muss. Mit einer Korpusstudie auf der Grundlage des British National Corpus liefert Zi Huang Evidenz dafür, dass dieser Unterschied zwischen POSS-*ing* und ACC-*ing*-Gerundien eher eine Tendenz als eine Regel ist und die deverbale Nominalisierung als solche zu textkonstituierenden Unterschieden führt.

Im Vortrag von **Fabian Fleißner**, **Regina Ruf** und **Elena Smirnova** (Université de Neuchâtel) ging es um deverbale Nominalisierungen (*Beendigung*, *Bezug*, *Erscheinen*), ihre Rolle bei der Verdichtung von Informationen und die Frage, warum dieser im Deutschen heute sehr frequente Prozess erst im 19. Jahrhundert zur Informationsverdichtung verwendet wurde. Auf der Grundlage von Daten aus dem Deutschen Textarchiv gingen die Autor:innen der Hypothese nach, dass deverbale Nominalisierungen in deutschen Texten

neben der Informationsverdichtung auch der Herstellung von Referenz und der textinternen Vernetzung dienen und sie diese Textfunktionen mithilfe von Konstruktionen ausüben (*in Bezug auf*). Auch die Erkenntnisse aus diesem Vortrag zeigen also, dass es einen Zusammenhang von diskurspragmatischen Mustern und morphosyntaktischen Regularitäten gibt.

Mauro Le Donne (Università per Stranieri di Perugia) sprach in seinem Vortrag "(Con)textaholic: a corpus-based study on Italian blends among textuality and word formation" über Wortkreuzungen im Italienischen (*fantastiliardo* < *fantasti*(co) + (m)*iliardo*). Er präsentierte eine Korpusstudie, in der mithilfe von Sketch Engine sechs italienische Korpora abgefragt wurden, und illustrierte an zehn Beispielen, dass solche Wortkreuzungen meist auf eine Dekodierungshilfe im Kontext angewiesen sind und sich zudem im Text anders verhalten als Komposita. So können Wortkreuzungen wie *fantastiliardo* gleichzeitig auf mehrere Konzepte im selben Ko-Text referieren und zeigen darüber hinaus eine inkonsistente Genus- und Numerusmarkierung.

Um die Produktivität und die pragmatische Funktion von Eigennamenkomposita mit einem Eigennamen als Zweitglied (*Helikopter-Cem*) im Deutschen ging es im Vortrag "Personal name compounding in German – On the relationship between productivity and pragmatic functions across discourse domains and text types" von **Milena Belosevic** und **Sabine Arndt-Lappe** (Universität Trier). Die Autorinnen stellten eine Korpusstudie vor, die diesen bisher nur wenig beachteten Kompositumstyp mit Daten aus dem DeReKo, dem DWDS sowie von Twitter untersucht, und vertraten die These, dass die Produktivität dieses Wortbildungstyps mit externen Faktoren, etwa der Textsorte, zusammenhängt, sowie mit internen Faktoren, etwa der Semantik des Modifizierers. Auch präsentierten Belosevic und Arndt-Lappe einen Vorschlag zur Unterscheidung verschiedener Lesarten solcher Eigennamenkomposita (eventiv, partitiv, qualifizierend) und ordnen ihnen unterschiedliche Funktionen zu. Demnach referieren Eigennamenkomposita mit partitiver Lesart tatsächlich auf den Namensträger, während die der eventiven und qualifizierenden Lesart eher zur Bewertung des Namensträgers verwendet werden. Die Autorinnen diskutierten zudem die These, dass die Auftretenshäufigkeit der drei Lesarten mit der Textsorte, in der die Eigennamenkomposita auftreten, zusammenhängt.

Zu Beginn des zweiten Workshoptags standen methodische Fragestellungen im Mittelpunkt. In seinem Vortrag "A Proposal for Explicit Word Formation Annotation in Discourse Corpora" bot **Mark-Christoph Müller** (Heidelberg Institute for Theoretical Studies) Antworten auf die Frage, wie man komplexe Wörter und ihren Beitrag zur Herstellung von Textstruktur und -kohärenz korpusbasiert besser analysieren kann. Müller präsentierte mit MMAX2 ein Tool, das eine flexible Textannotation auf verschiedenen sprachlichen

Ebenen ermöglicht und mit dem Texte sowohl manuell als auch automatisch annotiert werden können.

Im anschließenden Vortrag exemplifizierten **Wei Zhao, Kevin Mathews** und **Haixia Chai** (Heidelberg Institute for Theoretical Studies) am Beispiel der Auflösung von Koreferenz, wie die Berücksichtigung von Wortbildungsprozessen zu einer Verbesserung automatischer Sprachverarbeitung führen kann. In ihrem Ansatz werden Nominalphrasen aus dem Koreferenz-Datenset CoNLL (Pradhan et al. 2012) mit Wortbildungsmustern verknüpft und die Koreferenz im Anschluss mithilfe des Systems von Joshi et al. 2020 aufgelöst. Die Autor:innen messen auf diese Weise den Beitrag, den die Wortbildungsmuster an dem Ergebnis haben, und schlagen infolgedessen eine Methode vor, mit der das Koreferenzauflösungssystem mithilfe des kontextualisierten Sprachmodells BERT (Devlin et al. 2019) verbessert werden kann.

Wie bedeutungslose Affixe als Wortbildungseinheiten zur Herstellung von Textkohärenz beitragen können, erörterte **Adele Baltuttis** (Universität Leipzig) in ihrem Vortrag “Affixes in Discourse”. Ausgangspunkt war hier die Beobachtung von Lim (2004), dass Wortbildungselemente wie etwa das Affix *zer-* Topikrelationen bilden können. Baltuttis präsentierte erste Ergebnisse aus ihrem laufenden Dissertationsprojekt, die die vorherrschende Annahme relativieren, dass wiederkehrende Elemente in Texten nur auf der Grundlage einer semantischen Verbindung eine Kohärenzfunktion ausüben können und bedeutungslose Affixe in Texten höchstens als Kohäsionsmarker fungieren.

Im Anschluss diskutierte **Katrin Menzel** (Universität des Saarlandes) in ihrem Vortrag “Word formation and cohesion in specialized languages” die diskursrelevante Bedeutung von Wortbildung und beleuchtete das Zusammenspiel von Wortbildungsmustern und Kohäsion aus diachroner, synchroner und sprachvergleichender Perspektive. In der Analyse der Autorin ist dieses Zusammenspiel wichtig für die Übersetzung von fachsprachlichen Texten aus dem Englischen ins Deutsche und andersherum, da Wörter mit gleicher oder ähnlicher morphologischer Struktur eine semantische oder assoziative Verbindung zu anderen Wörtern im Text signalisieren. Das Wissen um Wortbildungsprozesse ist darum relevant für die Textkompetenz der Sprachteilnehmer:innen, insbesondere für Übersetzer:innen fachsprachlicher Texte.

Der Vortrag von **Julia Lukassek, Anke Lüdeling, Anna Shadrova** und **Shujun Wan** (HU Berlin) bildete den Abschluss des Workshops. Sie gingen der Frage nach, wie fortgeschrittene Lerner:innen des Deutschen den Nominalstil erwerben und wie sie den adäquaten Gebrauch von Wortbildungsmustern in diesem Register meistern. Die Autorinnen präsentierten in ihrem Vortrag “Complex nouns as markers of academic register in L1-and L2-authored

essays” eine auf dem Falko Essay corpus (Reznicek et al. 2012) basierende Studie, in der akademische Aufsätze von DaF-Lerner:innen mit Aufsätzen von Muttersprachler:innen des Deutschen in Bezug auf die verwendeten Wortbildungsmuster verglichen werden. Ein Ergebnis der Studie ist, dass komplexe Nomina in den Aufsätzen der beiden Gruppen ähnlich häufig verwendet werden, die Varianz in beiden Vergleichsdatensets aber sehr hoch ist. Hiervon lässt sich die weiterführende Frage ableiten, ob die Verteilung über komplexe Nominalklassen tatsächlich eine Frage des Registers ist oder doch eher eine Frage der subjektiven Präferenz.

Auch wenn die Veranstalter:innen des Workshops die Teilnehmer:innen mit einigen noch nicht gelösten Fragen entließen, machte der Workshop deutlich, dass der Wortbildung eine große Bedeutung für die Text-/Diskursstruktur zukommt. Während des Workshops wurden anhand vieler verschiedener Sprachphänomene wesentliche Problemstellen aufgezeigt und die Anwendung von Methoden und Modellierungen diskutiert. Die Vorträge und der Austausch bei diesem Zusammentreffen können somit als eine Grundlage verstanden werden, auf der künftig eine verbindende Betrachtung der Domänen ‚Wortbildung‘ und ‚Textstruktur‘ aufbauen kann. Der Workshop warf Fragen auf, die an den Veranstaltungstagen zwar nicht vollständig beantwortet werden konnten, sich im Verlauf des Workshops aber als essenziell herausgestellt haben: Welchen Mehrwert liefert die Berücksichtigung von Wortbildungsprozessen für die adäquate Beschreibung von Textstruktur? Welche empirischen Methoden erlauben eine angemessene Auseinandersetzung mit der Interdependenz von Wortbildung und Text? Wie kann die in den Vorträgen bereits skizzierte Bandbreite an Wortbildungsphänomenen mit der schwierig zu erfassenden Konstitution von Textkohärenz und -kohäsion korpusbasiert zusammengebracht werden? Wie kann das Zusammenspiel von Wortbildung und Text für die Verbesserung automatischer Sprachverarbeitung fruchtbar gemacht werden? Diese Fragen können nun im Rahmen weiterführender Studien angegangen werden.

Das detaillierte Programm des Workshops samt Vortragstiteln und Abstracts ist unter folgendem Link abrufbar: <https://home.uni-leipzig.de/grammatik-des-deutschen/WFDS/>

Literatur

Devlin, Jacob, Ming-Wei Chang, Kenton Lee & Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. Minneapolis, Minnesota.

- Joshi, Mandar, Omer Levy, Daniel S. Weld & Luke Zettlemoyer. 2019. BERT for coreference resolution: Baselines and analysis. In *Empirical Methods in Natural Language Processing (EMNLP)*, 5803–5808. Hong Kong, China.
- Lim, Seong Woo. 2004. *Kohäsion und Kohärenz: Eine Untersuchung zur Textsyntax am Beispiel schriftlicher und mündlicher Texte*. Würzburg: Universität Würzburg Dissertationsschrift.
- Pradhan, Sameer, Alessandro Moschitti, Nianwen Xue, Olga Uryupina & Yuchen Zhang. 2012. CoNLL-2012 shared task: Modeling multilingual unrestricted coreference in OntoNotes. In *Joint Conference on EMNLP and CoNLL-Shared Task*, 1–40. Jeju Island, Korea.
- Reznicek, Marc, Anke Lüdeling, Cedric Krummes, Franziska Schwantuschke, Maik Walter, Karin Schmidt, Hagen Hirschmann & Torsten Andreas. 2012. *Das falko-Handbuch. Korpusaufbau und Annotationen*. Version 2.01. Tech. rep. Berlin: Humboldt-Universität.

Maximilian Frankowsky
Institut für Germanistik
Beethovenstraße 15
D – 04107 Leipzig
maximilian.frankowsky@uni-leipzig.de