



Methoden und Analyseoptionen in der digitalen Unternehmens- kommunikation

MICHAELA RUSCH

Abstract Um Aussagen zu linguistischen Fragen u. a. in der digitalen Unternehmenskommunikation und den daraus resultierenden multimodalen Texten treffen zu können, bedarf es Methoden zur Analyse für eine systematische Betrachtung. Einschlägige Werkzeuge zur Textanalyse auf der Makro- und Mikroebene sind zwar schon lange vorhanden, sie wurden aber in den letzten Jahren stark modifiziert. So zeigt sich besonders bei der Korpusanalyse der vermehrte Einsatz von entsprechender Software bzw. Anwendungen zum Erstellen von Korpora und zum Evaluieren der entstandenen Ergebnisse. Software- und anwendungsspezifisch kann zwischen diversen Einstellungen auf technischer Ebene gewählt werden. So ermöglicht zum Beispiel die computergestützte Untersuchung von Kommentarspalten in den Sozialen Medien von Firmenkanälen ein individuelleres Einordnen bzw. eine spezifischere qualitative oder diskursanalytische Interpretation. Neben der linguistischen Korpusanalyse haben aber auch weitere Methoden im Rahmen der Computerlinguistik und des Data Mining zu einer Erweiterung der Datenbasis geführt. Die Untersuchung stellt daher anhand von Beispielen aus dem Französischen, Englischen und Deutschen derzeitige Trends zu Methoden vor und verdeutlicht ihre Aktualität sowie eine mögliche Applikation für die Lehre. Verschiedene Forschungsbeispiele bereits etablierter Fachsprachenkorpora unterstreichen die große Bedeutung der Thematik.

Keywords Unternehmenskommunikation, multimodale Inhalte, linguistische Analyse, qualitative und quantitative Analyse, Medienlinguistik

Durch den technologischen Wandel und die Fortschritte in der Digitalisierung können inzwischen große Mengen authentischer Sprachdaten kompiliert, aufbereitet und analysiert werden. (Wicher 2019: 79)

1. Korpuslinguistik im (Technik-)Wandel

Die zahlreichen grundlegenden Veränderungen im Bereich der Korpuslinguistik können durch verschiedene Aspekte initiiert werden. Allen voran steht der technische Wandel. Seit ihren Anfängen hat sich

für die Korpuslinguistik die zu bearbeitende Datenmenge wiederholt als Herausforderung gezeigt. Händisches Auszählen, Transkribieren oder das Annotieren von Texten zur Analyse waren sehr zeitaufwendige Tätigkeiten und bedurften nicht minder einer großen Geduld seitens der Forschenden. Mit der Veränderung im Bereich der Computerlinguistik, also zum Beispiel auch der verbesserten Rechnerleistung, bieten sich nun vielfältige und schnelle Lösungen für ehemals langwierige Untersuchungsprozesse.

Mittlerweile ist es aber nicht nur die Hardware und/oder Softwareausstattung, die zum Gamechanger in der Analyse wird. Neue Fächer wie *Data Science* als innovatives Fachgebiet wirken hier außerdem unterstützend und fördern so durch die gewonnenen Erkenntnisse über Methoden wie Data Mining, Lokalisierung u.a. die Produktivität und Aussagekraft der Ergebnisse zu den untersuchten Texten.

Ein weiterer Grund für die Entwicklung und Veränderung in der Analyse kann bei den Korpuswerkzeugen gesehen werden. Diese erscheinen zunehmend intuitiv und ihre Bedienung wird erleichtert, wie das Beispiel von AntConc^{®1} (nun auch mit einer KI-Erweiterung verfügbar) zeigt. Mit einer stetig steigenden Analysepräzision wird der Bereich Korpuslinguistik als maßgeblicher Motor zur Analyse von Sprachdynamik immer relevanter. Gerade auch im übergreifenden Anwenden der neuen Lösungen können Fragestellungen zur „theoretischen“ Linguistik und auch in Teilen zur angewandten besser verfolgt und evaluiert werden.

[...] There can be no doubt that CL [Corpus Linguistics] has benefited from technological progress in the last decade. Computers work faster; software has, in the main, become more intuitively usable by scholars and others who are non-specialists in computer science; and the mathematical processes which yield results that linguists can interpret have become more sophisticated, and indeed CL has led many of these changes [...] (O’Keeffe & McCarthy 2022: 1)

1 AntConc[®](2024): <https://www.laurenceanthony.net/software/antconc/> (2024)

Auch die Gründe für die Notwendigkeit von Korpusanalysen können sich ändern, z.B. durch politische, soziale oder wirtschaftliche Veränderungen. Es können durch sie Erkenntnisse erzielt werden, die sprachliche Entscheidungen von Autor:innen wie Rezipient:innen beeinflussen. So gibt es bspw. Untersuchungen zum *Brexit* (Jamet & Rodet 2021), zu *Covid* (Davies 2021), zum *Krieg gegen die Ukraine* (Sonnenhauser, Eckhoff & Fortuin 2022) und zur *Energiewende* (Wiertz, Kuhn & Mattissek 2023).

It has become the norm that major political and societal events come under the corpus linguist's research gaze. (O'Keeffe & McCarthy 2022: 2)

Generell bieten Themen gesellschaftlicher Diskurse gute Ausgangspunkte zur Initiierung korpusanalytischer Untersuchungen, eben auch in der Unternehmenskommunikation. An diesem Punkt setzt nun auch der Artikel an bzw. stellt er diese Perspektive in den Mittelpunkt der Betrachtung. Daher sollen in einem ersten Schritt die verschiedenen Arten der Korpusanalyse dargestellt und kurz eingeordnet werden. Die methodische Herangehensweise folgt diesem Abschnitt. Einen weiteren Punkt bilden die Ergebnisse. Dabei werden verschiedene Quellen für Beispiele (DWDS, Webseiten von Unternehmen, Befragungen, Sentimentanalyse, u.a.) zur näheren Betrachtung genutzt. Nach dieser Darstellung erfolgt die Ergebnissicherung, bevor die Herausforderungen in Bezug auf digitale Kommunikate kurz beleuchtet werden. Der vorletzte Teil des Artikels geht außerdem auf den didaktischen Transfer ein. Gerade für die Perspektive der angewandten Wissenschaften ist der Transfer neben dem theoretischen Aspekt von großer Bedeutung, so zum Beispiel in Studiengängen, die Betriebswirtschaft und Sprache kombinieren, den Language and Business Studies. Studierenden aufzuzeigen, welche Mechanismen den jeweiligen Ausprägungen von Wirtschafts- und Unternehmenskommunikation zugrunde liegen, kann unter anderem durch die Korpuslinguistik erreicht werden.

Crosthwaite, Ningrum und Schweinberger (2023) demonstrieren in ihrer Untersuchung "Research trends in corpus linguistics", welche Themen mit zunehmender oder abnehmender Relevanz bzw. Einschlägigkeit in der Korpuslinguistik thematisiert werden, siehe Abbildung 1. So zeigt sich

im diachronen Vergleich eine erkennbare Tendenz zur differenzierten Betrachtung der L2 (Zweitsprache) mit dem Fokus auf Mehrsprachigkeit, zur Grammatik sowie Übersetzung. Ebenso zeigen sich Forschungsinteressen zu „lexical bundles“ (vgl dazu O’Flynn 2022) sowie zu „language contact“ (Matras 2009).

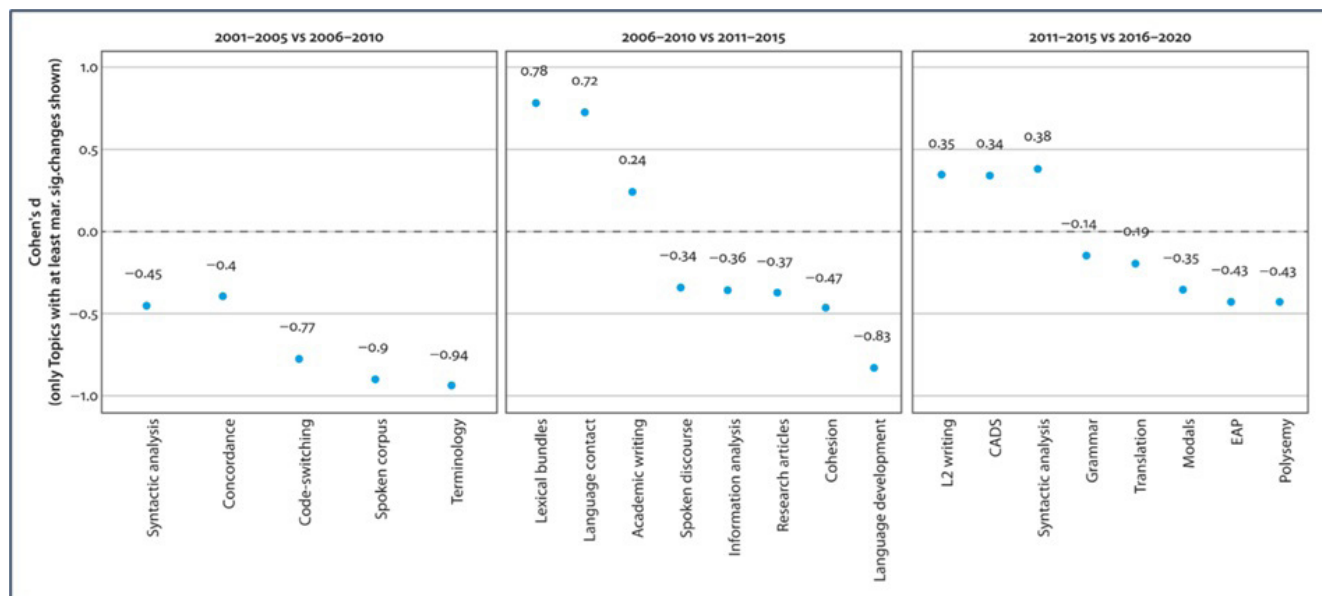


Abbildung 1: Themen mit zunehmender/abnehmender Bedeutung in der CL-Forschung bezogen auf die jeweilige Zeitspanne (Crosthwaite, Ningrum & Schweinberger 2023).

In Abbildung 2 ist dazu ein gemeinsames Vorkommen von Keywords statistisch dargestellt, ebenfalls bei Crosthwaite, Ningrum und Schweinberger (2023). Ersichtlich ist, dass zum Thema „Corpus Research“ Begriffe wie *register*, *grammar frequency*, *spoken discourse*, aber auch *research articles*, *metaphor* oder andererseits *terminology* und *political discourse* zu finden sind. Diese Einträge bilden somit auch eine Tendenz innerhalb der Forschungsdisziplin der Korpuslinguistik ab. Die jeweilige Bedeutung der entsprechenden Keywords wird dabei durch die Größe der Kreise dargestellt.



Nachdem der grundlegende Themenaufriss dargestellt wurde, soll nun die Herangehensweise dieser Überblicksdarstellung aufgezeigt werden. Diese besteht aus einer Bestandsaufnahme sowie einer Eingrenzung zu Korpuswerkzeugen, um das große Feld der vorhandenen Anwendungen und Softwaretools etwas limitieren zu können.

In der Literaturlauswertung wird außerdem vorrangig auf Studien zum Thema Korpusanalyse im Bereich digitaler/multimodaler Unternehmenskommunikation eingegangen. Dabei geht es um die Evaluation der Forschungsfelder sowie Herausforderungen im sprachübergreifenden Bereich. Ein besonderer Fokus in der Betrachtung gilt hierbei der Virtualität, der Digitalität und der Künstlichen Intelligenz (KI) als (neue) Variablen der Korpusanalyse. Dabei wird zum einen das DWDS in seiner Korpusfunktion betrachtet. Ein weiteres Beispiel bildet die Sentimentanalyse als Werkzeug des Social Listening. In einem dritten Beispiel rücken dann mit SNCF, L'Oréal und Peugeot drei ausgewählte französische Firmen in den Mittelpunkt der Beschreibung von Trends in der Analyse zur jeweiligen

Unternehmenskommunikation. Außerdem werden ein ingenieurwissenschaftliches Korpus sowie semantische Netzwerke im Sinne der angewandten Korpuslinguistik betrachtet. Im letzten Teil der Untersuchung werden didaktische Schlussfolgerungen für Module im Bereich Unternehmenssprache und Wirtschaftskommunikation gezogen.

2.1. Anwendungs- und Softwarelösungen für die Korpusanalyse

Eine erste Beurteilung hinsichtlich der Methode der Untersuchung ist in der Frage zu sehen, welche Software für die Analyse genutzt werden sollte und welche Eigenschaften dabei im Fokus stehen. So wäre hier primär das Genre einzugrenzen, wie etwa mündliche, schriftliche oder multimodale Kommunikation, multimodale Unternehmenskommunikation, multimodale Kommunikate, Videoformate (audio-visuell), Social-Media-Posts, Plattformen oder statistische Erhebungen. Bei der Frage, welche Eigenschaften der sprachlichen Daten im Fokus stehen, ist die linguistische Teildisziplin zu präzisieren, z.B. Phonetik, Morphologie, Syntax, Semantik bzw. die Analysekatgorie wie Frequenz, Keywords, n-Gramme, textlinguistische, pragmalinguistische Fragen oder soziolinguistische Variablen. Wie das folgende Zitat illustriert, ist die Liste der Werkzeuge eine sich stets entwickelnde, was auch durch den ‚Aufruf‘ seitens der Autoren der Corpus-Tool-Webseite verdeutlicht wird. So kommen kontinuierlich neue Werkzeuge/ Werkzeugvariationen (Open Source oder zu bezahlende) hinzu und erlauben damit differenziertere Betrachtungen.

Tools for Corpus Linguistics - A hopefully comprehensive list of currently 277 tools used in corpus compilation and analysis. This list is kept up to date by its users. Hence, please feel free to contribute by suggesting new tools. You can also make suggestions, e.g., corrections, regarding individual tools by clicking the  symbol. As this is a non-commercial side [...] project, checking and incorporating updates usually takes some time. (CorpAn 2023)

Im vorangegangenen Zitat wird außerdem deutlich, wie stark auch die Reflexionen innerhalb der linguistischen Disziplinen sind, erfolgt doch die

Aufforderung neue oder modifizierte Tools zu teilen, und diese damit einem Fachpublikum zugänglich zu machen.

Abbildung 3 zeigt zum besseren Verständnis eine Auflistung verschiedener Applikationen. Je nachdem wie die Analyse gestaltet werden soll, kann das entsprechende Werkzeug abgerufen und mit den dazugehörigen Parametern versehen werden. Auch ist zu sagen, dass es in der jeweiligen Kategorie natürlich unterschiedliche Anbieter von Software mit spezifischen Eigenschaften gibt und somit eine Modifikation der Werkzeuge durch die Anwendenden (Betaversionen, Coding, Einbettung bzw. Erweiterungen von Software, etc.) möglich ist.



Abbildung 3: Anwendungen von Korpus-Software (CorpAn 2023).

2.2. KI-Anwendungen im Vergleich zu Korpuswerkzeugen

Es besteht jedoch nicht nur die Möglichkeit der klassischen Korpusanalyse durch bestimmte ohnehin etablierte Werkzeuge. Crosthwaite und Baisa (2023) betonen allerdings im Folgenden:

[...] advantages that corpora and corpus tools still hold over chatbot-like GenAI applications: 1) Knowing the data, 2) Authenticity, 3) Replicability, 4) Multimodality, 5) Safety, 6) Hallucinations, 7) Active vs. passive learning [...].

Dennoch sprechen sie sich nicht gegen eine Nutzung von KI in diesem Zusammenhang aus.

KI kann trotz der gegebenen Herausforderungen (vgl. Gillings *et al.* 2024) als hilfreich in der Analyse von Korpora erscheinen (AntConc@2024), zum Beispiel im Bereich der Big Data. Eine Sammlung KI-generierter Texte kann durchaus auch als möglicher Referenzkorpus dienen und so aufschlussreiche Aspekte in der Untersuchung hinsichtlich semantischer oder stilistischer Fragen aufweisen. Schließlich besteht auch die Möglichkeit der Kombination

von Werkzeugen mit KI-Unterstützung bzw. Einbettung, um Aspekte, die konventionellen Herangehensweisen verschlossen bleiben, zu beleuchten. Crosthwaite und Baisa (2023) geben auch an:

[...] DDL [Data Driven Learning] researchers are well-placed to take advantage of this renewed mainstream interest in language data as we understand the power of such data for language teaching [...].

Das Argument des Transfers im vorangegangenen Zitat ist ein mehr als tragender Aspekt, der jedoch nicht ohne Herausforderungen erscheint, wie an anderer Stelle noch aufgezeigt wird.

3. Trends in der Korpusanalyse – Darstellung der Ergebnisse

3.1. Das Digitale Wörterbuch der deutschen Sprache DWDS – Beispiel Nachhaltigkeit

Betrachtet man nun die Korpusanalyse von der Seite der Terminologearbeit, so kann auch auf schon vorhandene Korpora referiert werden. Das *DWDS* (*Digitales Wörterbuch der deutschen Sprache*) ist eine solche Quelle. Gerade für den Auftakt in der Korpusarbeit ist die Plattform hilfreich, zeigt sie doch, welche Optionen sich in der linguistischen Betrachtung bieten. Auch muss berücksichtigt werden, dass durch Ressourcenknappheit (zum Beispiel Finanzierungsfragen bei spezifischen Projekten) die Erstellung, der Ausbau und die Pflege von Korpora nur bedingt realisiert werden kann, obwohl Analysen auf deren Basis notwendig sind, um Einblicke in eine Vielzahl linguistischer Erscheinungen zu erhalten. Am Beispiel des Begriffs *Nachhaltigkeit* soll der entsprechende Eintrag im *DWDS* nun näher betrachtet werden. Wie direkt unter der Überschrift ersichtlich ist, beginnt der Eintrag (Abbildung 4) mit der klassischen Darstellung zum Begriff, wie Grammatik (Flexion), Aussprache (Phonetik/Phonologie), Wortzerlegung und -bildung. Dem schließt sich die Angabe zum Thesaurus an, mit dem Verweis auf die Synonyme.

Nachhaltigkeit

Grammatik Substantiv (Femininum) · Genitiv Singular: **Nachhaltigkeit** · wird nur im Singular verwendet

Aussprache **ˈnɑːxˌhaltɪˈkʰeɪt**

Worttrennung Nach-hal-tig-keit

Wortzerlegung **ˈnachhaltig** / -keit

Wortbildung mit ›Nachhaltigkeit‹ als Erstglied: **Nachhaltigkeitsfaktor** ... 2 weitere

Thesaurus

Synonymgruppe

Aufrechterhaltbarkeit · Nachhaltigkeit · **Zukunftsfähigkeit** • **Enkeltauglichkeit** ugs.

Typische Verbindungen zu ›Nachhaltigkeit‹ (berechnet)

Detaillierte Informationen bietet das DWDS-Wortprofil zu ›Nachhaltigkeit‹.

Aspekt **Begriff** **Details** **Energieeffizienz** **Fremdwort**

Gedanke **Generationengerechtigkeit** **Gesichtspunkt**

Grundzug **Klimaschutz** **Konzept** **Kriterium** **Leitbild**

Modewort **Prinzip** **Regionalität**

Ressourcenschonung **Sinn** **Thema**

Umweltfreundlichkeit **Umweltschutz**

Umweltverträglichkeit **Wirtschaften** **Zukunftsfähigkeit** **finanzpolitisch** **forstlich**

offensive **Ökologie** **Übergabeprozess** **ökologisch**

Verwendungsbeispiele für ›Nachhaltigkeit‹

maschinell extrahiert aus dem DWDS-Korpus

Von einer **Nachhaltigkeit** in globalem Rahmen sind wir also weit entfernt. [o. A.: Ist Nachhaltigkeit nur eine Worthölze? In: Bild der Wissenschaft auf CD-ROM, Stuttgart: Dt. Verl.-Anst. 2000 [1999]]

Fachleute beschneigen der Regierung inzwischen einen echten Willen zu mehr **Nachhaltigkeit**. [Die Zeit, 19.11.2012, Nr. 47]

So groß war die Chance auf **Nachhaltigkeit** noch nie wie nach diesem »zweiten Olympia«. [Die Zeit, 09.09.2012 (online)]

Wie immer ist auch hier **Nachhaltigkeit** eine Frage des Bewusstseins. [Die Zeit, 09.01.2012, Nr. 02]

Die Sicherheit und **Nachhaltigkeit** dieser Wirkung war ihnen schier unbegreiflich. [May, Karl: Wittenberg IV, Berlin: Neues Leben 1993 [1910], S. 13]

Worthäufigkeit

selten häufig

Wortverlaufskurve

ab 1500 ab 1945

Frequenz / Mio. Texte

Geografische Verteilung

ZDL-Regionalkorpus Webmonitor

Verteilung über Areale

Bitte beachten Sie, dass diese Karten nicht redaktionell, sondern automatisch erstellt sind. Klicken Sie auf die Karte, um in der vergrößerten Ansicht mehr Details zu sehen.

Weitere Wörterbücher

Deutsches Wörterbuch (¹DWB)

Deutsches Wörterbuch, Neubearbeitung (²DWB)

Wörterbuch der deutschen Gegenwartssprache (WDG)

Belege in Korpora

Referenzkorpora

DWDS-Kernkorpus (1900-1999)

Abbildung 4: Screenshot zum Begriff Nachhaltigkeit auf der Webseite des DWDS (2023).

Ein nächster Punkt ist die typische Verwendung des Ausdrucks *Nachhaltigkeit* in Sätzen. Neben diesen Angaben findet sich noch eine Reihe weiterer Details zum Begriff wie die Wortverlaufskurve und die geografische Verbreitung. Auch werden hier mögliche Verwendungsbeispiele angeboten. Diese sind in Abbildung 5 dargestellt und können zum entsprechenden Eintrag im DWDS abgerufen werden.

Korpusbelege ZDL-Regionalkorpus (ab 1993)

Nachhaltigkeit

Korpus:

ZDL-Regionalkorpus (ab 1993)

Start: 1993

Ende: 2023

Areale:

ja

nein

☒ D-Nordwest

☒ D-Mittelwest

☒ D-Südwest

☒ D-Nordost

☒ D-Mittelost

☒ D-Südost

Anzeige:

☐ KWIC

☒ voll

☐ maximal

Sortierung:

Datum absteigend

Treffer pro Seite:

50

Verteilung über Areale:

Verteilung über Zeiten:

Histogrammansicht:

1-50 von 95345 Treffern

←

-10

-5

1

2

3

4

5

→

+5

+10

→

1: Döbelner Allgemeine Zeitung, 19.08.2023

Mit Dominanz und **Nachhaltigkeit**

2: Döbelner Allgemeine Zeitung, 19.08.2023

Nachhaltigkeit ist angesagt

3: Döbelner Allgemeine Zeitung, 19.08.2023

In den Regionalkonferenzen hat Dirk Neubauer mit Unternehmen, Institutionen und Bürgern Leitlinien diskutiert und dabei vier Kernthemen herausgearbeitet: Digitalisierung, Mobilität, **Nachhaltigkeit** und Ermöglichung.

4: Frankfurter Rundschau, 19.08.2023

Apropos Natur und **Nachhaltigkeit**:

5: Frankfurter Rundschau, 19.08.2023

Da werden 600 einzelne Abschnitte der **Nachhaltigkeit** untersucht, und im Jahr 2022 kam die Eintracht dabei auf 66 von 100 Punkten.

6: Frankfurter Rundschau, 19.08.2023

Regenwasserzisternen und Ökostrom zählen zu den Schritten, die die Eintracht auf ihrem Gelände in Richtung **Nachhaltigkeit** geht.

7: Frankfurter Rundschau, 19.08.2023

Gern würde der Klub auch in Sachen Verkehrswege rund ums Stadion einen Schritt in Richtung **Nachhaltigkeit** machen.

8: Frankfurter Rundschau, 19.08.2023

Mit **Nachhaltigkeit** im Handel kennt sich Lukas Glöckner von der "Lust auf besser leben GmbH" aus.

9: Fränkischer Tag, 19.08.2023

Hier kann auf zwei unterschiedliche Weisen **Nachhaltigkeit** betrieben werden:

10: Reutlinger General-Anzeiger, 19.08.2023

Nachhaltigkeit - Reparaturführer des AWB vertreten

11: Hamburger Abendblatt, 19.08.2023

Beim "ParkFunkeln" wird großen Wert auf **Nachhaltigkeit** und Ressourcenschonung gelegt.

12: Landshuter Zeitung, 19.08.2023

"Lärmschutzelemente aus Holzfasern" in einer "mächtigen Betonwanne" sind auch kein Ausweis von "**Nachhaltigkeit**"; der immense Flächenverbrauch des Bauwerks setzte bereits mehr Biomasse gebundenes CO2-frei.

13: Landshuter Zeitung, 19.08.2023

Christine Putz nannte als gegenwärtig vordringlich beachtete Themen MINT, Smart-City, Innenstadtbelebung und **Nachhaltigkeit** und konnte eine Reihe bereits realisierter und angepeilter wohldurchdachter Aktionen bilanzieren.

14: Leipziger Volkszeitung, 19.08.2023

Geografische Verteilung

... über Areale

... über Zeiten

Belege in Korpora

Referenzkorpora

DWDS-Kernkorpus (1900-1999) (27)

DWDS-Kernkorpus 21 (2000-2010) (40)

DTA-Kernkorpus (1598-1913) (48)

Metakorpora

DTA-Kern+Erweit. (1465-1969) (67)

Historische Korpora (1465-1998) (295)

Referenz- und Zeitungskorpora (frei) (3256)

ZDL-Regionalkorpus (ab 1993) (95345)

WebXL (726441)

Zeitungskorpora

Berliner Zeitung (1945-1993) (11)

Berliner Zeitung (1994-2005) (378)

neues deutschland (1946-1990) (10)

Der Tagesspiegel (ab 1996) (6007)

Die ZEIT (1946-2018) (2240)

Abbildung 5: Screenshot zum Begriff Nachhaltigkeit auf der Webseite des DWDS (2023).

Eine weitere Dimension zur Analyse auf der Basis von Korpora bildet die Sentimentanalyse, mit der auf Seiten der Rezipierenden vor allem deren Einstellung und Wahrnehmung erforscht werden sollen. Dass es sich dabei um ein sehr vielfältig nutzbares Werkzeug vor allem im Bereich der Wirtschaftskommunikation handelt, illustrieren die beiden folgenden Belege.

Die sogenannte Sentimentanalyse, das heißt das Erkennen von Stimmung in Texten ist mittlerweile zu einer zentralen Forschungsfrage der Computerlinguistik geworden. [Das damit verbundene] „Opinion Mining“ beschäftigt sich mit dem Bestimmen der Meinung, Empfindung oder der Haltung von Menschen über die Analyse von Texten. (Banholzer 2019: 167)

Anhand von Textquellen aus verschiedenen Plattformen und Medien ermöglicht sie auf Basis der Textanalyse zusammen mit Text Mining die Erstellung eines Stimmungsbildes.

Im Folgenden werden nun exemplarisch verschiedene Aspekte der Sentimentanalyse thematisiert. In Abbildung 6 ist dazu die Mediasentimentanalyse zum Thema „Tee“-Marketing bei Alexa und Siegel (2021) abgebildet. Mit der Bewertung von „positiv“ über „neutral“ zu „negativ“ bis „kein Sentiment“ kann so nachvollzogen werden, welche Meinung, Emotion oder Haltung zum Thema „Tee“ „pro Mediasegment“ (Blogs, Nachrichten, soziale Netzwerke, Microblogs, Web, Foren, Bilder, etc.) vorhanden sind. Die jeweilige farbliche Zuordnung erlaubt anhand ihrer Kategorisierung entsprechende Interpretationen und Rückschlüsse.

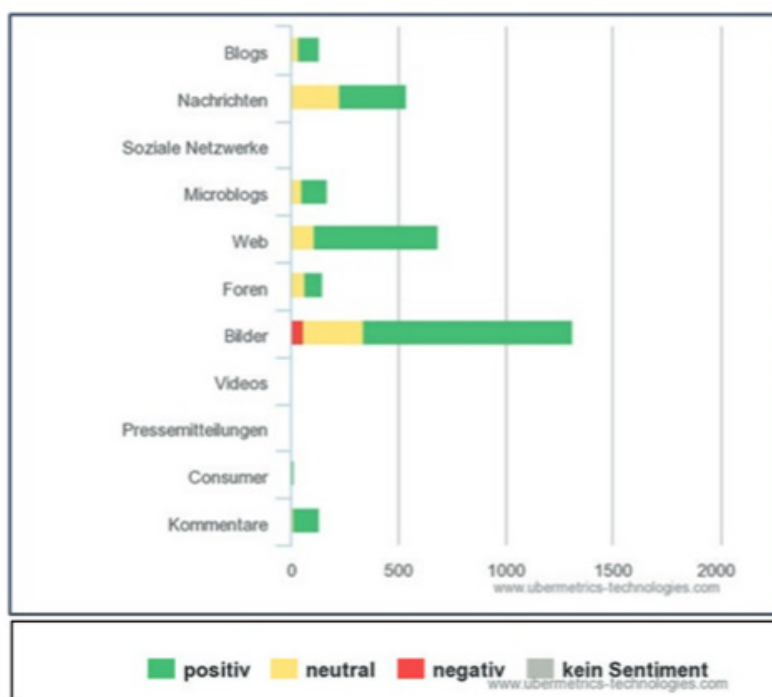


Abbildung 6: Ubermetrics Mediasentimentanalyse für „Tee“ Sentimentaufteilung pro Media-segment (Alexa & Siegel 2021: 134, Darstellung um Legende durch d. Verf. ergänzt).

In Abbildung 7 wird eine weitere Möglichkeit der Visualisierung von Ergebnissen der Sentimentanalyse dargestellt. Dabei geht es wiederum um das Vorkommen des Begriffs „Tee“, in diesem Fall allerdings im Bereich der Suchanfragen. Auffallend ist die überaus positive Wahrnehmung bzw. Bewertung von „Tee“ in diesem Zusammenhang. Damit lassen sich für eine marketingbezogene Erhebung Ergebnisse für ein vorhandenes oder wahrgenommenes Sentiment erkennen und zur anschließenden Verwendung aufbereiten.

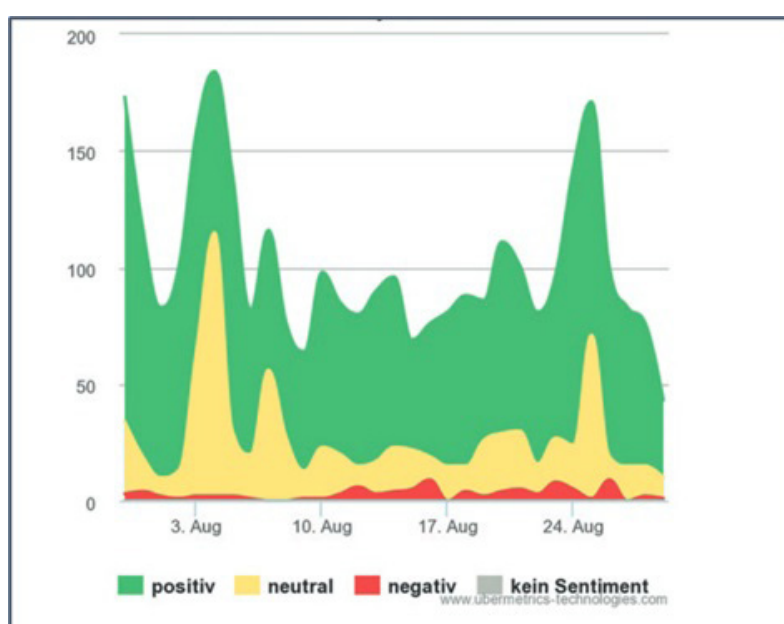


Abbildung 7: Ubermetrics Sentimentanalyse für „Tee“ (Alexa & Siegel 2021: 134).

Wie die beiden Diagramme in Abbildung 6 und 7 zeigen, können durch die Erkenntnisse aus der (Korpus)Linguistik für die Wirtschaftskommunikation Ergebnisse zum Beispiel im Marketing oder in der Unternehmenskommunikation gewonnen werden, die der besseren Umsetzung von Marketingideen und -konzepten dienen. In einem weiteren Schritt ließen sich anhand der Erhebung wie in den Abbildungen Schlussfolgerungen zum Thema „Tee“ ziehen: „Wie wird das Produkt wahrgenommen?“ oder „Welche Vorbehalte gibt es?“. Auch würde eine Korpusanalyse die Möglichkeit bieten durch eine quantitative Untersuchung der Häufigkeit des Begriffs „Tee“ sowie beispielsweise seiner Kollokationen weitere Rückschlüsse auf die Verwendung des Wortes in bestimmten Kontexten (Werbung) zu ziehen.

3.3. Französischsprachige Korpora und ihre Perspektiven

Ein Blick in die französischsprachigen Korpora soll nun einen weiteren Trend bezüglich der untersuchten Thematik – hier französische Gebärdensprache – aufzeigen. Im Gegensatz zur englischsprachigen Linguistik, in der in den letzten Jahren stark die akademische Sprache als Untersuchungsgegenstand diente, zeigt sich in der französischen Linguistik ein anderes Bild. Hier ist der Fokus eher im Bereich der Literatur zu finden, beziehungsweise auch in alternativen Themenfeldern. Zu nennen wäre hier das Korpus Mediapi-**RGB** (2017-2022), in dem es um Gebärdensprache in einem kontrastiven (bilingualen) multimodalen Zusammenhang geht.

Mediapi-RGB**** is a bilingual corpus of French Sign Language (LSF) and written French in the form of subtitled videos, accompanied by complementary data (various representations, segmentation, vocabulary, etc.). It can be used in academic research for a wide range of tasks, such as training or evaluating sign language (SL) extraction, recognition or translation models.

To build this corpus, we used videos from **Média'Pi!**, a bilingual online media with journalistic-type content in LSF with French subtitles. We collected 1230 videos dating from September 2017 to January 2022, representing a total of 86h. Based on the subtitles, we temporally segmented the videos into 50084 video segments (or extracts). We also automatically cropped the signer and harmonised the segments in terms of size (444x444) and frequency (25fps). (Mediapi-**RGB** (2017-2022) (Ortolang 2023))

Dies unterstreicht auch den interkulturellen Aspekt der Korpuslinguistik – die französische Gebärdensprache und die damit verbundene kulturelle Verknüpfung in die entsprechende Gehörlosengemeinschaft – sowie ihre thematische Vielfältigkeit. Mögliche Einblicke in die verschiedenen Themen französischsprachiger Korpora lassen sich zum Beispiel über Ortolang (2023) erhalten.

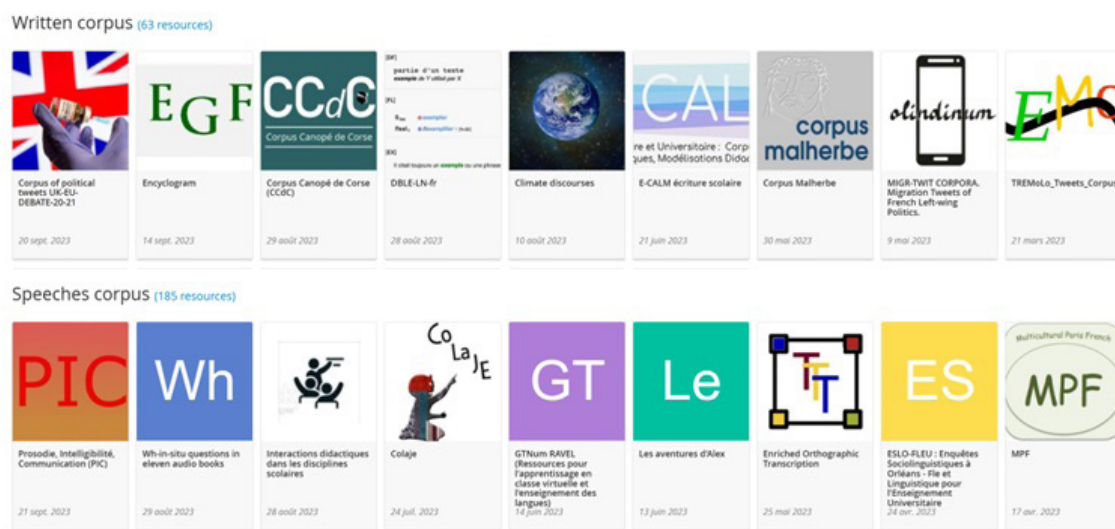


Abbildung 8: Screenshot Französischsprachige Korpora (Auswahl) (Ortolang 2023).

3.4. Unternehmenskommunikation als Datengrundlage: „Kundeninformation der SNCF“

Im Folgenden soll nun ein Beispiel der digitalen Unternehmenskommunikation der SNCF (Société Nationale des Chemins de Fer Français)² anhand ihrer Kundeninformation „En Coulisses“ (Blick hinter die Kulissen des Unternehmens) als Datengrundlage für eine potenzielle Korpusanalyse vorgestellt werden (siehe auch Große (2009) zur multimodalen linguistischen Analyse). Wie in Abbildung 9 ersichtlich ist, besteht die Kundeninformation „En Coulisses“ der SNCF, zum Verhalten und Verfahren bei der Beförderung von Passagieren mit Schienenfahrzeugen, jeweils aus einem Bild mit Überschrift und einem kurzen einleitenden Text mit den jeweiligen Informationen. An der unteren Seite dieser Texteinheit befindet sich als anklickbarer Hyperlink außerdem der Hinweis „*En savoir plus*“ (dt.: „Mehr erfahren“ [Übers. d. Verf.]). Über diesen gelangen Rezipient:innen zu weiteren spezifischeren Informationen zum ausgewählten Thema, wie „Accident de personne, Fortes pluies et inondations, Bagage abandonné, Feuilles mortes

2 dt.: Nationale Gesellschaft der französischen Eisenbahnen.

et problèmes d’adhérence, Fraude et usurpation d’identité, Adapter la circulation de vos trains“ (SNCF 2023).³

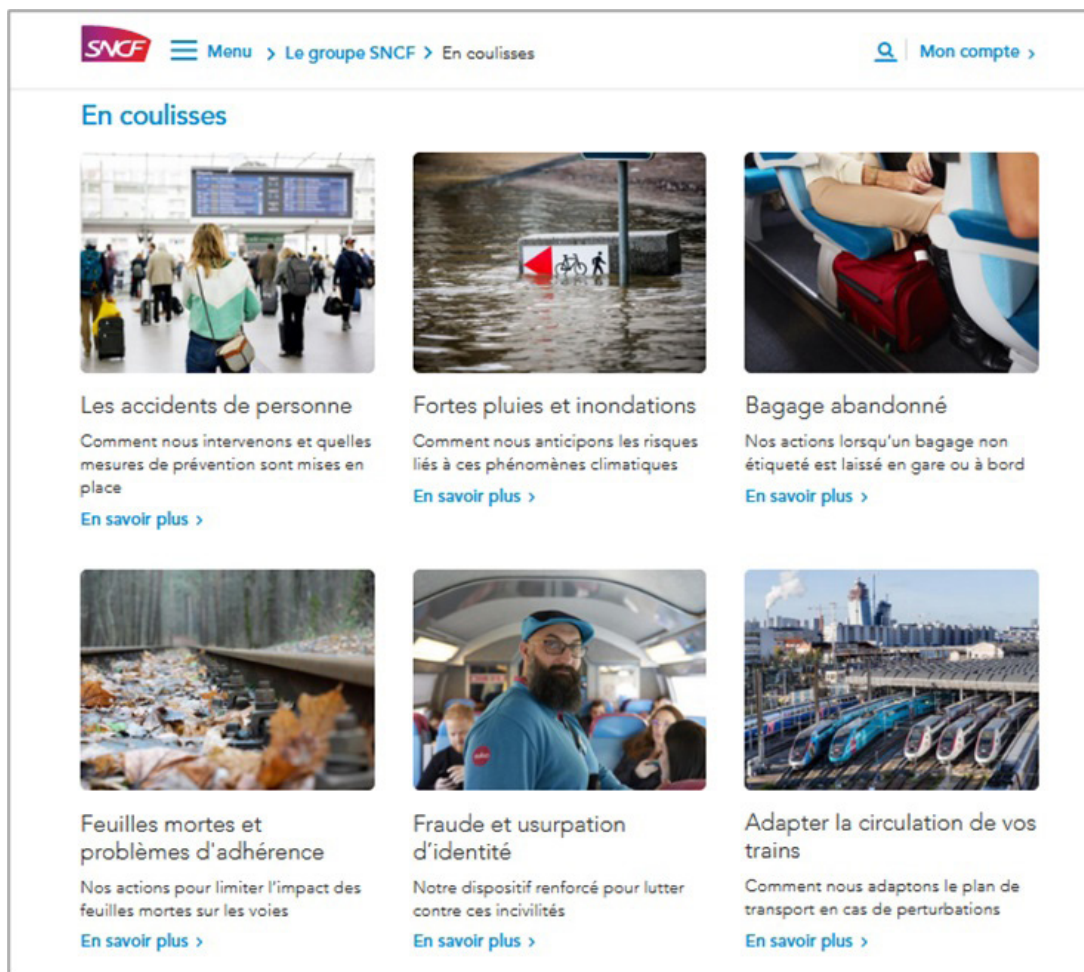


Abbildung 9: Screenshot der SNCF zum Thema «En Coulisses» (SNCF 2023).

Für eine korpusbasierte Untersuchung wären neben den Bildunterschriften auch die Texte unter dem Link „*En savoir plus*“ von besonderem Interesse. Sie könnten gesammelt, annotiert und entsprechend archiviert werden, um dann mit den jeweiligen Werkzeugen wie AntConc®(2024)⁴ auf sie zugreifen und Fragen hinsichtlich Struktur, Schlüsselwörtern, Kollokationen oder morphologischen Besonderheiten bzw. Markierungen klären zu können. Auch die Frage, wie die SNCF mit eher schwierigen Themen wie zum

³ „Personenunfall, Starkregen und Überschwemmungen, verlassenes bzw vergessenes Gepäck, Laub und Haftungsprobleme, Betrug und Identitätsdiebstahl, Anpassung des Verkehrs Ihrer Züge“ [Übers. d. Verf. unter Verwendung von DeepL] (SNCF 2023)

⁴ AntConc®(2024): <https://www.laurenceanthony.net/software/antconc/> (2024)

Beispiel Personenunfällen im Bereich des Schienenverkehrs oder auch Identitätsbetrug bei Kund:innen kommunikativ verfährt, wäre anhand dieser Beispiele als Forschungsfeld bearbeitbar.

Für Unternehmen wie die SNCF oder die DB (Deutsche Bahn) ist diese Kommunikation – wie der „Blick hinter die Kulissen“ – bedeutungsvoll, da sie Kommunikationsräume u.a. für Mitarbeitende und Kund:innen eröffnet. Die Modalitäten der Kommunikation der Unternehmensinhalte bzw. -kultur zeigen sich damit als zentral für die Wahrnehmung des Unternehmens nach außen. Darüber hinaus ergeben sich aber auch Einblicke in die Abläufe und Verfahrensweisen einer multimodalen Unternehmenskommunikation des jeweiligen Beförderungsunternehmens für Forschende. Auch lassen sich durch die Verwendung der angesprochenen Analyseverfahren die Veränderungen in der strategischen Kommunikation des Unternehmens verdeutlichen.

3.5. Unternehmenskommunikation als Datengrundlage: Nachhaltigkeit bei L'Oréal

Neben den Webseiten der Beförderungsunternehmen im vergangenen Beispiel lassen sich bezüglich der Wirtschaftskommunikation auch andere Dokumente von Unternehmen untersuchen. So wendet sich beispielsweise Kesselheim (2011) in seiner Analyse den Aktionärsbriefen (als Teil der Geschäftsberichte) zu. Anhand einer Korpusanalyse in – relativ zum Vorkommen – kleinem Umfang zeigt er exemplarisch die Aspekte, die bei einer Textanalyse von Bedeutung erscheinen, nämlich „Begrenzbarkeit“, „intratextuelle Verknüpfbarkeit“, „thematische Zusammengehörigkeit“, „pragmatische Nützlichkeit“, „intertextuelle Beziehbarkeit“ und „Musterhaftigkeit“ (Kesselheim 2011: 169).

Doch lassen sich neben den Geschäftsbriefen auch weitere Genres der Unternehmenskommunikation betrachten. Eine sich anreihende Textsorte im Rahmen dieser Untersuchung ist der Nachhaltigkeitsbericht. Am Beispiel der Firma L'Oréal lassen sich gut typische Merkmale skizzieren.

Die Abbildung 10 zeigt einen Auszug aus einem Text dieser Publikation. Die hervorgehobenen Sätze und Phrasen im Textauszug des Nachhaltigkeitsberichtes sollen mögliche Analyseansätze unterstreichen.

Zentral ist der Begriff „engagements“ (dt.: „Bekenntnis, Einsatz, Verpflichtung“ [Übers. d. Verf.]), der die Bemühungen des Unternehmens hinsichtlich seiner Nachhaltigkeitsbestrebungen zusammenfasst.

L'Oréal pour le Futur, nos engagements environnementaux pour 2030



Le réchauffement climatique et les changements environnementaux entraîneront une dégradation potentiellement définitive des habitats naturels et de nos conditions de vie. L'augmentation du niveau de la mer, la fonte des glaces, le réchauffement et l'acidification des océans ainsi que les événements météorologiques extrêmes vont s'accélérer. Pour répondre à des enjeux plus importants, il faut des engagements plus forts.

C'est pourquoi L'Oréal a décidé de renforcer ses efforts à travers un programme appelé L'Oréal pour le Futur. Nos engagements pour 2030 marquent le début d'une transformation plus radicale et incarnent notre vision du rôle et des responsabilités des entreprises pour relever les défis auxquels le monde fait face.

Téléchargez la brochure L'Oréal pour le Futur

Abbildung 10: L'Oréal nos engagements (L'Oréal 2023).

Die in der Abbildung gesetzten Hervorhebungen bilden zu einem großen Teil die Schlagwörter und Ausdrücke ab, die als typisch für diese Textsorte bezeichnet werden können.

Le réchauffement climatique et les changements environnementaux entraîneront une dégradation potentiellement définitive des habitats naturels et de nos conditions de vie. L'augmentation du niveau de la mer, la fonte des glaces, le réchauffement et l'acidification des océans ainsi que les événements météorologiques extrêmes vont s'accélérer. Pour répondre à ces enjeux plus pressants, il faut des engagements plus forts. (L'Oréal 2023)⁵

5 „L'Oréal für die Zukunft, unser Umweltbekenntnis für 2030 – Die globale Erwärmung und die Umweltveränderungen werden zu einer potenziell dauerhaften Verschlechterung der natürlichen Lebensräume und unserer Lebensbedingungen führen. Der Anstieg des Meeresspiegels, die Eisschmelze, die Erwärmung und

So bezieht sich das Unternehmen mit der Phrase „*réchauffement climatique*“ auf die globale Erwärmung und skizziert im weiteren Verlauf des Textes die sich verändernden Umweltbedingungen „*une dégradation potentiellement définitive des habitats naturels et de nos conditions de vie*“. Außerdem werden extreme Wetterlagen in „*événements météorologiques extrêmes*“ angesprochen und mit „*il faut des engagements plus forts*“ die Notwendigkeit zu einer verstärkten Einlassung des Unternehmens hinsichtlich nachhaltiger Vorgehensweisen unterstrichen.

C'est pourquoi L'Oréal a décidé de renforcer ses efforts à travers un programme appelé L'Oréal pour le Futur. Nos engagements pour 2030 marquent le début d'une transformation plus radicale et incarnent notre vision du rôle et des responsabilités des entreprises pour relever les défis auxquels le monde fait face. (L'Oréal 2023)⁶

Im zweiten Teil des Textes zeigt sich die Perspektive nun ganz fokussiert auf den expliziten Beitrag des Unternehmens zur Verbesserung der Umweltbedingungen „*C'est pourquoi L'Oréal a décidé de renforcer ses efforts*“. In der Darstellung ihres Einsatzes bezeichnet L'Oréal seine Einlassung bis 2030 als radikale Transformation „*2030 marquent le début d'une transformation plus radicale*“. In Verbindung mit dem Video, das ebenfalls in der Abbildung sichtbar und auf der Webseite eingebettet ist, unterstützt es den multimodalen Charakter der Textbotschaft.

Im Rahmen dieses Artikels sind die getroffenen Aussagen dennoch nur Impulse. Weitere korpuslinguistische Untersuchungen an diesem Text könnten im entsprechenden Umfang sicher gut und fundiert die mögliche

Versauerung der Ozeane sowie extreme Wetterereignisse werden sich beschleunigen. Um diesen drängenderen Herausforderungen zu begegnen, bedarf es stärkerer Verpflichtungen. [Übers. d. Verf. unter Verwendung von DeepL] (L'Oréal 2023)

- 6 Aus diesem Grund beschloss L'Oréal, seine Anstrengungen durch ein Programm namens „L'Oréal für die Zukunft“ zu verstärken. Unsere Verpflichtungen bis 2030 markieren den Beginn einer radikaleren Transformation und verkörpern unsere Vision von der Rolle des Unternehmens bei der Bewältigung der Herausforderungen, vor denen die Welt steht.“ [Verf. unter Verwendung von DeepL] (L'Oréal 2023)

Verwendung bestimmter Keywords oder Kollokationen akzentuieren sowie textkompositorische Fragen in der Unternehmenskommunikation aufwerfen. Aber auch eine Sentimentanalyse dieser Art von Unternehmenskommunikation würde hier sicher zahlenbasierte bzw. statistische Einblicke erlauben und könnte durch das daraus entstehende Stimmungsbild Textentscheidungen (Lokalisation oder Grad an Multimodalität der Botschaften und Textgestaltung) seitens des Unternehmens reflektieren. Ein weiteres mögliches Instrument zur Erhebung von derartigen Daten ist das Textmining, ausgeführt durch Werkzeuge wie Rapidminer, Lexalytics Intelligence Platform oder Google Natural Language API (Stadler 2024).

Abbildung 11 zeigt die englische Version des Textes aus Abbildung 10, die ebenfalls auf der Webseite von L'Oréal zu finden ist. Auch bei dieser Textversion finden sich die unter Abbildung 10 schon angesprochenen Diskurse zu Klimawandel „*Global warming is accompanied by environmental changes*“, die Notwendigkeit zur Intervention „*with higher stakes must come stronger commitments*“, Nachhaltigkeit „*L'Oréal for the Future*“ und das Schaffen eines Umweltbewusstseins des Unternehmens „*company's vision, purpose and responsibilities*“. Durch die multilinguale Verfügbarkeit der Texte (Französisch, Englisch und Chinesisch) können diese für die linguistische Analyse neben einer möglichen kontrastiven Betrachtung auch hinsichtlich multimodaler Aspekte genutzt werden. Besonders für korpuslinguistische Perspektiven bleibt der direkte Vergleich im Wording lohnenswert, also spezifische Aspekte der Analyse wie Keywords und auch Kollokationen. Da ein erster Zugang auch schon über verfügbare Suchmasken bzw. Werkzeuge zum Beispiel im Acrobat Reader oder auch im Foxit Reader möglich ist, kann diese Herangehensweise gerade für die Studierenden in der ersten Begegnung mit den Methoden der Korpuslinguistik als gewinnbringend erachtet werden. Dabei zeigt das Verwenden solcher Suchanfragen bereits Ergebnisse zur Frequenz eines bestimmten Begriffes oder den Positionen innerhalb des Textes auf.



Abbildung 11: L'Oréal Commitments and Responsibilities (L'Oréal 2023).

3.6. GINGKO-Projekt⁷

Neben der wirtschaftsbezogenen Dimension der korpuslinguistischen Analyse kann auch die ingenieursprachliche Perspektive interessante Einblicke bieten.

Das Projekt hat das Ziel verfolgt, Muster in den Fachsprache(n) der Ingenieurwissenschaften am Beispiel der Automobiltechnik systematisch zu erfassen und linguistisch zu beschreiben. [...] (GINGKO 2023)

So konnte das Projekt GINGKO (deutschsprachig) an der Universität Leipzig eine ebensolche Erfassung der ingenieursprachlichen Begriffe durch Sammlung und Evaluierung erreichen.

[...] Die Grundlage dafür bildet das Korpus Gingko (Geschriebenes ingenieurwissenschaftliches Korpus), das 2.498 wissenschaftliche Artikel aus zwei der renommiertesten deutschsprachigen Fachzeitschriften im Bereich der Automobiltechnik umfasst. Die

Ergebnisse des DFG-geförderten Projekts wurden für die Vermittlung des Deutschen als fremde Wissenschaftssprache nutzbar gemacht. (GINGKO 2023)

Ergebnisse, die sich auf solche Korpora beziehen, können eine Vielzahl von Erkenntnissen zur Sprachverwendung und -struktur bei Anwender:innen aus den Bereichen Deutsch als Fremdsprache, Unternehmenskommunikation und im Bereich des Automobilmarketing generieren. Die erlangten Erkenntnisse können Strukturen von Fachsprache, wie Idiomatisierungen, Nominalisierungen, Verbalisierungen oder Verb-Substantiv-Kombinationen verdeutlichen. Gerade für Studierende in den ingenieurwissenschaftlichen Studiengängen ist die Visualisierung der Ergebnisse dieser Fachsprachenbetrachtung von besonderem Gewinn, weil die Ingenieursprache durch dieses Projekt transparenter und sowohl für Studierende (mit linguistischem oder ingenieurwissenschaftlichem Bezug) als auch Dozent:innen (mit linguistischem oder ingenieurwissenschaftlichem Bezug) zugänglich gemacht wurde.

3.7. Semantische Netze

Korpuslinguistische Analysen, oder zumindest die Datenbasis, die sie schaffen, können unter anderem auch für das Erstellen semantischer Netze genutzt werden. Jiang et al. (2023) zeigten in ihrer Untersuchung zwischen “In-Person Positions” (also Stellen, an denen Arbeitnehmende von/in einem Büro in einer Einrichtung arbeiten) und “Remote Positions” (also Stellen, an denen Arbeitnehmende von zuhause oder einem anderen Platz als dem Büro arbeiten), wie die Beschreibung der beruflichen Tätigkeit jeweils variiert (vergleiche dazu Abbildung 12 und Abbildung 13). Semantische Netze können im einerseits externen, andererseits internen Teil der Unternehmenskommunikation wie Stellenausschreibungen durch ihre Darstellung von Zusammenhängen oder der Visualisierung von Bedeutungsinhalten für Entscheidungstragende wichtige Impulse – im Sinne einer Optimierung – zum Verständnis von schriftlicher und mündlicher Kommunikation im Unternehmen geben.



Im folgenden Abschnitt soll die Herangehensweise von Jiang et al. (2023) nochmals verdeutlicht werden. Die in der korpusbasierten Untersuchung vorkommende Textanalyse wird hier auf die angewandte Semantik übertragen.

Using text mining and semantic network analysis, this study analyzed the job descriptions of 34,787 positions about media analytics from Indeed.com to compare how the in-person and remote jobs differ to inform educators about integrating analytics in the media and communications curriculum. We found that the in-person positions emphasized more on the skills of verbal, interpersonal, and organizational communication, whereas the remote positions asked more for written communication. While the in-person positions had higher expectations of using general data management and analysis tools, the remote positions emphasized more on the use of social media analytics and digital marketing tools. (Jiang *et al.* 2023)

Es zeigt sich außerdem durch die Kombination der Methoden aus Textmining und semantischen Netzen, dass bei der Beschreibung der ausgeschriebenen Positionen der Fokus je nach dem Inhalt der Stellenbeschreibung entweder mehr auf den mündlichen („in-person“) oder den schriftlichen Fähigkeiten („remote position“) lag. Auch war der Umgang mit der technischen Komponente, wie Datenmanagement oder Analysetools bei der „in-person“-Stelle wichtiger. Bei der „remote“-Stelle war hingegen der Gebrauch von Social Media und Marketingwerkzeugen von größerer Bedeutung.

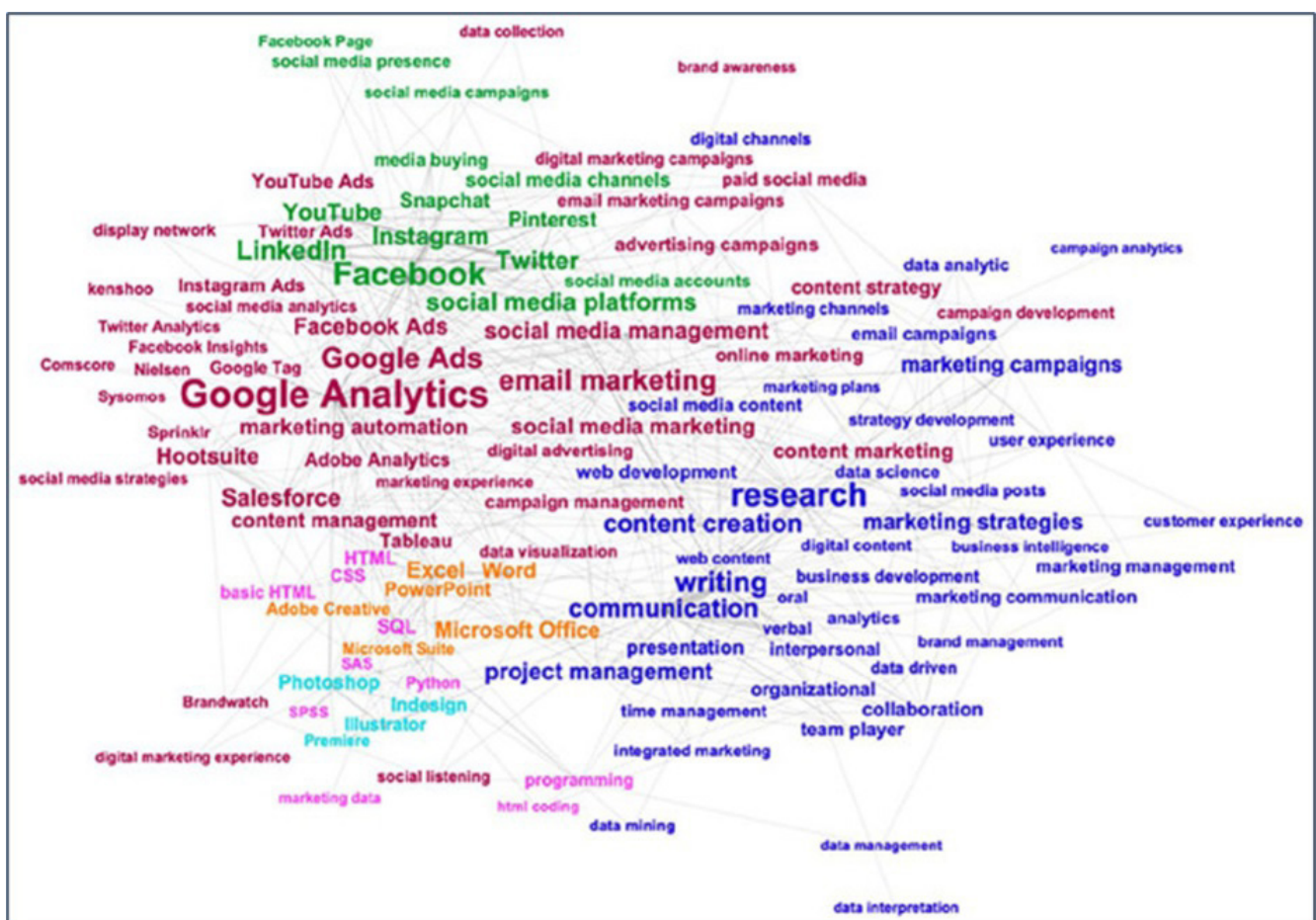


Abbildung 13: Semantisches Netz der Fähigkeiten und Werkzeuge in Media Analytics Remote Stellenbeschreibungen. (Jiang et. al 2023).

3.8. Webseite – Peugeot

Nach den besprochenen Beispielen mit Bezug auf Kosmetika, ingenieurwissenschaftlichem Inhalt und Mitarbeitendenakquise schließt sich nun ein weiteres aus dem Bereich der Technik an. Auf der Website von Peugeot lassen sich besonders zum Thema E-Mobilität interessante multimodale Text-Bild-

Kompositionen in der externen Unternehmenskommunikation entdecken. Dabei geht bereits die Überschrift der Auftaktseite in *medias res*, indem sie den Slogan „*Tout est plus beau avec allure*“ („Alles ist schöner mit Schwung“ (Übers. d. Verf.)) zum Hauptthema macht und Mobilität mit elektrischem Antrieb idealisiert. Mit dem näheren Betrachten des Slogans „*100% électrique, 100% irrésistible*“ („100% elektrisch, 100% unwiderstehlich“ (Übers. d. Verf.)) rundet sich dieser Eindruck einer äußerst stilvollen Perspektive auf Elektromobilität ab, die sich auch durch die entsprechende Visualisierung – Elektroautos als Designobjekte – auf der Webseite verdeutlicht. Die Texte der Website haben hier zwar nur einen geringen Umfang, doch kann mit ihnen im Vergleich in Form einer korpusbasierten Diskursanalyse dennoch ein vorläufiges Interpretationsergebnis erzielt werden. Der Aspekt des Umfangs der Textsammlung als Parameter bzw. Merkmal der korpuslinguistischen Analyse ist damit nicht weniger aussagekräftig. Allerdings zeigt sich hier auch eine Veränderung der Unternehmenskommunikation an sich. So sind Texte auf ebensolchen Webseiten von eher geringerem Umfang, was den Zugang für Kund:innen sehr erleichtert. Um dennoch eine größere Datenbasis erhalten zu können, eignen sich Referenztexte aus Webseiten anderer Autohersteller wie Renault, um mögliche Schlussfolgerungen hinsichtlich vorhandener Diskurse (wie zum Beispiel Vorteile der Elektromobilität oder Ausstattungsunterschiede der einzelnen Fahrzeugmodelle) verifizieren können.

3.9. Das IVO-Projekt als Beispiel für eine multimodale Korpusanalyse

Neben den bisher angesprochenen Beispielen für potenzielle Analysen soll nun der Blick auf ein sehr komplexes und auch multimodal angesiedeltes Projekt gelegt werden. Das *IVO-Projekt*⁸ – zwischen der Universität Cardiff und dem Mary Immaculate College (Limerick) als Projektleitende – betrachtet arbeitsplatzbezogene Diskurse in digitalen Umgebungen, wie Sitzungen von Gremien, Kulturorganisationen oder auch Bildungseinrichtungen. Die Analyse zum Erkenntnisgewinn für die Digital Humanities soll im Weiteren auch die Anwendung von neuen Methoden gerade in diesem Bereich

8 Projektwebseite: <https://ivohub.com/> (aufgerufen 21.09.24).

unterstützen und dabei deren praktische Relevanz für die pragmatische Dimension nachhaltig betonen. Durch die Untersuchung zur digitalen Kommunikation zum Beispiel in öffentlichen Einrichtungen wie dem City Council und den damit verbundenen aufkommenden Barrieren trägt das Projekt zum Abbau von Kommunikationshürden bei. Die Ergebnisse können Abläufe in öffentlichen Einrichtungen möglicherweise sogar verbessern und dabei noch weitere Aspekte aufdecken, die anschließende bzw. weiterführende Studien dafür initiieren.

4. Didaktische Implikationen

Adaptierte Anwendungen der korpuslinguistischen Analyseprinzipien und Methoden zum Beispiel im Bereich Wirtschaftskommunikation können die Terminologiarbeit von Lehrenden und Studierenden maßgebend unterstützen. Eine Vertiefung zur Thematik der Datenbank – u.a. begleitet von Methoden der Termextraktion aus Texten und systematischer Kategorisierung – im Rahmen des Studiengangformates Language and Business Administration oder Wirtschaftskommunikation Deutsch ist naheliegend. Gerade bei der Untersuchung der Wirtschaftskommunikation zeigen sich diese Herangehensweisen als zielführend. Dies ist nicht auf eine Sprache beschränkt und kann somit auch mehrsprachig bzw. kontrastiv durchgeführt werden. Ein Beginner's Guide zu Data Science und Data Mining eröffnet für die Studierenden weitere Räume, um die Komplexität sprachlicher Strukturen noch besser erfassen zu können und Prinzipien in der Verwendung zu erkennen. Den Umgang mit Werkzeugen der Korpuslinguistik bei den Studierenden zu stärken, führt so zum besseren Verständnis von Texten und ihren Mechanismen, besonders im Bereich Wirtschaft. Hier lassen sich außerdem digitale und virtuelle Kommunikate hin zur Diskussion von Multimodalität einbinden. Um den Studierenden einen ersten Einblick zu ermöglichen, sind Tutorials wie z. B. von Bubenhofer, Volk und Clematide ein guter Ausgangspunkt.⁹

9 Tutorialsseite: <https://www.bubenhofer.com/korpuslinguistik/kurs/> (aufgerufen 20.12.24)

Auch die Anwendung der Methoden auf Texte der Unternehmenskommunikation kann schrittweise erfolgen, indem Studierende erste kleine Korpora anlegen und nach einer Einführung zu den Optionen der Evaluation diese kleinschrittig untersuchen. Ein für die Studierenden ansprechendes Genre bildet dabei die Repräsentanz von Unternehmen (wie Volkswagen, Adidas, Ikea oder SNCF) in Sozialen Medien, die mit einfachen Mitteln hinsichtlich des Aufbaus der Profile und interlingualen Vergleichen näher betrachtet werden können. Die Ergebnisse festzuhalten und zu kategorisieren wäre ein nächster Schritt, der zugleich der Schreibpraxis dienlich wäre.

Je nach Studiengangsprofil können Studierende, unter fachlicher Begleitung durch die Lehrenden, fachspezifische Fragen an Textsammlungen stellen und eigenen Forschungsinteressen folgen.

5. Diskussion und Limitationen der Betrachtung

Nachdem im vorangegangenen Teil die verschiedenen Beispiele vorgestellt und evaluiert wurden, soll nun die Diskussion anschließen. Dabei kann geschlussfolgert werden, dass es sich im Fall der Korpuslinguistik um ein sehr relevantes und sich stetig weiterzuentwickelndes Feld handelt, das wiederholt neue Technologien integriert. So können durch diese Art der linguistischen Analyse Rückschlüsse auf die aktuelle Verwendung, den Sprachwandel und durchaus auch den Bedeutungswandel bzw. semantische Verschiebungen gezogen werden. Die Ergebnisse von korpuslinguistischen Untersuchungen geben damit wichtige Einblicke (gesellschaftlich, sprachlich und sprachenpolitisch). Sie werden in ihrer Methodik nun auch durch Entwicklungen wie KI-Einbettung (siehe auch Leible et al. 2024) erweitert und durch Data Mining mit entsprechender Software unterstützt. Grundsätzlich bleibt die Wahl der Werkzeuge entsprechend der Datenlage und Fragestellung ausschlaggebend. Dafür gibt es allerdings hinsichtlich der Ausgangspunkte kein festes Schema. Bei Bubenhofer (2018) waren für die Zukunft der Korpusanalyse und den neuen Themen, wie KI und Datenmining dazu folgende vier Thesen zu finden:

These 1: In der Korpuslinguistik sollte es nicht (nur) um Belege gehen,
[...]

These 2: Bei der quantitativen Analyse von Sprache droht die Linguistik bedeutungslos zu werden, [...]

These 3: Um der Bedeutungslosigkeit zu entgehen, sind zwei Dinge wichtig: Linguistische Theorie und die hermeneutische Deutung der Ergebnisse quantitativer Analysen, [...]

These 4: Korpuslinguistisches Arbeiten ist diagrammatisches Operieren [...] (Bubenhofer 2018)

Mit seiner ersten These trifft Bubenhofer einen Nerv, denn natürlich ist der Beleg von großer Bedeutung. Dennoch gibt es mittlerweile eine Reihe an Fragestellungen, in denen die Abstraktion von Belegen eine größere Rolle spielt und dabei Aussagen nicht mehr zu dezidierten Belegen selbst getroffen werden, sondern zu Tendenzen und Korrelationen in Texten und somit zum Sprachgebrauch im Allgemeinen (vgl. Bubenhofer 2018).

Die zweite These erscheint auf den ersten Blick absolut im Hinblick auf den drohenden Bedeutungsverlust im Bereich der quantitativen Analyse durch Data Science und ihre Methoden wie Data Mining. Wie schon bei der ersten spielt der Autor auch hier mit der Bedeutungsdimension. Gerade die quantitative Analyse in der Linguistik eröffnet sehr oft Felder, die im zweiten Schritt qualitativ untersucht werden können, eben mit Blick auf einzelne Belege und das Vorkommen im Text. Natürlich muss hier auch der wachsende Einfluss der Computerlinguistik hervorgehoben werden, da durch sie die Be- und Verarbeitung von großen Datenmengen signifikant erleichtert, aber auch automatisiert wird. Diese Automatisierung beantwortet allerdings nicht unbedingt die Fragen der Linguistik bzw. kommt sogar ohne sie aus (vgl. Bubenhofer 2018).

Die in der dritten These angesprochenen Fähigkeiten sind von essenzieller Bedeutung. Auch hier kann eine erste Annäherung erfolgen. Dazu können bereits simple Werkzeuge genutzt werden, um zum Beispiel Studierende an das Thema Datenbanken heranzuführen. In vertiefender Art und Weise lohnt sich dann die Auseinandersetzung mit der Thematik des Korpus an sich und im Weiteren mit seiner Funktion und Darstellung. In kleineren Projektgruppen lassen sich solche Arbeitsweisen schon gut bei Studierenden

etablieren und im späteren Verlauf auf Master- und Promotionsebene vertiefend ausbauen. Grundlegend bleibt dabei für die Analyse, deren fundierte Kriterien zu kennen und entsprechend valide Analysekategorien zu bestimmen (vgl. Bubenhofer 2018). Dass die Korpuslinguistik gerade auch in Verbindung mit der Diskursanalyse Aspekte (siehe dazu Keller 1997; Traue et al. 2019) aufdecken kann, zeigt sich u.a. gut bei den Untersuchungen von Social Media im Bereich Unternehmenskommunikation (siehe dazu auch Zerfaß und Pleil 2015). Aber auch bei Webseiten und multimodaler Kommunikation kann die korpusbasierte Methode Erkenntnisse wie Details zu u. a. Kommunikationsstrategien, semantischen Clustern, Zielgruppenorientierung und Lokalisation hervorbringen. Ansätze für Studierende zu interdisziplinären Projekten können hier zu aussagestarken Ergebnissen führen und den Blick auf das Forschungsgebiet und den -gegenstand positiv verändern.

Mit seiner vierten These greift Bubenhofer noch ein großes Thema der korpuslinguistischen Arbeit auf. Die Darstellung von Ergebnissen auf eine reine Diagramm-an-Diagramm-Reihung zu reduzieren, wird den komplexen Ergebnissen nicht gerecht, besonders auch wenn es um maschinell erstellte Ergebnisse geht. Wissenschaftstheoretisch weitere Räume der Darstellung zu eröffnen scheint hier eine lohnende Überlegung (vgl. Bubenhofer 2018).

6. Fazit und Ausblick

Zusammenfassend lässt sich nach Betrachtung der verschiedenen Analysetrends in und um die Korpusanalyse konstatieren, dass durch sie aussagekräftige Kommentare zu Texten ermöglicht werden – sei es auf der qualitativen oder auf der quantitativen Ebene (zu computer- und onlinegestützten Methoden der Kommunikationsforschung siehe auch Welker 2019; Zeller 2017). Dabei werden die quantitativen Herangehensweisen wohl im Laufe der Zeit von der KI in Form von LLM, sogenannten Large Language Models¹⁰ (wie zum Beispiel ChatGPT), abgelöst. Musterhaftigkeit von Texten in ausgewählten Textsorten u.a. der Wirtschaftskommunikation auf der Basis

¹⁰ Zur Erläuterung LLMs vgl. auch <https://www.iese.fraunhofer.de/blog/large-language-models-ki-sprachmodelle/> (aufgerufen am 19.09.24).

der quantitativen Auswertung feststellen zu können, zeigt sich vor allem auch in Bezug auf gesellschaftliche und politische Diskurse als wertvolles Mittel. Durch qualitative Untersuchungen hingegen kann vielfach durch individuelle Fragestellungen (u.a. der Unternehmenskommunikation) Forschungsfragen zu Sentimenten oder auch Präferenzen nachgegangen werden und es können dementsprechende Auswertungen für die Unternehmen erfolgen. Die Relevanz der Korpusanalyse – mit KI-Unterstützung – in der Kompilation von Textdaten umgesetzt, erweist sich auch in der Zukunft als gewinnbringende Methode zur linguistischen Betrachtung von Sprachzusammenhängen. **N**

MICHAELA RUSCH

WESTSÄCHSISCHE HOCHSCHULE ZWICKAU /

UNIVERSITY OF APPLIED SCIENCES ZWICKAU

Literatur

Bücher und Journale

- Alexa, Melpomeni & Melanie Siegel 2021. *Tools für Social Listening und Sentiment-Analyse Einsatzfelder und Praxisbeispiele für die Analyse deutschsprachiger Online-Textdaten*. Wiesbaden: Springer. <https://doi.org/10.1007/978-3-658-33468-0>
- Banholzer, Volker M. 2019. Trend- und Sentiment-Analyse des Begriffs ‚Industrie 4.0‘–Social-Media-Monitoring von Innovationskommunikation. *Digitalisierung und Kommunikation. Europäische Kulturen in der Wirtschaftskommunikation*, vol 31, hrsg. von Marcus Stumpf. 161–178. Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-26113-9_10
- Bubenhofer, Noah 2018. Wenn „Linguistik“ in Korpuslinguistik bedeutungslos wird. Vier Thesen zur Zukunft der Korpuslinguistik. *Osnabrücker Beiträge zur Sprachtheorie* (OBST) 92: 17–30.
- Davies, Mark 2021. The Coronavirus Corpus Design, construction, and use. *International Journal of Corpus Linguistics* 26 (4): 593–598. <https://doi.org/10.1075/ijcl.21044.dav>
- Jamet, Denis & Pauline Rodet 2021. How is Brexit Linguistically-Constructed? A Corpus Linguistic Analysis of Speeches by David Cameron, Theresa May and Boris Johnson. *Observatoire de la société britannique* 27 : 21–53. <https://doi.org/10.4000/osb.5290>
- Crosthwaite, Peter, Sulistya Ningrum & Martin Schweinberger 2023. Research trends in corpus linguistics: A bibliometric analysis of two decades of Scopus-indexed corpus linguistics research in arts and humanities. *International Journal of Corpus Linguistics* 28 (3): 344–377. <https://doi.org/10.1075/ijcl.21072.cro>
- Crosthwaite, Peter & Vit Baisa 2023. Generative AI and the end of corpus-assisted data-driven learning? Not so fast!. *Applied Corpus Linguistics* 3 (3). <https://doi.org/10.1016/j.acorp.2023.100066>
- Gillings, Mathew, Tobias Kohn & Gerlinde Mautner 2024. The rise of large language models: challenges for Critical Discourse Studies. *Critical Discourse Studies*: 1–17. <https://doi.org/10.1080/17405904.2024.2373733>
- Große, Franziska 2009. Linguistische Bildanalysen am Beispiel von Marken- und Werbekommunikation. *Marke und Gesellschaft*, hrsg. von Nina Janich. 149–160. Wiesbaden: Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91683-5_10
- Jiang, Ke, Quian Xu & Ashleigh Afromsky 2023. What Do Employers Expect for Jobs Requiring Media Analytics? A Semantic Network Analysis of Job Descriptions of In-Person and Remote Positions During the COVID-19 Pandemic. *Journalism & Mass Communication Educator* 78 (1): 5–24. <https://doi.org/10.1177/10776958221135969>
- Keller, Reiner 1997. Diskursanalyse. *Sozialwissenschaftliche Hermeneutik*, hrsg. von Roland Hitzler &

- Anne Honer. 309–333. Wiesbaden: Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-663-11431-4_12
- Kesselheim, Wolfgang 2011. Textlinguistische Analyse und Textsortenkompetenz: Der Aktionsbrief. *Textsorten in der Wirtschaft: Zwischen textlinguistischem Wissen und wirtschaftlichem Handeln*, hrsg. von Sascha Demarmels & Wolfgang Kesselheim. 167–189. Wiesbaden: Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-94041-0_10
- Leible, Stephan, Gian-Luca Gücük, Dejan Simic et al. 2024. Zwischen Forschung und Praxis: Fähigkeiten und Limitationen generativer KI sowie ihre wachsende Bedeutung in der Zukunft. *HMD* 61: 344–370. <https://doi.org/10.1365/s40702-024-01050-x>
- Matras, Yaron 2009. *Language Contact*. Cambridge: Cambridge University Press.
- O’Flynn, James 2022. Lexical bundles in the academic writing of the Arts and Humanities: from corpus to CALL. *Yearbook of Phraseology* 13 (1): 81–108. <https://doi.org/10.1515/phras-2022-0006>
- O’Keefe, Anne & Michael J. McCarthy 2022. ‚Of what is past, or passing, or to come 1’: corpus linguistics, changes and challenges. *The Routledge Handbook of Corpus Linguistics*, hrsg. von Anne O’Keefe & Michael J. McCarthy, Michael J. 1–10. London: Routledge. <https://doi.org/10.4324/9780367076399>
- Sonnenhauser, Barbara, Hanne Eckhoff & Egbert Fortuin 2022. Ukrainian Linguistics. *Russian Linguistics* 46, 165–166. <https://doi.org/10.1007/s11185-022-09268-3>
- Traue, Boris, Lisa Pfahl & Lena Schürmann 2019. Diskursanalyse. *Handbuch Methoden der empirischen Sozialforschung*, hrsg. von Nina Baur & Jörg Blasius. 565–583. Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-21308-4_38
- Welker, Martin 2019. Computer- und onlinegestützte Methoden für die Untersuchung digitaler Kommunikation. *Handbuch Online-Kommunikation*, hrsg. von Wolfgang Schweiger & Klaus Beck. 531–572. Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-18016-4_21
- Wicher, Oliver (2019). Korpuslinguistik mit dem Corpus de référence du français contemporain. Forschungsergebnisse zur französischen Lexikogrammatik und fremdsprachendidaktische Anwendungen. *Digitalkulturen/Cultures numériques: Herausforderungen und interdisziplinäre Forschungsperspektiven/Enjeux et perspectives interdisciplinaires*, hrsg. von Julia Montemayor Gracia, Vera Neusius & Claudia Polzin-Haumann. 79–92. Bielefeld: transcript Verlag. <https://doi.org/10.1515/9783839442159>
- Wiertz, Thilo, Lilith Kuhn & Annika Mattissek 2023. A turn to geopolitics: Shifts in the German energy transition discourse in light of Russia’s war against Ukraine. *Energy Research & Social Science* 98. <https://doi.org/10.1016/j.erss.2023.103036>
- Zeller, Frauke 2017. Soziale Medien in der empirischen Forschung. *Handbuch Soziale Medien*, hrsg. von Jan-Hinrick Schmidt & Monika Taddiken. 167–187. Wiesbaden: Springer Verlag. https://doi.org/10.1007/978-3-658-03765-9_21

Zerfaß, Ansgar & Thomas Pleil 2015.
Strategische Kommunikation in
Internet und Social Web. *Handbuch
Online-PR, Strategische
Kommunikation in Internet und*

Social Web, hrsg. von Ansgar
Zerfaß & Thomas Pleil. 39–83
Konstanz: UVK.

URLs

Einführung in die Korpuslinguistik:
Praktische Grundlagen und
Werkzeuge. <https://www.bubenhofer.com/korpuslinguistik/kurs/>
(aufgerufen am 20.12.24)
CorpAn (2023). Corpus Analysis. –
<https://corpus-analysis.com/>
(aufgerufen am 14.09.23)
IVO Projekt – <https://ivohub.com/>
(aufgerufen am 19.09.23)
[https://www.jbe-platform.com/content/
journals/10.1075/ijcl.21072.cro](https://www.jbe-platform.com/content/journals/10.1075/ijcl.21072.cro)
(aufgerufen am 17.09.23)
Kelbert Patricia, Siebert, Julien und
Jöckel, Lisa (2023). „Was sind Large
Language Models? Und was ist bei
der Nutzung von
KI-Sprachmodellen zu beachten?“.
Blog des Fraunhofer-Institut für

Experimentelles Software
Engineering. <https://www.iese.fraunhofer.de/blog/large-language-models-ki-sprachmodelle/>
(aufgerufen am 19.09.24)
Nachhaltigkeit – [https://www.dwds.de/r/?
corpus=regional&q=Nachhaltigkeit](https://www.dwds.de/r/?corpus=regional&q=Nachhaltigkeit)
(aufgerufen am 23.09.23)
„Sprachtechnologie in den Digital
Humanities“ – <https://www.coursera.org/learn/digital-humanities>
(aufgerufen am 23.09.23)
Stadler, Max-Ludwig (2024).
„Sentimentanalyse“. [https://
mindsquare.de/knowhow/
sentimentanalyse/](https://mindsquare.de/knowhow/sentimentanalyse/) (aufgerufen am
20.12.24)

Korpora

GINGKO (2023). [https://www.philol.uni-leipzig.de/herder-institut/forschung/
projekte/laufende-projekte/gingko](https://www.philol.uni-leipzig.de/herder-institut/forschung/projekte/laufende-projekte/gingko)
(aufgerufen am 23.09.23)
Ortolang (2023) – <https://www.ortolang.fr/market/corpora> (aufgerufen am
23.09.23)
Mediapi-RGB (2017-2022).
Y. Ouakrim, H. Bull, M. Gouiffès, D.
Beautemps, T. Hueber, A. Braffort.
„Mediapi- RGB: Enabling Techno-

logical Breakthroughs in French
Sign Language (LSF) Research
through an Extensive Video-Text
Corpus“. VISAPP 2024 - 19th Inter-
national Joint Conference on Com-
puter Vision, Imaging and Com-
puter Graphics Theory and
Applications, Feb 2024, Rome, Italy.
DOI=10.5220/0012372600003660.
HAL=hal-04494094

Laboratoire Interdisciplinaire des
Sciences du Numérique - UMR 9015
(LISN) (2024). *Mediapi-RGB* [Cor-
pus]. ORTOLANG (Open Resources
and Tools for Language) - www.ortolang.fr, v1, [https://hdl.handle.
net/11403/mediapi-rgb/v1](https://hdl.handle.net/11403/mediapi-rgb/v1).

Unternehmenswebseiten

[https://www.loreal.com/fr/nos-
engagements/pour-la-planete/](https://www.loreal.com/fr/nos-engagements/pour-la-planete/)
(aufgerufen am 23.09.23)
[https://www.loreal.com/en/commitments-
and-responsibilities/for-the-planet/](https://www.loreal.com/en/commitments-and-responsibilities/for-the-planet/)
(aufgerufen am 23.09.23)
<https://www.peugeot.fr/> (aufgerufen am
23.09.23)
[https://www.peugeot.fr/marque/
univers-peugeot/peugeot-universe.
html](https://www.peugeot.fr/marque/univers-peugeot/peugeot-universe.html) (aufgerufen am 23.09.23)
<https://www.sncf.com/fr/groupe/coulisses>
(aufgerufen am 19.09.23)