



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2022

Exploration semantischer Räume im Corona-Diskurs

Bubenhofer, Noah

DOI: <https://doi.org/10.1515/9783110774306-013>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-217622>

Book Section

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:

Bubenhofer, Noah (2022). Exploration semantischer Räume im Corona-Diskurs. In: Kämper, Heidrun; Plewnia, Albrecht. Sprache in Politik und Gesellschaft : Perspektiven und Zugänge. Berlin, Boston: De Gruyter, 197-216.

DOI: <https://doi.org/10.1515/9783110774306-013>

Noah Bubenhofer (Zürich)

Exploration semantischer Räume im Corona-Diskurs

Abstract: Im Corona-Diskurs prallen völlig unterschiedliche Meinungen und Positionen zur Rolle des Staates aufeinander. Die Studie untersucht diese Positionen mit korpuslinguistischen Methoden anhand der Berichterstattung von Medien und Kommentaren von Leserinnen und Lesern in der Deutschschweiz. Dabei werden auch rechte und Corona-skeptische Plattformen in die Analyse einbezogen. Grundlage des korpuspragmatischen Zugangs ist die Berechnung und Interpretation von Word Embeddings, einer Methode zur Modellierung von semantischen Räumen. Es zeigt sich, wie sich im Diskurs inkommensurable Semantiken entwickeln.

1 Einleitung

Diskurse prägen semantische Räume: Während der Corona-Pandemie werden die Maßnahmen der Regierungen unterschiedlich bewertet. Die einen bezeichnen das Verhalten der Regierung als fürsorgerisch und an Werten wie Solidarität und Zusammenhalt orientiert, andere sprechen von Eigenverantwortung, nach der sich die Bürgerinnen und Bürger richten sollen und Dritte wiederum bezeichnen den Staat als Diktatur oder totalitären Polizeistaat. Diese unterschiedlichen Bezeichnungen stehen für verschiedene diskursiv geprägte semantische Räume.

In Europa zeichnet sich das während der Corona-Pandemie gegenüber den epidemiologischen Maßnahmen kritische Lager durch Heterogenität aus. Positionen der Neuen Rechten vermischen sich mit jenen von Libertären und Impfgegnern und Impfgegnerinnen, die wiederum selber höchst heterogen zusammengesetzt sind. Deshalb eignen sich Corona-Diskurse gut, um völlig unterschiedliche Positionen – und damit Semantisierungen – beim gleichen Großthema analysieren zu können.

Im Folgenden soll gezeigt werden, wie mit datengeleiteten, korpuslinguistischen Methoden semantische Räume exploriert werden können. Dafür werden Methoden der distributionellen Semantik verwendet, sogenannte Word Embeddings nach dem word2vec-Verfahren (Mikolov et al. 2013; Lenci 2018). Als Basis

für die Analysen wird ein Korpus von Texten aus Orientierungsmedien, dazugehörigen Kommentaren von Lesern und Leserinnen, sowie von Online-Medien und Blogs sogenannter ‚alternativer Medien‘ verwendet, die die Corona-Pandemie thematisieren. Der Fokus liegt auf dem Deutschschweizer Corona-Diskurs, was sich in den Daten widerspiegelt. Da allerdings vor allem auch die maßnahmenkritischen ‚Skeptiker und Skeptikerinnen‘ international vernetzt sind und sich generell über deutschsprachige Medien informieren, werden auch Quellen aus Deutschland berücksichtigt.

Methodologisch soll untersucht werden, ob die Berechnung und Analyse von Word Embeddings dazu taugt, differente Semantisierungen gleicher Ausdrücke zu untersuchen. Während das über klassische Kollokationsanalysen gut möglich und bereits vielfach gezeigt worden ist (vgl. z. B. Belica/Perkuhn 2015; Evert 2009), böte sich mit den Word Embeddings die Möglichkeit, semantische Räume, bestehend aus einer Vielzahl von Ausdrücken, systematisch zu vergleichen.

2 Datengrundlage

Wie eingangs erwähnt, sollen die Daten möglichst heterogene Positionen im Corona-Diskurs mit Fokus auf der Deutschschweiz repräsentieren. Das Korpus besteht deshalb aus den folgenden Quellen:

1. **CH-Medien** – Schweizer Massenmedien (Orientierungsmedien): Neue Zürcher Zeitung NZZ (liberal-konservativ), Tages-Anzeiger (links-liberal), Blick (Boulevard), AZ-Medien (Medienverbund zahlreicher regionaler Zeitungen), 20 Minuten (Online- und Print-Gratiszeitung) – die Titel gehören zu den meistgelesenen Tageszeitungen der Deutschschweiz.
2. **Kommentare** – ‚User Generated Content‘ bei Online-Zeitungen: Kommentare zu Artikeln zu Corona auf den Online-Newsplattformen Blick Online (blick.ch), Nau.ch, Watson.ch, 20minuten.ch und des öffentlich-rechtlichen Radios und Fernsehens SRF (srf.ch). Blick Online und 20minuten.ch sind teilweise deckungsgleich mit ihren gedruckten Ausgaben, wobei hier nur die Kommentare zu den Artikeln Eingang in das Korpus gefunden haben.
3. **Skep-Medien** – Medien der Corona-Skeptiker und Skeptikerinnen und Neuen Rechten: Compact Online, KenFM.de, PI-News, Rubikon, Inside Corona, corona-transition.org, Zeitpunkt – die ersten vier gehören zu bekannten und einschlägigen Medien der Neuen Rechten, letztere beide sind vom Schweizer Journalisten und bekennenden Maßnahmenkritiker Christoph Pfluger heraus-

gegebene Organe.¹ ‚Inside Corona‘ ist eine Plattform von Stephan Rietiker, der sich während der Pandemie ebenfalls als Maßnahmenkritiker profiliert hat.²

Die Datensammlung umfasst für die Teilkorpora CH-Medien und Skep-Medien den Zeitraum 1. Januar 2020 bis 31. Mai 2021. Das Kommentare-Teilkorpus enthält Daten ebenfalls vom 1. Januar 2020, jedoch nur bis 31. Januar 2021.

Als Suchausdrücke für die Artikel wurden folgende Zeichenketten verwendet (jeweils sowohl mit Groß- als auch Kleinbuchstaben): corona, covid, virus, pandemie, epidemie. Bei den Kommentaren mussten die Zeichenketten ‚corona‘, ‚covid‘ oder ‚virus‘ im Ausgangsartikel vorkommen.

Das Kommentare-Korpus wurde bereits in zwei Studien ausgewertet (Bubenhofner 2021; Bubenhofner et al. 2020) und für die vorliegende Studie mit den anderen Teilkorpora zu einem neuen ‚Corona-Korpus‘ vereinigt.

Tabelle 1 zeigt die genauen Angaben zu den Teilkorpora.

Tab. 1: Korpuszusammensetzung

Korpus	Anzahl Texte	Anzahl Tokens
CH-Medien (Corona)	43.147	30.902.449
AZ-Medien	14.567	
Blick	6.337	
NZZ	10.395	
Tages-Anzeiger	9.338	
20 Minuten	2.510	
Kommentare	840.926	30.251.840
20 Minuten	223.437	
Blick	226.365	
Nau	140.190	
Watson	168.967	
SRF	81.967	

¹ Im Impressum wird das Ziel der Plattform u. a. beschrieben mit: „Corona Transition macht auf öffentlich zugängliche, aber von Entscheidungsträgern kaum beachtete oder ignorierte Informationen zur Corona-Krise aufmerksam.“ (vgl. <https://corona-transition.org/uber>, Stand: 17.6.2021).

² Die beiden Quellen ‚Zeitpunkt‘ und ‚Inside Corona‘ wurden zwar für die Berechnung der Word-Embedding-Modelle genutzt, werden bei der Analyse der Verteilung der Cluster nicht berücksichtigt, da ihre Textanteile im Teilkorpus zu gering sind.

Korpus	Anzahl Texte	Anzahl Tokens
Skep-Medien	7.189	22.949.169
Compact Online	2.168	
Corona-Transition	2.511	
KenFM.de	421	
PI-News	1.247	
Rubikon	489	
Zeitpunkt	311	
Inside Corona	42	
Corona-Korpus Total	891.262	84.103.458

3 Methode

3.1 Grundlagen

Die angewendete Methode lässt sich in ein datengeleitetes, korpuslinguistisches Paradigma der Diskurslinguistik einordnen. Mit ‚datengeleitet‘ ist jedoch nicht gemeint, dass die Methode ohne Prämissen wäre. Im Gegenteil: Sie ist theoretisch begründet und trifft Vorannahmen, indem bestimmte Daten ausgewählt und aufbereitet werden.

Der theoretische Hintergrund der Methode ist im Kontextualismus und der Gebrauchsemantik begründet. Beide Theorien gehen davon aus, dass sich die Bedeutung von Ausdrücken im Gebrauch ergibt (Firth 1957). Damit sind diachrone und synchrone Perspektiven auf Gebrauchsveränderungen (z. B. im Rahmen einer ‚Historischen Semantik‘: Busse 1987) und Gebrauchsunterschiede (Spitzmüller/Warnke 2011) von Interesse. Methodisch wird diese gebrauchsemantische Perspektive als statistisches Phänomen greifbar: Typische Kontexte von Ausdrücken sind z. B. in Form von Kollokationen berechenbar (Evert 2009).

Mit der distributionellen Semantik wurde in den letzten Jahrzehnten ein Set von Methoden entwickelt, die auf der Basis großer Datenmengen Algorithmen des neuronalen Lernens nutzen und eine konsequente Weiterführung der klassischen Kollokationsanalyse darstellen (für einen Überblick und linguistische Einordnung: Lenci 2018). Eine dieser Methoden ist das sogenannte word2vec-Verfahren (Mikolov et al. 2013): Die über neuronales Lernen und über mehrere Abstraktionsebenen hinweg (‚Deep Learning‘) gelernten Vektoren repräsentieren die typischen Kontexte, in denen die Ausdrücke im Korpus auftreten. Ausdrücke mit ähnlichen

Vektoren sind also Ausdrücke, die in ähnlichen Kontexten auftreten und damit in semantischer Nähe zueinander stehen.

Auf der Basis des Corona-Korpus so berechneten sogenannten ‚Word-Embedding‘-Modells können nun für alle Types im Korpus die ‚nächsten Nachbarn‘ ausgegeben werden, also die Ausdrücke, die ähnliche Vektoren aufweisen. Für den Ausdruck *Maske* sind das mit abnehmender Ähnlichkeit: *Schutzmaske*, *Gesichtsmaske*, *Atemschutzmaske*, *Mundschutz*, *FFP2 Maske*, *Hygienemaske*, *Maskentragen*, *FFP2-Masken*, *FFP2*, *tragen*, *Maulkorb*, *Tragen*, *nichts nützen*, *Stoffmaske*, *Mund-*, *Nasenschutz*, *Maske tragen*, *OP-Masken*, *FFP 2* etc.

In Abbildung 1 ist dieser semantische Raum grafisch und reduziert auf drei Dimensionen dargestellt.

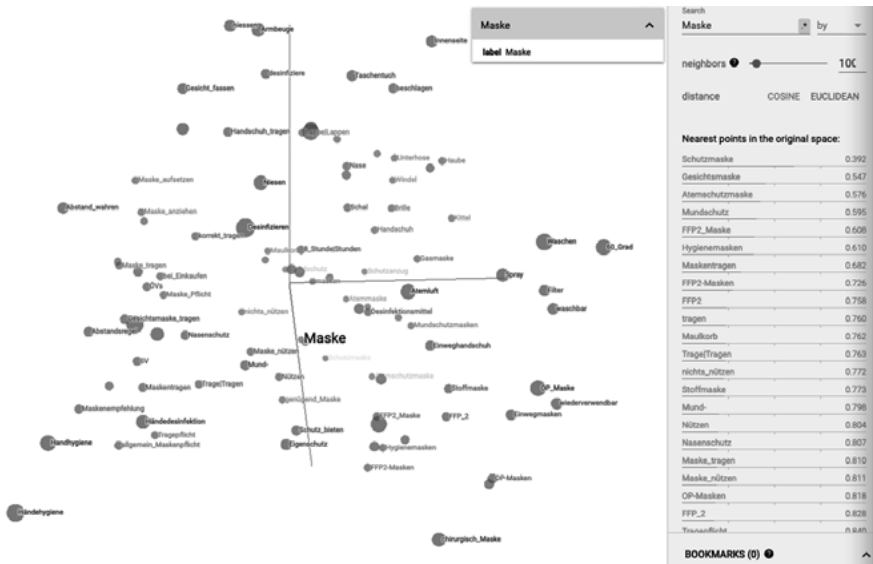


Abb. 1: Der semantische Raum um „Maske“ im Corona-Korpus

Die Liste der nächsten Nachbarn zu *Maske* macht deutlich, dass es sich bei den ersten Ausdrücken vielleicht um Synonyme oder Quasisynonyme handelt, mit Zunahme der Distanz (es handelt sich um den Winkel zwischen den Vektoren) treten jedoch auch Ausdrücke auf, die nicht mehr als Synonyme im engen Sinn akzeptiert würden – aber: Sie stehen zweifellos in einer semantischen Nähe, vor allem einer sehr kontextspezifischen, diskurspezifischen Nähe. Wenn nun auf der einen Seite es als Mangel der Methode verstanden werden kann, dass sie

semantische Relationen (auch z. B. Synonyme und Antonyme) vermischt, mitunter auch unterschiedliche Wortarten auftreten können, scheint es auf der anderen Seite aus gebrauchts- und diskursemantischer Sicht gerade besonders interessant, dass ein semantisches Kontinuum von Ähnlichkeit deutlich wird. Diese Beobachtung einer diskursspezifischen „funktionalen semantischen Äquivalenz“ (Bubenhofer 2020) gilt es auch in der vorliegenden Studie zu prüfen.

Der gebrauchtssemantische theoretische Hintergrund würde nun postulieren, dass sich die semantischen Räume je nach Datengrundlage unterscheiden müssten, ähnlich wie sich zwei Kollokationsprofile des gleichen Ausdrucks in unterschiedlichen Daten unterscheiden. Dies ist auch in den Word-Embedding-Modellen der Fall. Das auf der Basis des Skep-Medien-Korpus berechnete Modell gibt für *Maske* die folgenden nächsten Nachbarn: *Mundschutz*, *Maulkorb*, *Tragen*, *tragen*, *Schutzmaske*, *Abstand*, *Gesichtsmaske*, *FFP 2*, *Alltagsmaske*, *Staubsaugerbeutel*, *Mund*, *Alltagsmasken*, *Nase*, *Maske tragen*, *Virenschleuder*, *FFP2*, *Atemschutzmaske*, *Maskentragen*, *FFP2-Masken*, *Lappen* etc.

Die beiden Listen ähneln sich, es gibt jedoch auch Unterschiede: *Maulkorb* ist ähnlicher zu *Maske* als im Gesamtkorpus und *Virenschleuder* oder *Lappen* sind nicht unter den 20 nächsten Nachbarn von *Maske* im Gesamtkorpus.

Die Unterschiede – so die Hypothese – müssten sich diskurslinguistisch deuten lassen und stehen für die verschiedenen semantischen Räume. Unterschiedliche Positionen im Corona-Diskurs führen also zu unterschiedlichen semantischen Räumen – und umgekehrt: Die Argumentationen bewegen sich innerhalb geschlossener semantischer Räume und konstituieren so die Positionen.

3.2 Ausgestaltung der Methode

Die grundlegenden Überlegungen machen deutlich, dass die Word-Embedding-Modelle für eine qualitative Analyse aufbereitet werden müssen. In der vorliegenden Studie wird ein Vorgehen gewählt, das die datengeleitet berechneten Modelle mit einer Mischung von quantitativer Aufbereitung und qualitativer Analyse nutzt.

Es kommen zwei Verfahren zur Anwendung. Das erste Verfahren versucht ein Word-Embedding-Modell qualitativ interpretierbar zu machen, indem es über ein Clustering-Verfahren strukturiert und dann die Verteilung ausgewählter Cluster über die Quellen analysiert wird:

1. In einem ersten Schritt wurde das Word-Embedding-Modell mit dem word2vec-Algorithmus und der Skip-gram-Option (Mikolov et al. 2013, S. 7) berechnet unter Verwendung des Python-Gensim-Pakets (Řehůřek/Sojka 2010). Es

wurden Modelle für alle Teilkorpora und das Gesamtkorpus berechnet. Der Berechnung wurde eine n-Gramm-Analyse vorgeschaltet, um häufige Wortkombinationen bis zu einer Länge von $n = 3$ zu identifizieren. Zudem wurden die Berechnungen auf Basis der Grundformen (Lemma) und nicht der Wortformen vollzogen.

2. Im Anschluss wird ein Clustering eingesetzt, um das Word-Embedding-Modell in kleinere semantische Räume zu gliedern: Das Clustering soll Gruppen von Ausdrücken identifizieren, die im Vektorraum nahe beieinander stehen. Dafür wird der K-means-Algorithmus verwendet (Arthur/Vassilvitskii 2006), bei dem die Anzahl gewünschter Cluster vorgegeben wird. Da das Corona-Korpus-Modell gut 24.000 Types enthält, wurden 2.000 Cluster vorgegeben, so dass im Schnitt 12 Types pro Cluster enthalten sind. Dies ist eine Granularität, die sich bereits bei vorherigen Analysen als brauchbar erwiesen hat (Bubenhof 2020). Für jeden Cluster wird zudem ein Label vergeben, um ihn für die folgenden Analysen besser handhabbar zu machen. Dafür wurde der sogenannte Zentroid berechnet, also der Punkt, der im durch den Cluster beschriebenen Bereich im Vektorraum in der Mitte steht. Die drei diesem Punkt am nächsten liegenden Types bilden das Label.
3. Die so entstandenen Cluster können nun einer qualitativen Analyse unterzogen werden, wobei sie mit weiteren Daten angereichert sind: Zu jedem Cluster ist bekannt, wie viele Ausdrücke er enthält und wie häufig die Ausdrücke im Korpus sind. Die Cluster werden kategorisiert nach durch das Forschungsinteresse bestimmten Kriterien; im vorliegenden Fall liegt das Interesse bei Clustern mit thematischer Ähnlichkeit, bei denen jedoch unterschiedliche Semantisierungen sichtbar werden.
4. Manuell ausgewählte Cluster von Ausdrücken werden hinsichtlich ihrer Verteilung in den Daten visualisiert: Einerseits interessiert, wie häufig Ausdrücke des Clusters in den verschiedenen Quellen im Korpus verwendet werden, andererseits wie sie zeitlich verteilt sind (letzteres kommt in der folgenden Analyse nicht zur Anwendung).
5. Die vorhergehenden Schritte erlauben dann festzustellen, was quellenspezifische und gemeinsame semantische Räume sind; die Gründe dafür können diskurslinguistisch gedeutet werden.

Im zweiten Verfahren werden zwei Word-Embedding-Modelle verglichen mit dem Ziel, Ausdrücke zu finden, die in beiden Modellen vorkommen, jedoch unterschiedlich semantisiert sind (die Methode wurde mit Philipp Dreesen und Selena Calleri zusammen entwickelt und bereits in Bubenhof/Calleri/Dreesen 2019 beschrieben):

1. Ausgangspunkt ist die Überlegung, dass zwei Ausdrücke in zwei Word-Embedding-Modellen dann unterschiedlich semantisiert sind, wenn sie unterschiedliche nächste Nachbarn aufweisen. Potenziell müssen jedoch die nächsten Nachbarn in einem Modell auch im anderen als Types vorhanden sein.
2. Für beide Modelle werden für alle Types die 100 nächsten Nachbarn berechnet.
3. Es wird für jeden Type ausgegeben, ob seine 100 nächsten Nachbarn auch Types im anderen Modell sind (die sogenannte Out-of-Vocabulary-Kennzahl) und wie groß die Überlappung mit den 100 nächsten Nachbarn des gleichen Ausdrucks im anderen Modell ist (die sogenannte Overlap-Kennzahl).
4. Von Interesse sind nun Ausdrücke, bei denen die Überlappung der 100 nächsten Nachbarn und die Out-of-Vocabulary-Kennzahl sehr klein sind. Zweites Kriterium ist, dass der Ausdruck in beiden Korpora, deren Modelle verglichen werden, häufig genug vorkommt (andernfalls handelt es sich um Ausdrücke, die zu spezifisch für das eine Teilkorpus sind). Die gefundenen Ausdrücke werden dann qualitativ evaluiert und gedeutet.

Im folgenden Kapitel werden nun ausgewählte Beispiele von Clustern und ihre Verteilungen über Quellen und den Zeitraum Januar 2020 bis Mai 2021 dargestellt und diskutiert sowie die im Vergleich des CH-Medien-Korpus-Modells mit dem Skep-Medien-Korpus-Modell auffälligen Ausdrücke analysiert.³

4 Analysen

4.1 Diktatur

Die Frankfurter Allgemeine Zeitung titelt am 25. Februar 2021: „Die reichste Schweizer Unternehmerin hält ihr Land für eine Diktatur“.⁴ Damit ist die SVP-Politikerin Magdalena Martullo-Blocher gemeint, die in einem Interview mit der NZZ am 12. Februar 2021 meinte: „Der Bund hat eine Diktatur eingeführt. Er hat

³ Die vollständigen Daten der Analyse finden sich online auf: www.bubenhofer.com/publikationen/corona/ (Stand: 17.6.2021).

⁴ <https://www.faz.net/aktuell/wirtschaft/unternehmen/milliardaerin-aus-der-schweiz-haelt-ihr-land-fuer-eine-diktatur-17214698.html> (Stand: 17.6.2021).

die Demokratie ausgeschaltet“ und „Der Bund regiert mit Schreckensszenarien, damit er die Macht behalten kann“ (NZZ, 12.2.2021). Die Interpretation, den Bundesrat (die Exekutive) als Diktator zu bezeichnen, stößt auf große Kritik, zumal im Bundesrat auch die SVP (neben den anderen wählerstärksten Parteien) mit zwei von sieben Sitzen vertreten ist.

Die Lesart ist jedoch anschlussfähig an semantische Räume der Neuen Rechten und anderer Corona-Skeptiker und -Skeptikerinnen. Unter den semantischen Clustern finden sich einige, die den Staat thematisieren. Abbildung 2 zeigt sechs ausgewählte Cluster und ihre Verteilung auf die Quellen. Der prominenteste Cluster ist:

Label: demokratisch Staat, zentralistisch, totalitär Regime⁵ (Cluster 934)
 demokratisch (6924)⁶, regieren (3432), totalitär (1854), autoritär (1197), freiheitlich (813), diktatorisch (737), politisch_System (421), föderalistisch (415), undemokratisch (400), rechtsstaatlich (373), Bananenrepublik (329), Diktat (266), Gebilde (257), totalitär_Staat (207), zentralistisch (205), Rechtssystem (194), autokratisch (185), Staatsform (176), demokratisch_Staat (171), totalitär_System (169), souverän_Staat (141), Grundpfeiler (112), totalitär_Regime (101), Staatenbund (99), Gewaltentrennung (81)

Die Verteilung über die Quellen in Abbildung 2 zeigt nun deutlich, dass Ausdrücke des Clusters in den Skep-Medien deutlich häufiger verwendet werden. Die Kommentare (mit Ausnahme von 20 Minuten) liegen zusammen mit den CH-Medien NZZ und Tages-Anzeiger dazwischen. In den restlichen Quellen werden die Ausdrücke des Clusters seltener verwendet. Es muss berücksichtigt werden, dass mit *demokratisch* und *regieren* zwei Ausdrücke im Cluster vorhanden sind, die hinsichtlich eines politischen Standpunktes (in einer Demokratie) unauffällig sind. Die meisten der restlichen Ausdrücke jedoch sind Stigmawörter (Spitzmüller/Warnke 2011, S. 143) und kritisieren im Kontext der Corona-Pandemie die angeblich totalitäre Ausrichtung des Staates.

⁵ Wie in Abschnitt 3.2 erläutert, wurden Grundformen (Lemma) als Grundlage für die Berechnungen verwendet, weshalb auch in allen weiteren Angaben die Grundformen genannt werden.
⁶ Sortierung nach Häufigkeit absteigend, in Klammern Häufigkeit im Korpus (absolute Häufigkeiten).

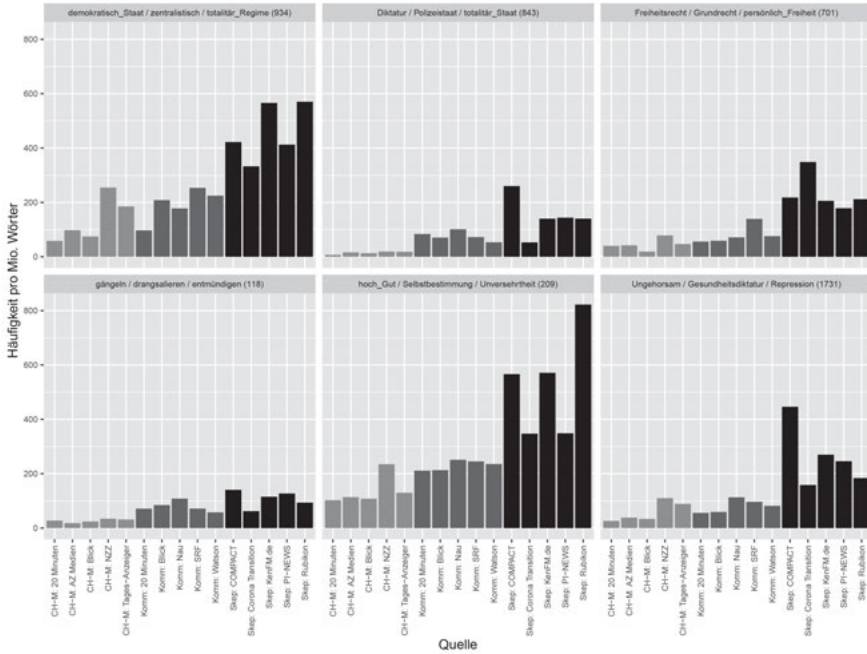


Abb. 2: Häufigkeiten semantischer Räume zur Beziehung zum Staat (Titel der Diagramme: Label des Clusters und Cluster-ID in Klammer)

Auch die Cluster 843 (*Diktatur, Polizeistaat, totalitär Staat*) und 1731 (*Ungehorsam, Gesundheitsdiktatur, Repression*) enthalten Stigmawörter für den Staat und werden in den Skep-Medien häufiger verwendet als in den anderen Quellen. Mit Cluster 118 (*gängeln, drangsaliieren, entmündigen*) ist ein Cluster vorhanden, das hauptsächlich Verben enthält und das Verhältnis zwischen Bürger und Bürgerinnen und Staat anspricht:

Label: gängeln, drangsaliieren, entmündigen (Cluster 118)

aufzwingen (690), bevormunden (548), erpressen (526), quälen (492), ohne_Rücksicht (417), drangsaliieren (417), schikanieren (374), grundlos (283), gefügig (265), gängeln (257), erdulden (212), aufdrücken (196), demütigen (194), Freiheit_einschränken (185), Geiselhaft (184), entmündigen (178), Freiheit_berauben (117)

Die Verteilung dieses Clusters über die Quellen (Abb. 2, unten links) zeigt, dass dessen Ausdrücke in den Kommentaren und den Skep-Medien ähnlich häufig verwendet werden, jedoch seltener in den CH-Medien.

Eine Diktatur oder ein totalitärer Staat ist wohl für die Mehrheit der Bewohnerinnen und Bewohner westlicher Demokratien eine unerwünschte Staatsform. Indem in den Texten der Corona-Skeptiker und -Skeptikerinnen und der Neuen Rechten jedoch die Schweiz (und Deutschland) als Diktatur(en) semantisiert wird (werden), ergibt sich ein semantischer Raum vieler daran anschließender Bezeichnungen und Verhaltensformen (vgl. die Verben wie *drangsalieren*, *bevormunden* etc.).

In den Kommentaren sind die Cluster aus Abbildung 2 mäßig vertreten – meist häufiger als in den CH-Medien, jedoch weniger als in den Skep-Medien. Das liegt wahrscheinlich daran, dass die Kommentarplattformen im Vergleich zu den Skep-Medien ein heterogeneres Meinungsspektrum ermöglichen. Die Textsorte ist von Dialog geprägt, und es sind in den Kommentaren intensive Diskussionen beobachtbar. Dies zeigt folgender Beleg zu einem Artikel vom 22. April 2020 auf blick.ch zum Entscheid des Schweizer Bundesrates, nur großen Geschäften eine komplette Öffnung zu ermöglichen, nicht jedoch den kleinen Läden.⁷ Im Folgenden ein Ausschnitt aus den Kommentaren:⁸

- (1) A: Ein **salomonischer Entscheid!** Und man hat damit rechnen müssen. Zum Glück lässt sich unser BR **nicht erpressen**
- B: Lasst den Bigler [Gewerbeverbands-Direktor Hans-Ulrich Bigler, NB] mal wettern, der Bundesrat **sagt wo es lang geht** und keine Profilierungsneurotiker!
- C: Befassen Sie sich mal mit seinen Argumenten, denn da hat er absolut Recht. Es ist ein Trauerspiel was abgeht wegen ein paar Risikogruppen. Es ist eine Weltweite Hysterie wegen eines recht Harmlosen Virus.
- D: **Der Virus diktiert**, was zu tun ist und sonst niemand. Der Bundesrat leistet solide Arbeit. Weiter so!
- [...]
- E: Der Bundesrat hat auf breiter Linie **versagt**. [...] Ein weiterer Angriff auf unsere **Freiheit** im Irrglauben, dass der Staat dann alles richten soll, ganz im Sinne der Linken mit dem Ruf nach mehr Staat
- [...]

⁷ Blick.ch: www.blick.ch/wirtschaft/unzufrieden-mit-shopping-kompromiss-des-bundesrats-gewerblern-geht-der-laden-runter-id15857215.html (Stand: 18.6.2021).

⁸ In allen Belegen: Originale Orthographie.

F: Irgendwem konnte man es sowieso nicht recht machen. **Der Bundesrat ist nicht perfekt**, niemand ist das. Er macht es trotzdem [sic!] gut und hat **vollstes Vertrauen**.

G: **Reine Machtdemonstration** des Bundesrats. Bezahlen wird es die Gesellschaft, als Folge der zugrunde gehenden Wirtschaft.

[...]

H: Tja, dass [sic!] ist nun das Resultat der Quänglerei. Dieser Bundesrat hat nun entschieden, dass bis auf weiteres niemand profitieren kann. [...] Und somit hat Er, zusammen mit den Lobbyisten, **das ganze Volk quasi in Geiselnhaft** genommen [...]. Einfach nur noch ein Trauerspiel das ganze.

(Kommentare von Blick.ch: www.blick.ch/wirtschaft/unzufrieden-mit-shopping-kompromiss-des-bundesrats-gewerblern-geht-der-laden-runter-id15857215.html, Stand: 18.6.2021; Hervorh. NB).

Es ist deutlich, dass mit diesen Kommentaren diametral entgegengesetzte Positionen aufeinandertreffen, was sich auch auf Ausdrucksebene zeigt: Die in den Kommentaren verwendeten Bezeichnungen für die Regierung und ihr Handeln sind heterogener als in den Skep-Medien. Dort ist ein homogenerer semantischer Raum zu erwarten.

4.2 Totalversagen

Eng mit der Kritik an der Regierung verbunden ist der Topos des ‚Totalversagens‘ der politischen Führung. In Abbildung 3 sind vier Cluster aufgeführt, die diesen Topos teilweise bedienen.

Die beiden Cluster 285 (*komplett versagen, schlecht vorbereiten, verpennen*) und 1060 (*total Versagen, Totalversagen, Fehlentscheide*) verteilen sich auf die Quellen bezüglich Skep-Medien und Kommentaren mehr oder weniger gegenläufig, was sowohl an stilistischen (Kommentare vs. Mischung von redaktionellen Texten und Kommentaren) als auch nationalen Spezifika (eher Bezüge zur Schweizer Regierung vs. zur bundesdeutschen Regierung) liegt.

Beide Cluster enthalten meist Ausdrücke, die das Handeln der Regierung aufs Heftigste kritisieren (Cluster 285: *überfordert, wursteln, verpennen, verschlampen* etc.; Cluster 1060: *Unfähigkeit, Inkompetenz, Fehlleistung, Schlamperei* etc.). Diese Kritik ist in dieser ausgeprägten Form im CH-Medien-Korpus deutlich seltener zu finden, doch auch dort findet Kritik statt, wie der Cluster (*zaudern, lasche Maßnahme, überhasten*) zeigt.

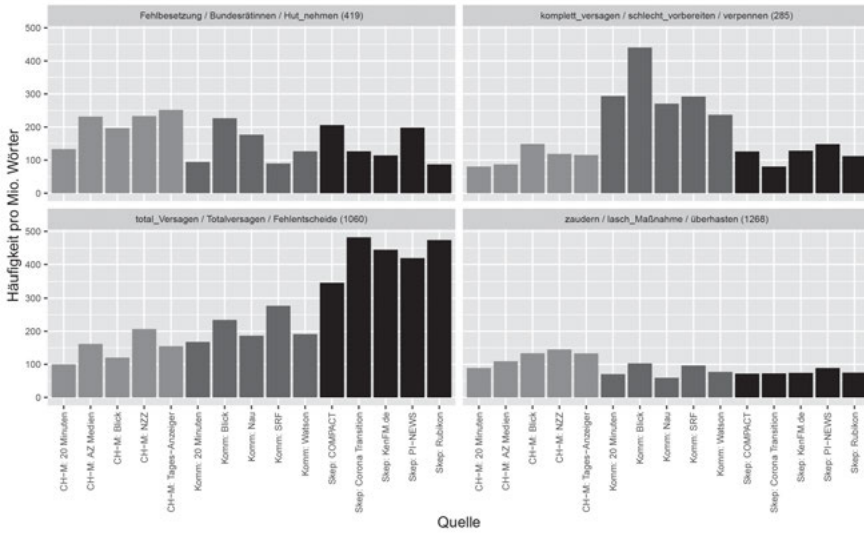


Abb. 3: Häufigkeiten semantischer Räume zu Regierungskritik (Titel der Diagramme: Label des Clusters und Cluster-ID in Klammer)

Cluster 419 (*Fehlbesetzung, Bundesrätinnen, Hut nehmen*) ist ziemlich gleichmäßig über die Quellen verteilt, was jedoch daran liegt, dass das Cluster eine recht heterogene Sammlung von Ausdrücken enthält: Einerseits Ausdrücke, die ebenfalls der Regierungskritik zuzuordnen sind, jedoch auch Bezeichnungen für die verschiedenen Funktionen einer Regierung (*Bundesrätinnen, Politikerin* etc.).

Im Skep-Medien-Korpus und auch in den Kommentaren bedeutet Kritik an der Regierung also Fundamentalkritik und mit dem Topos der ‚überforderten Regierung‘ öffnet sich ein semantischer Raum vieler ähnlicher Ausdrücke.

4.3 Verschwörungserzählungen

Nicht überraschend finden sich im Word-Embedding-Modell eine Reihe von Clustern, deren Ausdrücke auf Verschwörungserzählungen (Butter 2018) verweisen. In Abbildung 4 sind fünf davon abgebildet, wobei vier davon eher von kritischen Auseinandersetzungen mit Anhängerinnen und Anhängern von Verschwörungserzählungen zeugen und nur Cluster 1262 tatsächlich Verschwörungserzählungen nennt:

Label: Weltregierung, Great Reset, neu Weltordnung (Cluster 1262)

global (7893), Revolution (2489), Transformation (954), Great Reset (852), NWO (597), Reset (533), neu Weltordnung (499), Geldsystem (307), Weltregierung (287)

Der Ausdruck *neue Weltordnung*, das Akronym *NWO* („New World Order“) und auch *Weltregierung* verweisen auf die bekannte Verschwörungserzählung einer angeblich im Aufbau begriffenen supranationalen Weltregierung durch Eliten.⁹ Die Bezeichnungen *Great Reset* und *Reset* verweisen allerdings auf eine Initiative des Weltwirtschaftsforums, die Corona-Pandemie als einen Neustart des Kapitalismus zu nutzen – sie wurden jedoch schnell in die Verschwörungserzählung der ‚neuen Weltordnung‘ integriert.¹⁰ Allerdings wird auch *global* diesem Cluster zugeordnet, was den frequentesten Ausdruck des Clusters darstellt und die Verteilung beeinflusst: Abbildung 4 zeigt eine Version der Verteilung ohne das Adjektiv *global*, denn ansonsten wären die Häufigkeiten des Clusters in der NZZ und in Corona Transition etwa 2,5 Mal höher. Dies zeigt jedoch, dass *global* im Kontext der Neuen Rechten und der Corona-Skeptiker und Skeptikerinnen tendenziell anders semantisiert und im Zusammenhang mit Verschwörungserzählungen genannt wird, während es für die liberal-konservative NZZ ein neutralerer Wirtschaftsterminus ist.

Die Verteilung der Cluster macht zudem deutlich, dass die Kontroversen um Verschwörungserzählungen etwas weniger im Skep-Medien-Korpus geführt werden, jedoch deutlich häufiger in den Kommentaren. Besonders Cluster 245 (*Stuss, wirr Zeug, panikmache* [sic!]) ist in den Kommentaren deutlich übervertreten. Auf der anderen Seite ist Cluster 1869 (*stigmatisieren, andersdenkend, verunglimpfen*) im Skep-Medien-Korpus am häufigsten vertreten: Dort wird der Topos gepflegt, dass das Vertreten von Minderheitsmeinungen vom ‚Mainstream‘ stigmatisiert würde.

⁹ Vgl. dazu Wikipedia: [https://de.wikipedia.org/wiki/Neue_Weltordnung_\(Verschwörungstheorie\)](https://de.wikipedia.org/wiki/Neue_Weltordnung_(Verschwörungstheorie)) (Stand: 24.6.2021).

¹⁰ Vgl. dazu Wikipedia: https://de.wikipedia.org/wiki/The_Great_Reset (Stand: 24.6.2021).

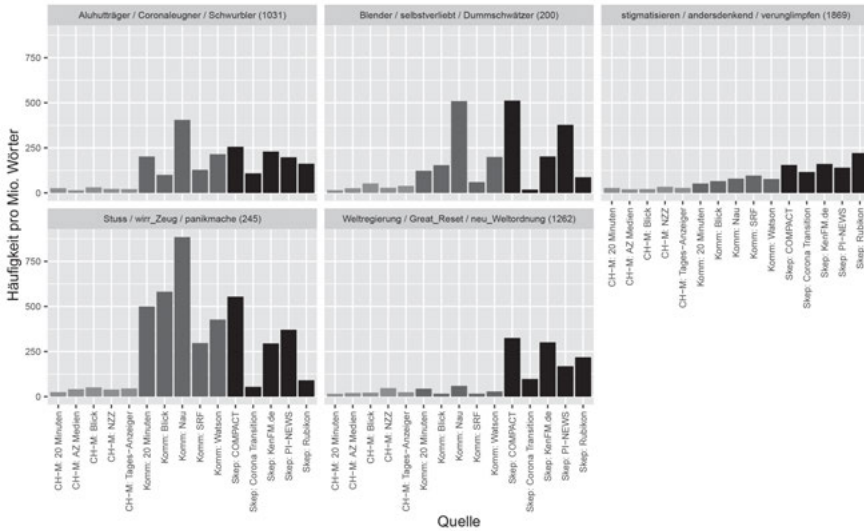


Abb. 4: Häufigkeiten semantischer Räume zu Verschwörungserzählungen (Titel der Diagramme: Label des Clusters und Cluster-ID in Klammer)

4.4 Unterschiedliche Semantisierungen

Ein systematischer Vergleich der Word-Embedding-Modelle des CH-Medien-Korpus und des Skep-Medien-Korpus soll nun zeigen, welche Ausdrücke in den beiden Quellentypen unterschiedlich semantisiert sind.

Für den Vergleich der CH-Medien- und Skep-Medien-Modelle ist die komplette Tabelle online verfügbar.¹¹ Dabei fällt zunächst auf, dass es keine Ausdrücke gibt, bei denen beide Kennzahlen (Overlap und OOV) null sind. Die Liste lässt sich nun aber filtern und es gilt, die richtige Balance zwischen den Kennzahlen zu finden. Werden Overlap und OOV niedrig gewählt, finden sich vor allem Ausdrücke, die in den beiden Korpora sehr unterschiedlich häufig auftreten.¹² Wenn Overlap und

¹¹ Vgl. www.bubenhofer.com/publikationen/corona/.

¹² Beispielsweise gibt es im CH-Medien-Korpus deutlich mehr Sportberichterstattung, in der die Pandemie am Rande auch häufig Thema ist. Dadurch ist die Semantik von Ausdrücken wie *Partie* oder *Team* jedoch im Vergleich zum Skep-Medien-Korpus viel stärker durch den Sport beeinflusst. Das Interesse liegt in der vorliegenden Studie jedoch auf Ausdrücken, die in beiden Korpora häufig verwendet werden.

OOV maximal bei 15 liegen, ergibt sich geordnet nach Häufigkeit absteigend im Skep-Medien-Korpus folgende Auswahl (erste 60 Ausdrücke):¹³

- (2) Bürger, Macht, angeblich, falsch, Wahrheit, Diktatur, Nazi, zerstören, Recht|Rechte, böse, Elite, Feind, hängen, Bill Gates, vermeintlich, kriminell, Lüge, sozialistisch, Zerstörung, illegal, Unsinn, lächerlich, Mainstream, jeglich, Herrschaft, erheben, Böse, befreien, mißbrauchen, entziehen, Arm|Arme, Vorgang, aggressiv, schreien, Verschwörung, Sklave, unterdrücken, Zensur, Unrecht, Versagen, Handlung, Vorwand, Mächtige, Menschlichkeit, Stalin, entlarven, Obrigkeit, Narrativ, Mitbürger, Unterdrückung, Andersdenkende, betrügen, selbsternannt, Selbstmord, befehlen, bösartig, Feststellung, Ausbeutung, attackieren, suggerieren

Anhand des Ausdrucks *Macht* sei der Unterschied der Semantisierung in den beiden Korpora gezeigt:

CH-Medien: Nächste Nachbarn (Top 35) zu *Macht* (absteigende Nähe, gerundet): Herrschaft (0.8), Sozialismus (0.8), politisch Elite (0.8), Demokratie (0.8), Revolution (0.8), frei Wahl (0.8), Militärdiktatur (0.8), freiheitlich (0.8), demokratisch Institution (0.7), allmächtig (0.7), totalitär (0.7), politisch System (0.7), Nationalist (0.7), liberal Demokratie (0.7), Westen (0.7), Multilateralismus (0.7), Reformier (0.7), Autokrat (0.7), Kommunist (0.7), kommunistisch (0.7), Innenpolitik (0.7), Populisten (0.7), unantastbar (0.7), Führer (0.7), Deutungshoheit (0.7), Sowjetunion (0.7), Vorherrschaft (0.7), Schalthebel (0.7), Nationalstaat (0.7), Rechtsstaat (0.7), verhaßt (0.7), populistisch (0.7), autokratisch (0.7), Islam (0.7), sozial Gerechtigkeit (0.7)

Skep-Medien: Nächste Nachbarn (Top 35) zu *Macht* (absteigende Nähe, gerundet): Schalthebel (0.8), Hebel (0.8), eigen Vorteil (0.8), Nutznießer (0.8), Herrschaft (0.8), Mächtige (0.8), verschwören (0.8), Knute (0.8), Idiologie [sic!] (0.8), Finanzelite (0.8), kapitalistisch System (0.8), Eigennutz (0.8), Geldadel (0.8), festigen (0.8), Machterhalt (0.8), Machtelite (0.8), Ausbeuter (0.8), Hierarchie (0.8), nützlich Idiot (0.8), Machtmittel (0.8), Herrscher (0.8), Ziel erreichen (0.8), Plutokrat (0.8), eigen Interesse (0.8), frei Marktwirtschaft (0.8), echt Demokratie (0.8), erringen (0.8), Großkapital (0.8), Weltherrschaft (0.8), Zwietracht (0.8), Besitzende (0.8), Weltfrieden (0.8), Ruhm (0.8), Mittel recht (0.8), über Leiche (0.8)

¹³ Vgl. für die vollständigen Angaben und eigene Möglichkeiten der Filterung die Begleitseite mit den Daten zur Studie (www.bubenhofer.com/publikationen/corona/).

Im CH-Medien-Korpus sind die semantisch ähnlichen Ausdrücke zu *Macht* im Vergleich deutlich weniger kritisch. Im Skep-Medien-Korpus befinden sich Ausdrücke wie *Schalthebel*, *eigenen Vorteil*, *Knute*, *Eigennutz*, *Machterhalt* etc. in ähnlichen Kontexten wie *Macht*, also Ausdrücke, die den Missbrauch von Macht beschreiben. Das Verb *verschwören* bringt explizit Verschwörungstheorien mit Macht in Verbindung und *nützlich[er] Idiot* verweist auf Dritte, die andere ausnutzen:

- (3) Doch Sven Román, einer der unterzeichnenden Ärzte, ist skeptisch gegenüber den offensichtlichen Verbindungen der Behörden zur Pharmaindustrie. Richard Bergström war beispielsweise zuvor der CEO der Pharmaceutical Industries and Associations (EFPIA).
 «[...] Die Schweden müssen erkennen, dass diese Menschen Interessenbindungen haben und nicht unparteiisch sind», twitterte Román über Bergström.
 [...] Für Sven Román sind viele Experten **nützliche Idioten** für die Pharmaindustrie.
 „Schwedische Ärzte warnen davor, Junge und Gesunde zu impfen“, in: Corona Transition, 2021-05-10.

Ein Vergleich der Kollokationsprofile von *Macht* in den beiden Korpora zeigt zwar auch Unterschiede, jedoch weniger deutlich.¹⁴ Im Hinblick auf die 50 signifikantesten Kollokatoren sind im Skep-Medien-Korpus *Antifa*, *Arroganz*, *Gier*, *Herrschaft*, *links* oder *Untergrund* beispielsweise statistisch signifikante Kollokatoren zu *Macht*. Auch im CH-Medien-Korpus gibt es negativ konnotierte Kollokatoren zu *Macht* (z. B.: *klammern*, *korrumpieren* oder *Verschwörung*), daneben jedoch weit mehr primär Macht spezifizierende Kollokatoren wie *Exekutive*, *Militär*, *militärisch*, *Nationalstaat*, *Partei*, *Politik* oder *wirtschaftlich*. Natürlich gibt es auch viele gemeinsame Kollokatoren wie *abgeben*, *ausüben*, *dunkel*, *Einfluss*, *Geld*, *missbrauchen*, *politisch*, *Reichtum* oder *Staat*. Es ist schwieriger, aufgrund der Kollokationsprofile Differenzen der Semantisierung von *Macht* zu identifizieren; der Vergleich der jeweils semantisch ähnlichen Ausdrücke in den beiden Korpora macht jedoch deutlich, dass in den Skep-Medien hinter der Kritik an Macht (die natürlich auch in den CH-Medien vorkommt) viel deutlicher eine Sicht steht, die hinter der Macht eine Verschwörung sieht.

¹⁴ Kollokationen auf Lemma-Basis, Kontext 5 Wörter links/rechts, Log Likelihood Ratio LLR.

Die weiteren Ausdrücke in der Belegliste (2) liefern bei genauerer Analyse ebenfalls Belege für diese Hypothese, so etwa beim Ausdruck *Bill Gates*: Dort verweisen dazu semantisch ähnliche Ausdrücke im CH-Medien-Korpus zwar explizit auf die Verschwörungserzählungen (*Fake-News*, *Verschwörung*, *Desinformation* etc.), im Skep-Medien-Korpus wird anhand der semantisch ähnlichen Ausdrücke jedoch deutlich, dass dort die Verschwörung tatsächlich erzählt wird: *Rockefeller*, *Soros*, *Schwab*, *Big Pharma*, *Musk*, *Rothschild* etc.

Der Ausdruck *Narrativ* soll an dieser Stelle noch genauer untersucht werden: In den beiden Modellen sind nur *Halbwahrheit*, *instrumentalisiert*, *Massenmedium*, *Populismus* und *Verschwörung* semantisch ähnliche Ausdrücke, die in beiden Modellen vorkommen. Ansonsten unterscheiden sich die Modelle deutlich. Nächste Nachbarn wie *Angsterzeugung*, *aufbauschen*, *böse Absicht*, *dahinter stecken*, *Fiktion*, *infam*, *Inszenierung*, *krampfhaft*, *Leugnung*, *Lügenpropaganda*, *Massenpsychologie*, *psychologische Kriegsführung*, *Unwahrheit*, *Wahn* oder *Weltverschwörung* zeichnen einen semantischen Raum, bei dem *Narrativ* als bösertige Manipulation semantisiert ist. Im CH-Medien-Korpus gibt es ebenfalls Ausdrücke, die diese Semantisierung nahelegen, jedoch auch andere wie *westliche Demokratie*, *Ideologie*, *Deutungshoheit* oder *politische Debatte*, die der sozialwissenschaftlichen Definition von *Narrativ* als sinnstiftende Erzählung näher kommen (Lyotard 2015).

Die beiden Beispiele *Macht* und *Narrativ* zeigen, wie die Neuen Rechten und Corona-Skeptiker und -Skeptikerinnen diese Ausdrücke systematisch anders semantisieren als es in den Orientierungsmedien geschieht. In beiden Fällen ist die bösertige Manipulation, teilweise als *Verschwörung* aufgefasst, der dominante Hintergrund der Lesart.

5 Diskussion und Fazit

Die Analysen sind ein Versuch, in einem Wechsel von datengeleiteter, quantitativer Prozessierung und qualitativer Analyse systematische Differenzen in der Ausgestaltung der semantischen Räume in unterschiedlichen Diskursausschnitten zu identifizieren. Die Verwendung von Word-Embeddings stellt dabei eine Methode dar, die theoretisch über den Kontextualismus und die Prämissen einer Gebrauchssemantik mit diskursiver Prägung begründbar ist und auf einer abstrakten Ebene Semantik fasst. Im Zentrum steht nicht die Gebrauchsbedeutung einzelner Ausdrücke, sondern eine umfassende Analyse verschiedener Ausdrücke, die ähnlich kontextualisiert werden.

Wichtig ist dabei, die berechneten Modelle ‚lesbar‘, also einer qualitativen Analyse zugänglich zu machen. Das Clustering der Modelle und die Analyse der Verteilung der Cluster in den Korpora, jedoch auch der systematische Vergleich hinsichtlich unterschiedlicher Semantisierungen sind Versuche, dies zu erreichen.

Der Einsatz von Word Embeddings für gebrauchsssemantische Fragestellungen ist vielversprechend, wie auch bereits Arbeiten von Kozłowski/Taddy/Evans (2018) und auch eigene Analysen gezeigt haben (Bubenhofer 2020, 2021; Bubenhofer/Calleri/Dreesen 2019; Bubenhofer et al. 2020). Sie zeigen über die Verwendung von einzelnen Ausdrücken hinausgehend, dass diese Teil eines semantischen Raums sind, der in sich schlüssig aber inkommensurabel zu anderen semantischen Räumen ist.

Damit kann im Anschluss an Joachim Scharloth (vgl. seinen Beitrag in diesem Band) von einer ‚Schließung rechter Diskurswelten‘ gesprochen werden: Die systematisch im Vergleich zu anderen Diskursen differenten Kontextualisierungen führen zu eigenen semantischen Räumen, die Kommunikation innerhalb desselben Kommunikationssystems, hier also der Corona-Skeptikerinnen und -Skeptiker, problemlos anschließbar macht, Anschlüsse an andere Kommunikationssysteme jedoch erschwert. Eine besondere Stellung nehmen die relativ breit konsumierten Newsplattformen ein: Über die Kommentare dort interagieren Akteure und Akteurinnen miteinander, die in den anderen Medien seltener aufeinandertreffen. Prinzipiell wäre dort deshalb auch der Dialog zwischen den verschiedenen Positionen möglich, allerdings bewegt sich ein Teil der Akteurinnen und Akteure in inkommensurablen semantischen Räumen, was einen Dialog verunmöglicht. Statt Dialog werden bloß Meinungen geäußert.

Wo ein Dialog nicht mehr stattfindet, sind neurechte politische Positionen kaum mehr an andere Positionen anschließbar und gleichzeitig wirken für die Insider und Insiderinnen die entstehenden semantischen Räume in sich schlüssig.

Literatur

- Arthur, David/Vassilvitskii, Sergei (2006): k-means++: the advantages of careful seeding. Technical Report Nr. 2006–13. In: SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Stanford: InfoLab, S. 1027–1035.
- Belica, Cyril/Perkuhn, Rainer (2015): Feste Wortgruppen/Phraseologie I: Kollokationen und syntagmatische Muster. In: Haß, Ulrike/Storjohann, Petra (Hg.): Handbuch Wort und Wortschatz. (= Handbücher Sprachwissen 3). Berlin/Boston: De Gruyter, S. 201–225. DOI: 10.1515/9783110296013-009.

- Bubenhofer, Noah (2020): Semantische Äquivalenz in Geburtserzählungen: Anwendung von Word Embeddings. In: *Zeitschrift für germanistische Linguistik* 48, 3, S. 562–589. DOI: 10.1515/zgl-2020-2014.
- Bubenhofer, Noah (2021): Masken und Küsschen: Korpuslinguistische Exploration des Corona-Diskurses in der Deutschschweiz. In: *Mitteilungen des Deutschen Germanistenverbandes* 68, 2, S. 127–140.
- Bubenhofer, Noah/Calleri, Selena/Dreesen, Philipp (2019): Politisierung in rechtspopulistischen Medien: Wortschatzanalyse und Word Embeddings. In: *OBST (Osnabrücker Beiträge zur Sprachtheorie)* 95, S. 211–242.
- Bubenhofer, Noah/Knuchel, Daniel/Sutter, Livia/Kellenberger, Maaïke/Bodenmann, Niclas (2020): Von Grenzen und Welten: Eine korpuspragmatische COVID-19-Diskursanalyse. In: *Aptum. Zeitschrift für Sprachkritik und Sprachkultur* 16, 2–3, S. 156–165.
- Busse, Dietrich (1987): *Historische Semantik. Analyse eines Programms.* (= Sprache und Geschichte 13). Stuttgart: Klett-Cotta.
- Butter, Michael (2018): »Nichts ist, wie es scheint«. Über Verschwörungstheorien. Berlin: Suhrkamp.
- Evert, Stefan (2009): Corpora and collocations. In: Lüdeling, Anke/Kytö, Merja (Hg.): *Corpus Linguistics. An international handbook*. Bd. 2. (= Handbücher zur Sprach- und Kommunikationswissenschaft (HSK) 29.2). Berlin/New York: De Gruyter, S. 1212–1248.
- Firth, John Rupert (1957): Modes of meaning. In: Firth, John Rupert: *Papers in Linguistics 1934–1951*. London: Oxford University Press, S. 190–215.
- Kozłowski, Austin C./Taddy, Matt/Evans, James A. (2018): The geometry of culture: analyzing meaning through word embeddings. In: *American Sociological Review* 84, 5, S. 905–949. CoRR. abs/1803.09288.
- Lenci, Alessandro (2018): distributional models of word meaning. In: *Annual Review of Linguistics* 4, 1, S. 151–171. DOI: 10.1146/annurev-linguistics-030514-125254.
- Lyotard, Jean-François (2015): *Das postmoderne Wissen: ein Bericht*. 8., unveränd. Aufl. Wien: Passagen.
- Mikolov, Tomas/Chen, Kai/Corrado, Greg/Dean, Jeffrey (2013): Efficient estimation of word representations in vector space. In: arXiv:1301.3781 [cs].
- Řehůřek, Radim/Sojka, Petr (2010): software framework for topic modelling with large corpora. In: *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Valletta, Malta. Malta: University of Malta, S. 45–50.
- Spitzmüller, Jürgen/Warneke, Ingo H. (2011): *Diskurslinguistik: eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse*. Berlin/New York: De Gruyter.