

## Semantische Netzwerke als Hilfsmittel für die Untersuchung des Begriffskomplexes *Arbeit* im öffentlichen Diskurs Österreichs

Klaus Hofmann, Andreas Baumann (Universität Wien); Anna Marakasova, Julia Neidhardt (TU Wien); Tanja Wissik (ÖAW)

Dieser Beitrag nähert sich dem Thema ‚Wert der Arbeit‘ aus sprachwissenschaftlicher Perspektive an. Mittels *Machine-Learning*-basierter semantischer Netzwerke wird untersucht, welche Konzepte im österreichischen öffentlichen Diskurs mit *Arbeit* verknüpft sind und inwieweit diese sprachlichen Nahebeziehungen Rückschlüsse auf Wertzuschreibungen zulassen, die mit der Verwendung von Ausdrücken für verschiedene Tätigkeitsbereiche (z.B.: *Erwerbsarbeit*, *Hausarbeit*, etc.) einher gehen.

Wie über Dinge gesprochen wird, hat einen maßgeblichen Einfluss darauf, wie sie wahrgenommen werden, welche Wertzuschreibungen sie erfahren und welche Weltbilder durch sie transportiert werden. Dies gilt spätestens seit dem Einzug der Diskursanalyse in den wissenschaftlichen Methodenapparat als eine der wichtigsten Grundlagen für die kritische Auseinandersetzung mit sozialen, politischen und kulturellen Themen.<sup>1</sup> Ausgehend von der kognitiven Sprachwissenschaftsforschung ist in den letzten Jahrzehnten besonders die subtile Wirkkraft umgangssprachlicher Metaphorik und des sogenannten ‚Framings‘, also der sprachlichen (anstatt der rein thematischen) Kontextualisierung, in den Mittelpunkt des Interesses getreten.<sup>2</sup> Im politischen und medienöffentlichen Diskurs haben diese Zugänge ein reichhaltiges Quellenmaterial gefunden.

Im vorliegenden Beitrag sollen Ressourcen und Anwendungen aus der Computerlinguistik vorgestellt werden, die die meist qualitative Arbeit in diesem Bereich um eine quantitativ/datenbasierte Dimension erweitern und dadurch bereichern sollen. Konkret werden der Begriff *Arbeit* und damit verwandte Ausdrücke in Form von semantischen Netzwerken dargestellt. Diese werden mithilfe bewährter *Machine-Learning* Algorithmen<sup>3</sup> und umfangreicher Textdatenbanken (*Korpora*) erstellt und eignen sich sehr gut für die graphische Darstellung komplexer Zusammenhänge (siehe Abbildungen unten). Die Knoten (*nodes*) und Verbindungslinien (*edges*) in den Netzwerken repräsentieren die semantischen Affinitäten zwischen Begriffen, wie sie sich aus den Verteilungswahrscheinlichkeiten im Textmaterial ergeben: Wörter, die außergewöhnlich oft gemeinsam in Texten auftreten, werden durch dichte Verknüpfungsverbindungen dargestellt und formen so Bedeutungs-*Cluster*. Dadurch wird sichtbar gemacht, welche Konzeptmuster häufig mit den gesuchten Ausdrücken in Verbindung gebracht werden und dadurch potenziell für die Wahrnehmung der Begriffe relevant sein können.

Das Textmaterial, das für die Erstellung der Netzwerke herangezogen wird, ist zwei großen Textkorpora entnommen, welche beide an der Österreichischen Akademie der Wissenschaften (ÖAW) beheimatet sind: das *Austrian Media Corpus* (AMC) beinhaltet österreichische Zeitungen, Magazine und Fernsehbeiträge aus den vergangenen Jahrzehnten;<sup>4</sup> *ParLAT* umfasst die Transkripte aller

---

<sup>1</sup> Vgl. Foucault, Michel: *Archäologie des Wissens*, 17. Auflage. Suhrkamp: Frankfurt am Main 2015; Fairclough, Norman: *Critical Discourse Analysis. The critical study of language*. Longman: London 1995; Wodak, Ruth: *Politics of fear. Analyzing right-wing popular discourse*. London: SAGE 2015.

<sup>2</sup> Vgl. Lakoff, George; Johnson, Mark: *Metaphors we live by*. Chicago: University of Chicago Press 1980; Wehling, Elisabeth: *Politisches Framing. Wie eine Nation sich ihr Denken einredet – und daraus Politik macht*. Köln: Herbert von Halem Verlag 2016.

<sup>3</sup> Mikolov, Thomas; Sutskever, Ilya; Chen, Kai; Corrado, Greg; Dean, Jeffrey: *Distributed representations of words and phrases and their compositionality*. In: *Advances in Neural Information Processing Systems* (2013), S. 3111-3119 (arXiv:1310.4546v1); Wang, Shirui; Zhou, Wenan; Jiang, Chao: *A survey of word embeddings based on deep learning*. In: *Computing* 2020 (102), S. 717-740.

<sup>4</sup> Ransmayr, Jutta; Karlheinz Mörth; Ďurčo, Matej: *AMC (Austrian Media Corpus)*. In: C. Resch und W. U. Dressler (Hrsg.): *Korpusbasierte Forschungen zum österreichischen Deutsch*. In *Digitale Methoden der Korpusforschung in Österreich*. Verlag der Österreichischen Akademie der Wissenschaften: Wien 2017, S. 27-38; Siehe <https://amc.acdh.oeaw.ac.at/about-amc/>.

Redebeiträge im österreichischen Parlament seit 1996.<sup>5</sup> Beide Korpora wurden lemmatisiert, nach Wortart markiert (*Part-of-Speech Tagging*) und folgen dem internationalen Dokumentenformat der *Text Encoding Initiative* (TEI).

Beispielsweise lässt sich das diskursive Umfeld des Begriffs *Arbeit* in Tageszeitungen des Jahres 2017 (Abb. 1) in mehrere Bedeutungscluster unterteilen. Hervorzuheben sind hier einerseits der Bedeutungskomplex *Knochenjob*, *Verdienen*, der eine Nahebeziehung zwischen den Konzepten Arbeit, (körperliche) Anstrengung und finanziellem Verdienst suggeriert, sowie die dichte Verknüpfung im Bereich *Freiwilligenarbeit*, zu dem auch *Flüchtlingsarbeit* gehört. Andererseits ist interessant, dass die Begriffe *Angestelltenverhältnis* und *Erwerbsarbeit* im engen Verbund mit *Zweitjob*, *-Stunden-Job* (für diverse Teilzeitarbeitsverhältnisse) und *Saisonarbeit* auftreten, und dass dieser Cluster sich wiederum in abgegrenzter aber dennoch relativ enger Beziehung mit dem Komplex *Pflegearbeit*, *Erziehungsarbeit*, *Landschaftspflege* befindet. In dem Netzwerk für *Arbeit* in Parlamentsreden (Abb. 2) sticht besonders der Cluster *sinnstiftend*, *Selbstwert(gefühl)*, *Bestreiten* ins Auge, allesamt Begriffe, die den psychologischen/sozialen Wert von Arbeit für das Individuum jenseits rein finanzieller Gesichtspunkte ansprechen.

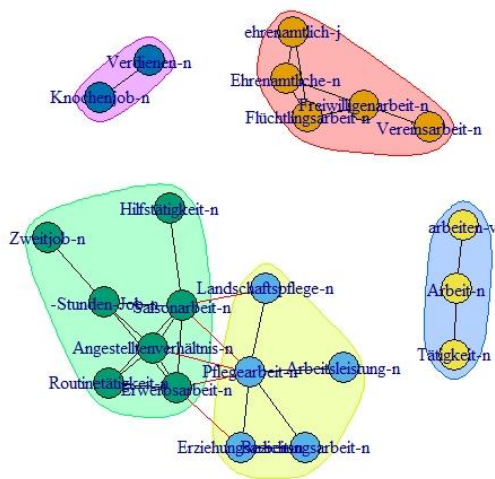


Abb. 1: Semantisches Netzwerk für *Arbeit* in österreichischen Tageszeitungen 2017.

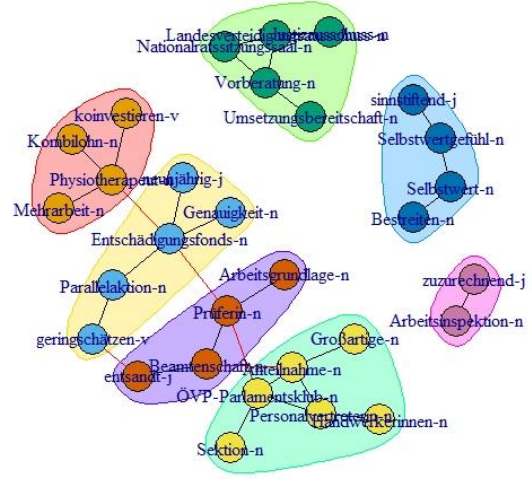


Abb. 2: Semantisches Netzwerk für *Arbeit* in österreichischen Parlamentsreden 2017.

Das Ziel des hier vorgestellten Ansatzes ist es, die unüberschaubare Textfülle für qualitative Untersuchungen zugänglich zu machen, um so auf sprachliche Gebrauchsmuster aufmerksam zu werden, die ansonsten in der schier Masse an Textmaterial verborgen blieben. Zusätzlich bietet die umfangreiche Annotation der Korpora (nach Medium, Publikationsdatum, Autor\*in, etc.) die Möglichkeit, verschiedene Segmente des Medien- und Parlamentsdiskurses getrennt zu analysieren. Einerseits lassen sich so zum Beispiel Unterschiede im Sprachgebrauch verschiedener Medienpublikationen betrachten oder verschiedener im Parlament vertretener Fraktionen. Andererseits kann damit abgebildet werden, wie sich der Diskurs um einen Begriff herum über Jahre und Jahrzehnte hinweg entwickelt hat, indem die semantischen Netzwerke verschiedener Jahre miteinander verglichen werden. Wir hoffen, damit einen nutzbringenden Beitrag zur Diskussion über den Stellenwert von Arbeit in der österreichischen Gesellschaft leisten zu können.

<sup>5</sup> Wissik, Tanja; Pirker, Hannes: ParlAT beta Corpus of Austrian Parliamentary Records. In: Proceedings of the LREC 2018 Workshop ParlaCLARIN: LREC2018 workshop on creating and using parliamentary corpora, 2018, S. 20-23.