

Marcel Dräger

## **Auffindbarkeit, Sichtbarkeit, Usability und Zukunft von digitalen Sprachressourcen**

Erkenntnisse und Perspektiven aus dem Forschungsprojekt  
„Digitale Sprachressourcen“ unter der Leitung von Prof. Dr. Christa Dürscheid  
und der Mitarbeit von Gerard Adarve, Dr. Patrizia Sutter, Dr. Marcel Dräger,  
Klaus Rothenhäusler, Nicole Zellweger und Desirée Noser

 **NET.WORX 90**

**FNSNF**

SCHWEIZERISCHER NATIONALFONDS  
ZUR FÖRDERUNG DER WISSENSCHAFTLICHEN FORSCHUNG

# NETWORX

## IMPRESSUM

---

<b>Herausgeber</b>	Dr. Jens Runkehl, Prof. Dr. Peter Schlobinski, Dr. Torsten Siever
<b>Editorial-Board</b>	Prof. Dr. <b>Jannis Androutsopoulos</b> (Universität Hamburg) für den Bereich Medienanalyse; Prof. Dr. <b>Christa Dürscheid</b> (Universität Zürich) für den Bereich Mobile Kommunikation; Prof. Dr. <b>Nina Janich</b> (Technische Universität Darmstadt) für den Bereich Werbesprache; Prof. Dr. <b>Ulrich Schmitz</b> (Universität Duisburg-Essen) für den Bereich Digitale Kommunikation
<b>ISSN</b>	1619-1021
<b>Anschrift</b>	Leibniz Universität Hannover, Deutsches Seminar, Königsworther Platz 1, 30167 Hannover Internet: <a href="http://www.mediensprache.net/networx/">www.mediensprache.net/networx/</a> E-Mail: <a href="mailto:networx@mediensprache.net">networx@mediensprache.net</a>

## ZU DIESER ARBEIT

---

<b>Autor &amp; Titel</b>	Marcel Dräger (2020). Auffindbarkeit, Sichtbarkeit, Usability und Zukunft von digitalen Sprachressourcen.
<b>Version</b>	1.0 (2020-12-21)
<b>DOI</b>	10.15488/10317
<b>Zitierweise</b>	Dräger, Marcel (2019). Auffindbarkeit, Sichtbarkeit, Usability und Zukunft von digitalen Sprachressourcen. < <a href="http://www.mediensprache.net/networx/networx-90.pdf">http://www.mediensprache.net/networx/networx-90.pdf</a> >. In: <i>Networx</i> . Nr. 90. ISSN: 1619-1021. DOI: 10.15488/10317.
<b>Zitiert nach</b>	Runkehl, Jens und Torsten Siever (2001). Das Zitat im Internet. Ein Electronic Style Guide zum Publizieren, Bibliografieren und Zitieren. Hannover

## MANUSKRIPTE

---

<b>Einsendung</b>	Die Einsendung von Beiträgen und Mitteilungen sind an folgende E-Mail-Adresse zu richten: <a href="mailto:networx@mediensprache.net">networx@mediensprache.net</a> oder an die Postadresse: Prof. Dr. Peter Schlobinski, Leibniz Universität Hannover, Deutsches Seminar, Königsworther Platz 1, 30167 Hannover.
<b>Autorenhinweis</b>	Mit der Annahme des Manuskripts zur Veröffentlichung in der Schriftenreihe Networx räumt der Autor dem Projekt <a href="http://www.mediensprache.net">mediensprache.net</a> das zeitlich, räumlich und inhaltlich unbeschränkte Nutzungsrecht ein. Dieses beinhaltet das Recht der Nutzung und Wiedergabe. Ein Recht auf Veröffentlichung besteht nicht.
<b>Begutachtung</b>	Die Begutachtung eingesandter Beiträge wird von den Herausgebern sowie den Vertretern des Editorial Board vorgenommen.

## Networx

ist die Online-Schriftenreihe des Projekts [mediensprache.net](http://www.mediensprache.net). Die Reihe ist eine eingetragene Publikation beim Nationalen ISSN-Zentrum der Deutschen Bibliothek in Frankfurt am Main und ist ausgestattet mit einem DOI.

## Einsenden?

Möchten Sie eine eigene Arbeit in der Networx-Reihe veröffentlichen? Dann senden Sie uns Ihren Text an folgende E-Mail-Adresse: [networx@mediensprache.net](mailto:networx@mediensprache.net) oder an die Postadresse: Prof. Dr. Peter Schlobinski, Leibniz Universität Hannover, Deutsches Seminar, Königsworther Platz 1, 30167 Hannover.

## Homepage:

Alle Arbeiten der Networx-Reihe sind kostenlos im Internet downloadbar unter:

<https://www.mediensprache.net/networx/>

## Copyright

© Projekt [mediensprache.net](http://www.mediensprache.net)  
Die Publikationsreihe Networx sowie alle in ihr veröffentlichten Beiträge und Abbildungen sind urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne ausdrückliche Zustimmung des Projekts [mediensprache.net](http://www.mediensprache.net) unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

## Informationsstand

Stand der hier angegebenen Informationen – soweit nicht anders vermerkt ist: **Februar 2020**

# Inhalt

1	Vorbemerkungen – Lexikographie im digitalen Wandel.....	5
2	Perspektiven auf digitale Sprachressourcen .....	7
2.1	Steht die Lexikographie vor einem disruptiven Wandel?.....	7
2.2	Die Filterblasen-Gefahr .....	10
2.3	Websites sind kein nachhaltiges Publikationskonzept .....	11
2.4	Mehr als Klickzahlen.....	12
2.5	Informationen als Wert der wissenschaftlichen Lexikographie .....	13
3	Wie werden digitale Sprachressourcen genutzt? .....	13
3.1	Für wen konzipiert man eine Sprachressource? .....	14
3.2	Suchmaschinen verändern die Nutzung und den Nutzendenkreis .....	17
3.3	Benutzungsforschung für digitale Sprachressourcen neu denken.....	19
3.4	Datenbanken können mehr als statische Texte ausgeben .....	19
3.5	Kein Widerspruch: schneller Informationsgewinn und Qualität der Information .....	20
4	Werk oder Informationsspeicher? – Ein neuer theoretischer Ansatz für die digitale Lexikographie ..	22
4.1	Vom Werk zur digitalen „Ressource“ .....	22
4.1.1	Nachschlagen im Wörterbuch hat die Lexikographie geprägt .....	22
4.1.2	Die Finanzierung lexikographischer Online-Angebote mit Werbung.....	25
4.1.3	Signale der Vertrauenswürdigkeit .....	25
4.1.4	Auffindbarkeit und Sichtbarkeit.....	26
4.2	Zwei Identitäten einer digitalen Sprachressource .....	27
4.2.1	Im Internet werden Informationen direkt gesucht und gefunden .....	27
4.2.2	Ist das Werk noch wichtig? .....	30
4.2.3	Der Nutzen dieser Unterscheidung .....	30
4.2.4	Lexikographie-kompatibles Wissenschaftsmarketing .....	31
5	Auffindbarkeit, Sichtbarkeit und Usability digitaler Sprachressourcen .....	32
5.1	Auffindbarkeit – Digitale Werke und Informationen in Suchmaschinen .....	32
5.1.1	Das Werk finden .....	33
5.1.2	Informationen finden.....	35
5.2	Sichtbarkeit – Das Werk hinter den Informationen wahrnehmen.....	45

5.2.1	Was soll im Suchergebnis dargestellt werden? .....	45
5.2.2	Sichtbarkeit in der Suchmaschine .....	46
5.3	Usability – Bieten, was versprochen ist .....	48
5.3.1	Neue Anforderungen an die Usability .....	50
5.3.2	Nutzende auf der Seite halten .....	50
6	Die Zukunft digitaler Sprachressourcen .....	55
6.1	Usability aus nicht-lexikographischen Umgebungen importieren .....	55
6.2	Plattform statt Einzelressourcen?.....	60
6.3	Digitale Suche, Filter und Ranking .....	62
6.3.1	Digitale Suche.....	62
6.3.2	Filter und Kategorisierungen .....	63
6.3.3	Ranking der Ergebnisse .....	63
6.4	Lexikographische Videos .....	64
6.5	Beurteilung der Wertigkeit .....	65
6.6	Finanzierung lexikographischer Angebote im Internet .....	67
6.7	Vision: Ein Marktplatz für Sprachinformationen .....	68
7	Literatur .....	70

# 1 Vorbemerkungen – Lexikographie im digitalen Wandel

In dem Moment, in dem eine Sprachressource im Internet verfügbar ist, wird sie anders, von anderen Personen und auch zu anderen Zwecken genutzt, als wenn sie nicht online zugänglich wäre. Die Lexikographie muss dieser veränderten Nutzungsweise Rechnung tragen: Denn mit der Digitalisierung verändern sich nicht nur die lexikographisch-redaktionellen Prozesse sowie die Darstellung und die Rezeption von Sprachinformationen, es verändert sich auch das Distributions- und Publikationskonzept.

Im Folgenden geht es um Fragen, Ideen und Vorschläge zur zukünftigen Gestaltung, Publikation und Distribution wissenschaftlich erarbeiteter digitaler Sprachressourcen. Grundlage hierfür sind Erkenntnisse aus der Forschung, aus der eigenen lexikographischen Praxis und aus einem Forschungsprojekt, das sich mit der Auffindbarkeit, Sichtbarkeit und Usability von digitalen Sprachressourcen beschäftigt hat (zum Projekt vgl. die folgende Infobox).

## **INFO: Das Projekt „Digitale Sprachressourcen. Empirische Analysen und Perspektiven“ (im Weiteren „DL-Projekt“)**

### KURZBESCHREIBUNG:

Digitale Sprachressourcen müssen in erster Linie qualitativen Ansprüchen genügen, sie müssen aber auch so gestaltet sein, dass sie im Internet schnell gefunden werden und dass sie intuitiv leicht zugänglich sind. Im Projekt wird untersucht, wie solche Sprachangebote benutzerfreundlich konzipiert und im Netz sichtbar dargestellt werden können und wie sich die für die Nutzung dieser Ressourcen erforderliche Recherche-Kompetenz am besten vermitteln lässt. Hierzu wurden auf der Basis eines ausgewählten Online-Nachschlagewerkes, der *Variantengrammatik des Standarddeutschen* (im Weiteren nur *Variantengrammatik*), einige Fallstudien durchgeführt. Als Grundlage hierfür wurden theoretische Überlegungen angestellt, wie sich das Informieren und Nachschlagen im Digitalen wandelt. Die Erkenntnisse aus dem Projekt sind in diesen Text eingeflossen und sollen Forschergruppen nützliche Hinweise zur Konzeption, Diffusion und Sichtbarmachung von digitalen Sprachressourcen im Open-Access geben.

### FALLSTUDIEN UND WORKSHOPS:

- Log-File-Analysen zur Nutzung der *Variantengrammatik* über den Zeitraum von 6 Monaten
- Experimente zur Auffindbarkeit von Artikeln aus der *Variantengrammatik* via *Google*
- Adaption und Testung der SEO-Empfehlungen (SEO = Suchmaschinenoptimierung) von *Google* auf die spezifische Struktur, Gestalt und Anforderung der wissenschaftlichen Lexikographie
- Umfrage zur Nutzung von (digitalen) Sprachressourcen
- Vergleichende Analyse der Online-Auftritte digitaler Sprachressourcen und der Startseite von Websites mit Informations- und Rechercheangeboten
- Theoretische Auseinandersetzung mit dem digitalen „Nachschlagen“ von Sprachinformationen
- Austausch mit kooperierenden lexikographischen Expertinnen und Experten
- Workshops zur Recherchekompetenz mit Studierenden der Journalistik, der Unternehmenskommunikation und des höheren Lehramts

### PROJEKT-TEAM:

Projektleitung: Prof. Dr. Christa Dürscheid

Mitarbeitende: Gerard Adarve, Dr. Patrizia Sutter, Dr. Marcel Dräger, Klaus Rothenhäusler

Hilfskräfte: Nicole Zellweger, Desirée Noser

### FÖRDERUNG:

Das Projekt wurde im Rahmen des Programms *Digital Lives* vom Schweizerischen Nationalfonds gefördert.

### PROJEKTLAUFZEIT:

September 2018 bis Oktober 2020

Die wissenschaftliche Lexikographie vernetzt sich seit Beginn der digitalen Publikation von Wörterbüchern in zahlreichen kleineren und größeren Kooperationen.<sup>1</sup> So gibt es Bestrebungen, auf der Grundlage der TEI-Guidelines<sup>2</sup> zum Markup von Texten ein allgemeines lexikographisches Daten-Austausch-Format zu entwickeln. Auch Kollaborationen der Wörterbuchredaktionen sind vielversprechende Schritte in eine Zukunft, in welcher wissenschaftliche Sprachressourcen weiterhin eine wichtige Anlaufstelle für Personen mit entsprechenden Fragestellungen sind. Grundlage all dieser Pläne ist, zentrale Plattformen für sprachbezogene Fragestellungen zu schaffen (vgl. dazu Kap. 6.2). Solche Plattformen gleichen dem Bücherregal einer gut sortierten Bibliothek mit Nachschlagewerken, in dem eben nicht nur die großen Gesamtwerte zugänglich sind, sondern auch kleinere und spezialisierte Fachpublikationen. Ein Ziel der Lexikographie muss sein, durch Kooperationen und gegebenenfalls Kompetenzzentrierung auch diesen wertvollen wissenschaftlichen Publikationen eine größere, weit über Fachgrenzen hinausragende Sichtbarkeit im Digitalen zu verschaffen. Doch auch wenn wissenschaftliche Plattformen für Sprachfragen jeglicher Art sehr wünschenswert sind, die digitale Auffindbarkeit, Sichtbarkeit und Benutzbarkeit einzelner Angebote ist von solchen Meta-Angeboten nicht abhängig.<sup>3</sup>

Gerade das Internet führt dazu, dass man nicht mehr das Buch besitzen und es aus dem Regal nehmen muss, um etwas nachzuschlagen. Das Angebot an qualifizierten Antworten auf Sprachfragen ist enorm, potentiell sind alle Informationen erreichbar, die im Internet publiziert wurden. Die allgemeine Online-Verfügbarkeit macht es möglich, dass die Interessen von Nachschlagenden und passgenaue Wissensangebote schneller bzw. überhaupt zusammenfinden. Aber ebenso groß wie dieser Vorteil ist auch die Gefahr, dass kleinere lexikographische Angebote in der Online-Informationsflut nicht gefunden, gesehen und genutzt werden.

Für das einzelne Projekt ist es eine große Anstrengung, Suchende auf die eigene Seite der Sprachressource zu führen. Das hierfür nötige Knowhow umfasst keine spezifisch lexikographische Kompetenz, sondern hängt mit den Bedingungen für eine gute Positionierung in Suchmaschinen zusammen. Doch gerade die Optimierung der eigenen Website für Suchmaschinen (auch *SEO* für *search engine optimization*) gilt als kosten- und aufwandsintensiv. Und sie wird als ein Instrument des Informationsmarketings von der wissenschaftlichen Lexikographie zwiespältig beurteilt. So sehr diese Bedenken nachvollziehbar sind: In einer digitalen Zukunft, in welcher Suchmaschinen – derzeit überwiegend *Google* – das Recherchieren von Informationen dominieren, ist es wichtig, dass die mit viel Engagement und Fachwissen erarbeiteten und qualitativ herausstehenden Informationen auch gefunden werden. Für die Anbieter von Sprachressourcen mit spezialisierten Inhalten liegt in dieser digitalen Zugangsform sogar die große Chance, dass ihr Produkt häufiger genutzt wird, als es mit den überwiegend nur in Bibliotheken verfügbaren Fachwörterbüchern der Fall wäre.

Im Fokus dieses Textes steht daher die Frage: Wie lassen sich die Auffindbarkeit, die Sichtbarkeit und die Benutzbarkeit einer bestehenden oder geplanten Sprachressource verbessern (vgl. v.a. Kap. 5)? Dabei geht es auch darum, zu zeigen, dass digitale Sprachressourcen den Algorithmen von

---

<sup>1</sup> Beispiele für solche Kooperationen sind (in unvollständiger Auflistung): Forschungsschwerpunkt Vernetzung am *IDS Mannheim* (<https://www1.ids-mannheim.de/lexik/bzvelexiko/vernetzung.html>); *ZHistLex – eHumanities-Zentrum für Historische Lexikographie* der deutschen Akademien der Wissenschaft und der Universität Frankfurt (<https://zhistlex.de/>); *Wörterbuchnetz des Trier Center for Digital Humanities* (<http://www.woerterbuchnetz.de/cgi-bin/WBNetz/setupStartSeite.tcl>); Mehrere ehemals gedruckte Wörterbücher sind über die Seite des *DWDS* zugänglich (<https://www.dwds.de/d/woerterbuecher>); vgl. auch Boelhouwer/Dykstra/Sijens (2018) für einen internationalen Überblick.

<sup>2</sup> TEI = Text Encoding Initiative (<https://tei-c.org/guidelines/>).

<sup>3</sup> Dass die Bereiche der Sichtbarkeit, der Auffindbarkeit und möglicher Portallösungen bei einem Blick in die Zukunft eng zusammenwirken, zeigt das abschließende Kapitel.

Suchmaschinen sogar entgegenkommen – im Prinzip passen sie also gut in das Konzept des derzeit stark forcierten Content-Marketings. Wichtig ist aber, dass bei aller Optimierung die lexikographischen Inhalte und die Autonomie des zugrundeliegenden Werkes unangetastet bleiben. Den Überlegungen zur Auffindbarkeit und Sichtbarkeit von digitalen Sprachressourcen vorausgeschickt werden daher Fragen zur Benutzung und zu Nutzenden (vgl. Kap. 3) und prinzipielle, theoretische Überlegungen zum Nachschlagen und zu Nachschlagewerken im Digitalen (vgl. Kap. 4). Vorab aber seien die Veränderungen in der Lexikographie, die möglicherweise eine lange bestehende lexikographische Tradition und Denkweise grundlegend in Frage stellen, diskutiert (vgl. Kap.2). Der Text schließt mit einem Blick in die Zukunft, der die zentralen Fragen aus den vorhergehenden Kapiteln aufgreift und zusammenführt (Kap. 6).

## 2 Perspektiven auf digitale Sprachressourcen

### 2.1 Steht die Lexikographie vor einem disruptiven Wandel?

„Disruption“ ist ein Schlagwort im Digitalisierungsdiskurs. Gemeint ist damit der Umstand, dass neue Technologien alte, oftmals traditionelle ersetzen. So schreibt *Der Spiegel* (19/2018): „Künstliche Intelligenz revolutioniert die Sprachübersetzung. Die Kölner Firma DeepL kann das besser als GoogleTranslate. [...] DeepL macht damit das Gleiche, was GoogleTranslate oder Microsofts Bing-Übersetzer schon lange machen – nur eben viel besser.“ Diese Revolution – durch die eine digitale Technologie eine ebenfalls digitale Vorgängertechnologie abzulösen scheint – könnte sich demnach auch im Übersetzungswesen zeigen. In ihrem Beitrag von 2012 zu Übersetzungswörterbüchern hat Bowker noch nicht ansatzweise eine Software erwähnt. Doch genau eine solche Software hat keine sechs Jahre später den gesamten Übersetzungsprozess mit einer überraschenden Qualität automatisiert. Diese nun wachsende Konkurrenz durch die neue Technologie von *DeepL* stellt die Branche vor Probleme: Die guten Argumente für eine nicht automatisierte Übersetzung werden schwächer. Die Situation ähnelt der Entwicklung in der gesamten Lexikographie, die ebenfalls eine stabile, kontinuierliche Tradition aufweist und nun auf die technologisch bedingten Informationsmöglichkeiten und auf ein neues Informationsverhalten reagieren muss.

Wenn man die Geschichte der Lexikographie verfolgt, dann lässt sich darin eine kontinuierliche publikationstechnologische Weiterentwicklung erkennen: Die Lexikographie und ihre Werke haben sich von Sammlungen bestimmter Ausdrücke (Kodizes) zu Wörterbuch und Lexikon entwickelt, dann über deren digitale Varianten auf CD-ROM, anschließend kamen PDF-basierte Online-Publikationen dieser Werke auf den Markt, und heute sind datenbankbasierte digitale Nachschlagewerke gängig. Diese Geschichte lässt sich holzschnittartig beschreiben als eine Entwicklung von Makrostrukturen zu Mikrostrukturen, es folgte eine Terminologie zu den Artikelstrukturen und damit einhergehend eine Theorie der Lexikographie.<sup>4</sup>

Mit den ersten digitalen Produkten auf CD-ROM kamen neue Erschließungsstrategien der Wörterbuchinhalte in den Blick<sup>5</sup> (vgl. bspw. Dräger 2009) und mit der Online-Publikation die XML-basierte Auszeichnung der lexikographischen Informationen und Strukturen. Aus linear strukturierten

---

<sup>4</sup> Bis in die 90er Jahre ist die Geschichte der Lexikographie gut dokumentiert und aufgearbeitet, bspw. in Haß-Zumkehr (2001) oder Wiegand (1998). Für die Entwicklung der Internetlexikographie siehe Klosa/Müller-Spitzer (2016).

<sup>5</sup> In Dräger (2009) habe ich gezeigt, welche „verschollenen“ Informationen sich gewinnen lassen, wenn man digitalisierte Wörterbücher als Korpora benutzt.

Büchern wurden Datenbanken mit Suchmasken. Mit den Wiki-Systemen steht derzeit sogar ein allgemeines und kostenloses lexikographisches Redaktions- und Publikationswerkzeug zur Verfügung. Diese Entwicklung war immer wieder von bedeutenden Fortschritten gekennzeichnet (bspw. durch die Einführung einer alphabetischen anstelle einer thematischen Ordnung oder durch das Angebot detaillierter digitaler Suchstrategien), und dennoch lässt sich zwischen den einzelnen Schritten stets Kontinuität erkennen. Das mag daran liegen, dass in der Regel frühere Publikationsstufen (bspw. ein gedrucktes Wörterbuch) in neue Medien überführt wurden (bspw. in ein durchsuchbares PDF online). Zudem mag die Kontinuität darauf zurückzuführen sein, dass die beteiligten Lexikographinnen und Lexikographen durch eine bestimmte Tradition und Denkweise geprägt sind. So stehen auch die meisten derzeit publizierten digitalen Sprachressourcen noch erkennbar in dieser Tradition.

*DeepL* allerdings ist eine Sprachressource, die nicht mehr einer lexikographischen, sondern einer informationstechnischen Grundidee entsprungen ist: Es basiert auf Deep Learning, einer Form von künstlicher Intelligenz. Im Falle von *DeepL* kann man durchaus von einer Disruption sprechen. Ist diese Entwicklung auch für die Lexikographie denkbar? Basiert die Lexikographie und basieren ihre Werke auf einem überholten und digital ersetzbaren Konzept?

Die Stärke der wissenschaftlichen Lexikographie sind redaktionell-lexikographisch verknüpfte Daten über Sprache, Sprachstrukturen und den Sprachgebrauch. Diese Informationen wurden aus den bekannten Wörterbüchern heraus digitalisiert und sind damit einfacher, schneller und zunehmend online erschließbar. Die entsprechenden Online-Plattformen hatten und haben aufgrund ihrer Beziehung zum gedruckten Buch eine entsprechende Reputation (bspw. *Duden* und die Online-Version *duden.de*).<sup>6</sup> Dann aber entstanden und entstehen neue digitale Nachschlagewerke, denen kein gedrucktes Werk entspricht (bspw. *Wikipedia*, *Wiktionary*, aber auch spezialisierte Fachpublikationen wie *www.variantengrammatik.net*). Diese digitalen Werke wollen und müssen von den Personen gefunden werden, die Fragen zur Sprache und zum Sprachgebrauch haben. Und sie wollen von diesen Menschen genutzt und gekannt werden. Über klassische Distributionswege wie den Buchhandel oder Bibliothekskataloge sind rein digitale Sprachressourcen – die kein Buch, sondern eine Website sind – nicht mehr auffindbar. Sollen neue Werke wie die *Variantengrammatik* (vgl. die folgende Infobox) über Fachkreise hinaus wahrgenommen und genutzt werden, bleiben derzeit zwei Möglichkeiten: eigene Werbung (bspw. via *Twitter* oder Medienberichte) und eine gute Positionierung in Suchmaschinen. Ein Blick in die Log-Files der *Variantengrammatik* zeigt: Besuchende, die die Seite direkt aufrufen, gibt es im Vergleich zu den Besuchen über Suchmaschinen nur wenige. Zusammen mit den Besuchenden über Links von externen Seiten machen sie nur elf Prozent aller Besuche aus. PR-Maßnahmen wie Zeitungsartikel über das *Variantengrammatik*-Projekt<sup>7</sup> oder die regelmäßigen Beiträge auf dem *Twitter*-Account der *Variantengrammatik* führen zu kurzfristigen, aber deutlichen Anstiegen der Besuche, was oftmals daran erkennbar ist, dass dieselben, in dem Medienbericht oder dem jeweiligen Tweet erwähnten, Phänomene nachgeschlagen werden. Wie nachhaltig diese Maßnahmen sind, ließ sich aufgrund der kurzen Beobachtungszeit nicht evaluieren. Offensichtlich ist aber, dass auch bei den Besuchen, die länger als drei Sekunden dauern, 64% über Suchmaschinen auf die Website kommen.

---

<sup>6</sup> In einer Umfrage unter den Studierenden (Fächer: Journalismus, Unternehmenskommunikation und Lehramt Deutsch), die an den im Projekt „Digitale Sprachressourcen“ durchgeführten Workshops teilgenommen haben, wurden *duden.de* und *wikipedia.de* als die bekanntesten Online-Nachschlagewerke genannt.

<sup>7</sup> Eine Übersicht findet sich auf der Projekt-Website: <http://www.variantengrammatik.net/publikationen.html#medienberichte>.

## INFO: Die Variantengrammatik als typisches Beispiel einer digitalen Sprachressource

Die *Variantengrammatik* ist eine digitale Sprachressource, die in langjähriger internationaler Zusammenarbeit von Fachleuten aus Variationslinguistik, Grammatik und Computerlinguistik erarbeitet wurde. Sie ist für die Sprachwissenschaft und Sprachdidaktik, aber auch für einen interessierten Nutzendenkreis jenseits dieser Fachdisziplinen konzipiert. Für das Projekt wurde der deutschsprachige Raum in 15 Großregionen (= Areale) gegliedert, um regionale Unterschiede im Bereich der Grammatik herauszuarbeiten, also z.B. Fragen der Wort- und Satzgliedstellung und der Wortbildung. Die Datengrundlage hierfür lieferten die Online-Ausgaben von 68 regional verbreiteten Zeitungen aus dem gesamten deutschsprachigen Raum. In solchen regionalen Zeitungen wird eine regional geprägte Standardsprache verwendet, aber eben kein Dialekt und keine Umgangssprache. Die Ergebnisse und die lexikographischen Ausführungen und Erläuterungen wurden mit dem Wiki-System von *MediaWiki.org* erstellt und publiziert.

The image shows two screenshots of the 'Variantengrammatik' website. The left screenshot is the homepage, featuring a search bar, navigation links, and three main sections: 'Suche', 'Alle Artikel', and 'Projekt'. The right screenshot is an article titled 'Virus', which includes a text introduction, a list of 'Beispielsätze', and three maps of Germany showing the distribution of the virus in different dialectal areas.

Startseite der *Variantengrammatik des Standarddeutschen*

Artikel zur Verbreitung von „der“ bzw. „das Virus“ im dt. Sprachraum

Das Gesamtprojekt *Variantengrammatik* bestand aus drei nationalen Teil-Projekten, die zwischen 2012 und 2018 von Institutionen aus der Schweiz, Österreich und Deutschland (Schweizerischer Nationalfonds, Wissenschaftsfonds FWF und Deutsche Forschungsgemeinschaft) gefördert wurden. Damit entsprach die Erstellung der *Variantengrammatik* einem typischen lexikographischen Wissenschaftsprojekt, das nicht in eine Langzeitförderung (bspw. an einer Akademie) eingebunden ist. In solchen Projekten stellen sich zu Anfang viele Fragen (beispielsweise hinsichtlich der Erstellung einer Datengrundlage, bei der Ausarbeitung eines lexikographischen Konzepts, bei der Installation eines Redaktionssystems, bei der parallelen Entwicklung von prototypischen Probeartikeln und bei der Einarbeitung eines Redaktionsteams). Die Vermarktung spielt in dieser Phase oder auch generell in wissenschaftlichen Projekten in der Regel keine Rolle, auch weil das bei Printprodukten von den Verlagen organisiert wird. In Bezug auf digitale Publikationen ohne Verlagskontext fällt dieser Aufgabenbereich aber auf das wissenschaftliche Team zurück.

Mit dem Projekt „Digitale Sprachressourcen“ bot sich nun die Möglichkeit, die *Variantengrammatik* als typische Stellvertreterin für wissenschaftlich-lexikographische Publikationen hinsichtlich ihres digitalen Auftritts exemplarisch zu evaluieren und Fragen, Ideen und Vorschläge für Projekte mit einer ähnlichen Ausgangslage zu erarbeiten.

## 2.2 Die Filterblasen-Gefahr

Bekannt ist die Diskussion um sogenannte Filterblasen („filter bubbles“)<sup>8</sup>, die dazu führen, dass man aufgrund sozialer und medialer Einflüsse nur die Informationen auffindet, die man schon kennt. Als das Projektteam begann, die Log-Files der *Variantengrammatik* zu analysieren, zeigte sich schnell, was der wissenschaftliche Diskurs und die Forschung in diesem Bereich schon publiziert und diskutiert hat.<sup>9</sup> Passend zu dieser Feststellung, dass sich vieles bestätigt, beschreibt Lew (2015, 8) einen „dictionary user cycle“, wonach mit den Strukturen zufriedene Nutzende Innovationen verhindern, solange die Verlage diese Menschen als Referenzgruppe für die Akzeptanz von Innovationen heranziehen. Er schreibt (ebd.): „One way to get out of the vicious circle is to employ user studies in order to try out novel, experimental lexicographic features.“ Nur das Aufbrechen des Kreises könne zu innovativen neuen Strukturen führen. Im Projektteam haben wir<sup>10</sup> uns gefragt, ob wir mit unseren methodischen Ansätzen nicht auch in einem Kreis, einer Filterblase, stecken, der Wörterbuchstrukturen anhand eines eingegrenzten Publikums evaluiert: Können wir mit den klassischen wissenschaftlichen Methoden wie Umfragen, Log-File-Analysen oder einer Eye-Tracking-Studie<sup>11</sup> etwas herausfinden, was jenseits der Filterblase liegt? Können wir mit einer wissenschaftlich erarbeiteten Sprachressource, die auf einem klassischen Wiki-Redaktionssystem aufgebaut ist, etwas herausfinden über Möglichkeiten, die jenseits der systembedingt vorgegebenen Strukturen liegen? Können wir durch die Testung von Artikelstrukturen, die der Tradition der Buchlexikographie entstammen, etwas zum Potential der Digitalisierung aussagen? Merken wir vielleicht gar nicht, welche disruptiven Elemente sich jenseits der Filterblase entwickeln? Unser Ziel war, einen Beitrag zur Beschreibung einer zukunftsfähigen Entwicklung von digitalen Sprachressourcen zu leisten. Lässt sich aber ein solcher Beitrag leisten, indem lediglich eine lineare Weiterentwicklung bestehender Konzepte beschrieben wird?

Ohne den Weg der bisherigen Lexikographie in Frage stellen zu wollen, haben wir uns entschieden, eine andere Perspektive auf digitale Sprachressourcen einzunehmen. Es ist die Perspektive der Menschen, die digitale Sprachressourcen bisher nicht oder selten nutzen, die sich jenseits der Filterblase bewegen und die in dieser auch nicht (oder nur ganz kurz und anonym in den Log-Files) sichtbar werden. Die grundsätzliche Einstellung dabei ist folgende positive Grundannahme: **Es gibt sehr viele Personen, für die die Inhalte digitaler Sprachressourcen relevant wären und die an diesen Inhalten interessiert wären.** Es sind dies Personen, die noch nichts von der Sprachressource wissen, für die sie sich interessieren könnten. Weil dieser Personenkreis derart unspezifisch ist, gehen wir davon aus, dass es potentiell jede Person sein kann, die im Internet nach einer (Sprach-)Information sucht. Im Sinne der freien Verfügbarkeit (Open-Access) von wissenschaftlich erarbeiteten Informationen sollten digitale Sprachressourcen von allen auffindbar und für alle zugänglich sein. Das schien uns nicht nur eine zeitgemäße wissenschaftliche Positionierung, sondern auch eine Konsequenz aus der Digitalisierung. Doch dabei zeigte sich bald, dass eine freie Verfügbarkeit zwar schnell hergestellt ist, wenn eine Sprachressource im Netz frei verfügbar ist, dass die fehlende Auffindbarkeit aber weiterhin eine Zugangsbarriere darstellt.

---

<sup>8</sup> <https://de.wikipedia.org/wiki/Filterblase>, 05.10.2020; Begriff nach Pariser (2011).

<sup>9</sup> Abweichungen, wie ein hohes Maß an Nutzenden, die sich scheinbar wahllos durch die *Variantengrammatik* klicken, ließen sich schlüssig erklären. In diesem Fall lag es wohl an der Tatsache, dass die Online-Ressource neu war und in einigen Zeitungsberichten darüber berichtet wurde.

<sup>10</sup> Das Personalpronomen „wir“ bezieht sich im Text auf das Team des *DL-Projekts*, mit „ich“ wird auf den Autor, Marcel Dräger, Bezug genommen.

<sup>11</sup> Die geplante Studie konnte aufgrund hygienischer Auflagen während der Pandemie nicht in der Projektzeit durchgeführt werden.

## 2.3 Websites sind kein nachhaltiges Publikationskonzept

Im Vergleich zu den Publikationskosten von gedruckten Nachschlagewerken erscheinen die Kosten für Online-Publikationen niedrig. Doch mit den Kosten für die Publikation eines gedruckten Nachschlagewerks ist auch dessen Auffindbarkeit in Bibliothekskatalogen, in Buchhandlungen und Online-Shops sowie die Vermarktung durch den Verlag enthalten. Eine Online-Präsenz leistet all das nicht, insbesondere wenn es keine der oben erwähnten PR-Bestrebungen gibt (diese sind im Rahmen der Wissenschaftsförderung nicht vorgesehen). Die Sichtbarkeit und Auffindbarkeit wissenschaftlicher Online-Ressourcen ist daher nicht per se gewährleistet.

Anders als beim gedruckten Nachschlagewerk kann jede und jeder relativ ortsunabhängig und niederschwellig Online-Sprachressourcen nutzen. Das ist ein Vorteil. Faktisch aber müssen die Betreiber oder Betreiberinnen einer solchen Website dafür sorgen, dass die Site und damit das dort konservierte Wissen von möglichst vielen gefunden werden. Für diese neue Situation fehlt es aber noch an Zukunftsideen und Finanzierungskonzepten: Werbefinanzierte Websites werden in der wissenschaftlichen Community (noch) weitgehend abgelehnt und die finanzierenden Forschungsgemeinschaften übernehmen den Unterhalt von Websites nicht. Die Bereitschaft potentieller Nutzender, für den Zugang zu einer Online-Ressource etwas zu zahlen, ist deutlich niedriger als bei Büchern. Dazu kommt das projektorientierte Denken und Finanzieren in der Forschung, weshalb über die Laufzeit eines lexikographischen Projekts hinaus kaum Mittel für die Pflege und Publikation des Produkts bereitstehen. Eine Website anzubieten und aktuell zu halten reicht nicht aus – es braucht vielmehr eine zukunftsfähige Idee für die Distribution und Öffentlichkeitsarbeit zu den unterschiedlichen digitalen Sprachressourcen.

Solche Gedanken an die Distribution bzw. Vermarktung scheinen nicht in der Konzeption und Evaluation von lexikographischen Projekten vorgesehen: Sie spielen in den Beiträgen über „[I]lexikographische Prozesse bei Internetwörterbüchern“ (Hildenbrandt/Klosa 2016) keine Rolle. Auch Tarp (2015, 20) fordert zwar, den Blick auf eine „extra-lexicographical pre-consultation phase“ zu richten, stellt aber nicht die Frage, wie eine digitale Sprachressource überhaupt in den Fokus von potentiellen Nutzenden gelangt. Dziemianko (2012, 319) orientiert sich beim Vergleich von Print-Wörterbüchern und elektronischen Wörterbüchern an den Kategorien „decoding, encoding, speed, look-up frequency, learning and appreciation“. Wie bei den anderen Autorinnen und Autoren wird der Prozess des Auffindens eines geeigneten Werkes bei ihm ausgeklammert. Damit können auch die vermarktungsspezifischen Eigenschaften digitaler Sprachressourcen nicht herausgearbeitet werden, obwohl sie bis in die Artikelgestaltung hinein relevant sein können (s.u.). Bei Printwerken delegierte man die Aufgabe der Vermarktung an den Verlag, bei Online-Ressourcen gehen viele davon aus, dass potentielle Nutzende das Werk schon finden werden. Im Prinzip kann man den Erfolg dieser Annahme sehr gut an den Klickzahlen ablesen. Die Frage ist nur: Welche Zahlen sind erstrebenswert und wie bewertet man sie? Der gefährliche Trugschluss kann sein, dass man geringe Klickzahlen als Qualitätskriterium einer digitalen Sprachressource heranzieht, die Ursache aber in einer schlechten Auffindbarkeit beziehungsweise Vermarktung liegt.

## 2.4 Mehr als Klickzahlen

Wenn Suchmaschinen den größten Anteil an Besuchenden auf digitale Sprachressourcen führen, liegt es nahe, sich um eine gute Positionierung im Ranking der Suchmaschinen zu bemühen. Selbstverständlich sollen Klickzahlen nicht den Wert wissenschaftlicher Forschung bestimmen. Deshalb sollten sie nicht – oder zumindest nur sehr vorsichtig und kundig interpretiert – zur Bewertung wissenschaftlicher digitaler Sprachressourcen eingesetzt werden (vgl. Müller-Spitzer 2016b, 81). Doch bei aller Vorsicht und Distanzierung muss man auch sehen: Hinter jedem von Menschenhand erzeugtem Klick auf einen lexikographischen Artikel steckt eine potentiell am Inhalt interessierte Person. Und genau darum geht es bei der Suchmaschinenoptimierung: zu ermöglichen, dass Menschen mit einem relevanten Interesse das in digitalen Sprachressourcen vorhandene und zu ihrer Frage passende Wissen auch finden. Welche Lexikographin und welchen Lexikographen freut es nicht, zu erfahren, dass das sorgfältig zusammengetragene und dokumentierte Wissen auch Anklang findet? Enttäuschend hingegen ist, wenn trotz vorhandenem Interesse an einem Thema ein weniger passendes, populäres Informationsangebot aufgrund einer besseren Vermarktung in den Vordergrund drängt – das heißt bei *Google* besser gefunden wird – und dies die Rezeption eines wissenschaftlichen Werkes verhindert. Weiter zuspitzen könnte sich das in Zukunft, wenn die wissenschaftlich erarbeiteten Informationen von populären Online-Angeboten kopiert werden und diese dann möglicherweise sogar mit den fremden Informationen Werbeeinnahmen erzielen.

Der Wert der wissenschaftlichen Lexikographie ist die Verlässlichkeit und Nachvollziehbarkeit der präsentierten Informationen und die Bearbeitung von weniger populären, aber nicht weniger wichtigen Themenfeldern. Diesen Mehrwert gilt es im Rahmen der Publikation wissenschaftlich-lexikographischer Projekte herauszustellen und nachvollziehbar zu machen (vgl. hierzu Kap. 4.2.4). Untersuchungen zeigen, dass sich genau diese Merkmale sorgfältig erarbeiteter Sprachressourcen mit den primären Interessen und Erwartungen potentieller Nutzerinnen und Nutzer zu decken scheinen (vgl. Müller-Spitzer/Koplenig 2014; Adarve 2020). So schreibt Müller-Spitzer (2016a, 311f.) in einer Zusammenfassung ihrer Untersuchung von 2014, dass auch im Digitalen die inhaltliche Verlässlichkeit, eine regelmäßige Aktualisierung, die Übersichtlichkeit und eine längerfristige Erreichbarkeit – also klassische Merkmale von Wörterbüchern – hoch bewertet würden. Man kann die Situation so zusammenfassen: **Digitale Sprachressourcen müssen als verlässlich und als aktuell erkannt werden, in der Benutzung müssen sie verständlich und übersichtlich sein und ihre Inhalte müssen für die Nutzenden wertvoll und nützlich sein. Und vor all dem müssen sie aufgefunden werden.**

### ***INFO: Google – Das Dilemma der Publikation von digitalen Sprachressourcen***

Mit den gedruckten Werken fallen Distributionswege für Sprachressourcen weg. Der von der Reichweite vielversprechende, aktuell propagierte Ansatz heißt *Google* – eine in vielerlei Hinsicht problematische Suchmaschine. Es ist selbstverständlich nicht das Anliegen dieses Textes, eine Lanze für *Google* zu brechen. Die Intention ist, einen Publikationsweg für die wissenschaftliche Lexikographie der Zukunft zu skizzieren. Stand heute spielt *Google* dabei eine große Rolle, und die dabei wirksamen Mechanismen sind mit Sicherheit auch Teil der Informationssuche der Zukunft. An dieser Stelle sei auf den Sammelband „Die Googleisierung der Informationssuche. Suchmaschinen zwischen Nutzung und Regulierung“ (Stark et al. 2014) verwiesen, der als kritische Begleitlektüre zu diesem Text herangezogen werden kann. Aus pragmatischen Gründen (bspw. wegen der Reichweite und der Marktbeherrschung) steht in diesem Text *Google* oftmals stellvertretend für

Suchmaschine. Eine objektive Betrachtung der technischen Aspekte zeigt eine hohe Passung: *Google* will wertige, für die Suchenden nützliche Informationen präsentieren, Sprachressourcen enthalten wertige Informationen. Dass es auch eine problematische Kehrseite dieses Verhältnisses gibt, ist evident. Mit der Nennung einiger Kapitelüberschriften aus dem Buch von Stark et al. soll kurz angedeutet werden, wie vielfältig die Themen sind, die im Kontext der Publikation digitaler Sprachressourcen zu diskutieren wären: „Die Macht von Google und die Ohnmacht der Nutzer und Regulierer“, „Wie Nutzer im Suchprozess gelenkt werden. Zwischen technischer Unterstützung und interessengeleiteter Darstellung“, „Informationskompetenz als notwendige Voraussetzung zur Nutzung von Suchmaschinen“, „Gefangen in der Filter Bubble? Search Engine Bias und Personalisierungsprozesse bei Suchmaschinen“, „Ideologie des Algorithmus. Wie der neue Geist des Kapitalismus Suchmaschinen formt“.

## 2.5 Informationen als Wert der wissenschaftlichen Lexikographie

Der Wert der Lexikographie steckt in den erarbeiteten Informationen, also den durch wissenschaftliche Forschung verknüpften sprachlichen und außersprachlichen Daten (bspw. Wortform und Bedeutung). Genau diese Verknüpfung von zusammenhängenden und nützlichen Daten ist die Erfolgsformel zahlreicher Digital-Angebote, allen voran des Konzerns *Alphabet*, zu dem *Google* gehört. Wissenschaftliche Sprachressourcen enthalten solche verknüpften Daten, teilweise mit großem Mehrwert: Sie sind oft die einzige Quelle für diese Informationen. Außerdem ist in die lexikographischen Informationen die analoge Kompetenz der Redaktion eingeflossen, weshalb sie sich nicht per se algorithmisch reproduzieren lassen. Doch selbstverständlich kann die einmal publizierte Information kopiert werden – ein Vorgang, der die Geschichte der Lexikographie nicht unwesentlich geprägt hat und wahrscheinlich weiterhin prägen wird.<sup>12</sup> Es ist daher nötig – wenn die Lexikographie langfristig die Lorbeeren für ihre Arbeit nicht anderen überlassen möchte –, einen einfachen Zugang zu diesen Informationen zu ermöglichen. Dieses Anliegen, den Zugang zu den Informationen einer Sprachressource zu erleichtern, ist derzeit ein wichtiger Forschungsgegenstand der Wörterbuchbenutzungsforschung.

## 3 Wie werden digitale Sprachressourcen genutzt?

Zweck der Wörterbuchbenutzungsforschung ist es, „zukünftige Wörterbücher besser den Bedürfnissen der Nutzerinnen und Nutzer anzupassen.“ (Müller-Spitzer 2016a, 292) Daher plädiert Müller-Spitzer (2016a, 335f.) dafür, nicht nur die eigentliche Nutzung zu untersuchen, sondern schon die sprachlichen Schwierigkeiten in den Blick zu nehmen, die letztlich zu einem Nachschlagebedürfnis führen (vgl. auch Tarp 2015, 20). Eine solche Übernahme der Perspektive potentieller Nutzerinnen und Nutzer steht allerdings in Gegensatz zu dem dokumentarischen Ansatz vieler Nachschlagewerke und ihrer Redakteurinnen und Redakteure. Dieser Ansatz ist nicht am Nutzenden orientiert, sondern am sprachlichen Phänomen als solchem. Auch wenn es Wiegand (1998, 54) folgend heute unbestritten scheint, dass Wörterbücher als Gebrauchsgegenstände zu betrachten sind (vgl. Müller-Spitzer 2016a, 292), bleibt eine Diskrepanz: Der Gegensatz zwischen einer möglichst vollständigen Dokumentation von Sprach- und Sprachgebrauchswissen einerseits und der Aufbereitung eines Wissens für die Bedürfnisse und Interessen der Nutzerinnen und Nutzer andererseits prägt auch die Situation digitaler Sprachressourcen. Die digitalen Informationsangebote, die im Rahmen des DL-Projekts

---

<sup>12</sup> In Dräger (2010) habe ich diese Tradition und ihre Problematik des Abschreibens für den Bereich der phraseologischen Nachschlagewerke aufgezeigt.

analysiert wurden, entsprechen noch immer der traditionellen Vorstellung eines Werkes, das in seinen Artikeln Sprachwissen dokumentiert.

Digitale Sprachressourcen unterscheiden sich deutlich von anderen digitalen Angeboten, die konzentriert und fokussiert auf die Informationsbedürfnisse von Nutzenden ausgerichtet sind (bspw. Buchungsportale, Online-Shops oder Zeitungsportale; vgl. hierzu Kap. 6). Hinsichtlich der Nutzendenorientierung ist es allerdings ein Unterschied, ob man eine Struktur an den tatsächlich Nutzenden ausrichtet oder an den potentiell Nutzenden, die womöglich ganz andere Bedürfnisse und Gewohnheiten mitbringen. Müller-Spitzer (2016a, 336) deutet dieses Potential an, wenn sie schreibt: Es „werden sehr viele sprachliche Fragen im Internet recherchiert – vielleicht oder sehr wahrscheinlich mehr als es Nachschlagehandlungen in gedruckten Wörterbüchern gab. Die Frage ist also, wie man diese Nutzungshandlungen besser mit den vorhandenen lexikografischen Ressourcen zusammenbringen kann.“ Mir ist bislang keine digitale Sprachressource bekannt, die konsequent nach den Bedürfnissen und Gewohnheiten der im Internet Recherchierenden konzipiert wurde, also einer Gruppe, die weit über die tatsächlich Benutzenden hinausreicht. Letztlich scheint das Anliegen, einen bestimmten Ausschnitt einer Sprache oder eines Sprachgebrauchs zu dokumentieren, noch immer leitend in der Lexikographie. Die Adaption der erarbeiteten Inhalte an die Nutzenden ist dabei sekundär. Müssen wir also mehr über die Nutzerinnen und Nutzer erfahren?

### 3.1 Für wen konzipiert man eine Sprachressource?

Ein Wörterbuch wird oftmals direkt auf ein Zielpublikum hin konzipiert. Sprachlernende sind nicht das einzige, aber ein naheliegendes und bislang gut untersuchtes Wörterbuchpublikum, weil man ihre Interessen und Bedürfnisse relativ gut nachvollziehen kann (vgl. Müller-Spitzer 2016a, 293). Im Vorfeld der Konzeption eines vergleichsweise neuartigen Wörterbuchtyps, einem Kollokationenwörterbuch<sup>13</sup>, konnte beispielsweise durch Benutzungstests mit Personen aus der Zielgruppe herausgefunden werden, dass die ursprünglich geplante und in vergleichbaren Werken praktizierte Struktur nicht zur Zielgruppe passt: Die dokumentarische, an der syntaktischen Konstruktion der Wortverbindungen orientierte Mikrostruktur (bspw. werden alle Akkusativ-Verb-Kombinationen gruppiert), wie sie in schon bestehenden Kollokationen-Wörterbüchern zu finden war<sup>14</sup>, war für die Zielgruppe in den Tests nicht hilfreich. Wesentlich besser funktionierten die Testaufgaben zur Textproduktion und stilistischen Textüberarbeitung mit einer an der Semantik der Wortverbindung orientierten Gruppierung (vgl. Dräger et al. 2014). Genau diese Struktur wurde in Rezensionen<sup>15</sup> lobend hervorgehoben, was ein nachdrücklicher Hinweis darauf ist, dass es nicht die dokumentarische Logik des Wörterbuchgegenstandes sein sollte, sondern die Erkenntnisinteressen und die Bedürfnisse des Zielpublikums, die konstituierend für die lexikographischen Strukturen sein sollten. Doch das *Kollokationenwörterbuch* wurde bewusst als gedrucktes Werk mit einer klaren Zielgruppe konzipiert, bei vielen Online-Sprachressourcen ist das so nicht möglich und auch nicht sinnvoll.

Doch ist es überhaupt ratsam, sich beim Erstellen von Online-Sprachressourcen auf ein Zielpublikum zu fokussieren, wenn vorab unklar ist, welche Personen mit welchen Interessen und Bedürfnissen das Angebot schließlich nutzen werden? Für die *Variantengrammatik* ließ sich beispielsweise feststellen,

---

<sup>13</sup> Vgl. [www.kollokationenwoerterbuch.ch](http://www.kollokationenwoerterbuch.ch) bzw. Häcki Buhofer et al. (2014).

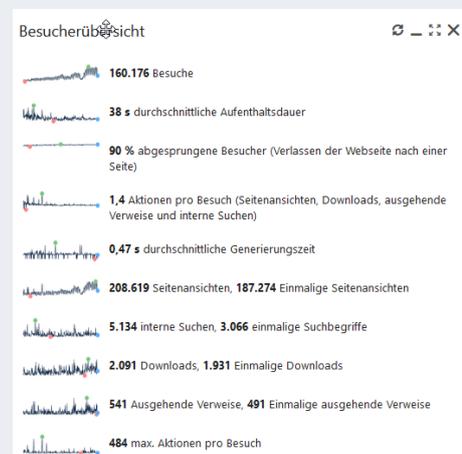
<sup>14</sup> Die beiden Referenzwerke waren McIntosh et al. (2009) und Quasthoff (2010).

<sup>15</sup> <https://kollokationenwoerterbuch.ch/web/aktivitaeten/presse/>, 29.09.2020.

dass über Suchmaschinen scheinbar unklare Schreibungen wie „öfter/öfters“ oder „Ellenbogen/Ellbogen“ gesucht und häufig gefunden und angeklickt wurden. Diese Fragestellungen decken sich mehr oder weniger zufällig mit dem Anliegen des Nachschlagewerks, sie werden auch in den entsprechenden Artikeln geklärt, aber die *Variantengrammatik* wurde nicht im Hinblick auf solche orthographischen Fragestellungen ausgearbeitet. Andere Artikel (bspw. zu „weil im Nebensatz“) werden wiederum mit einem ganz anderen Interesse und möglicherweise auch von einem anders zusammengesetzten Nutzendenkreis konsultiert. Der Grund für diese Diversität der Nutzenden ist, dass die Menschen durch Suchmaschinen direkt den zu ihrem Frageinteresse passenden Artikel finden, ohne vorab zu klären, ob das dahinterstehende Nachschlagewerk thematisch passend ist.

### INFO: Log-File-Analysen – Wie man das Verhalten der Nutzenden beobachten kann

Log-File-Analysen sind eine Möglichkeit, um besser zu verstehen, wie Nutzende sich in einem Web-Angebot bewegen. Für die CMS (Content-Management-Systeme), mit welchen Websites in der Regel erstellt werden, gibt es Plugins für eine anschauliche und verständliche Betrachtung und Auswertung der Log-Files. Selbstverständlich kann man die Daten für komplexere Analysen auch exportieren. Im Projekt haben wir das Plugin *Matomo* verwendet, aus welchem auch die hier gezeigten Screenshots stammen. Auf diese Weise lassen sich einerseits die Möglichkeiten solcher Analyse-Plugins bzw. der Log-File-Analyse im Allgemeinen erkennen und andererseits interessante Erkenntnisse über das Beispielprojekt *Variantengrammatik* thematisieren. Die Besucherübersicht zeigt – hier für einen Zeitraum von einem Jahr – allgemeine Daten zu allen Besuchen auf der Website. Diese Daten lassen sich dann beispielsweise nach den Kanaltypen aufgliedern:



KANALTYP	BESUCHE	AKTIONEN	AKTIONEN PRO BESUCH	DURCHSCHNITTSZEIT AUF DER WEBSITE	ABSPRUNGRATE
Suchmaschinen	140.233	162.968	1,2	26 s	93 %
Direkte Zugriffe	16.573	37.418	2,3	1 Minuten 48s	77 %
Websites	2.235	14.071	6,3	4 Minuten 44s	35 %
Soziale Netzwerke	1.135	1.928	1,7	59 s	83 %

Diese einfache Auflistung informiert darüber, von welcher Seite die Besuchenden auf die Seite der *Variantengrammatik* geführt wurden, wie viele Aktionen sie ausgeführt haben und wie lange sie durchschnittlich auf der Website geblieben sind. Die Absprungrate gibt an, wie viele Nutzende nach der ersten Seite die Website wieder verlassen haben. Die Übersicht lässt vermuten, dass von Websites und über direkte Zugriffe offensichtlich ein interessierteres Publikum gekommen ist als über die Suchmaschinen und über die Sozialen Netzwerke.

Von anderen Projekten ist bekannt, dass viele Besuchende von Suchmaschinen weniger als drei Sekunden auf der Seite bleiben, also wurde die obige Aufstellung nochmals generiert und ein Filter (Dauer des Besuchs mehr als 3 Sekunden) angewendet.

Kanaltypen					
KANALTYP	▼ BESUCHE	AKTIONEN	AKTIONEN PRO BESUCH	DURCHSCHNITTZEIT AUF DER WEBSITE	ABSPRUNGRATE
Suchmaschinen	9.820	32.399	3,3	6 Minuten 8s	0 %
Direkte Zugriffe	3.689	24.426	6,6	8 Minuten 3s	0 %
Websites	1.448	13.281	9,2	7 Minuten 18s	0 %
Soziale Netzwerke	196	987	5	5 Minuten 44s	0 %
<b>Gesamttotal</b>	<b>15.153</b>	<b>71.093</b>	<b>4,7</b>	<b>6 Minuten 42s</b>	<b>0 %</b>

Jetzt stellt sich das Bild augenscheinlich anders dar. Aber wie ist diese Veränderung zu interpretieren? Deutlich erkennbar ist, dass sich die Durchschnittszeiten auf der Website angleichen, dass allerdings die Aktionen pro Besuch zwischen 3,3 und 9,2 noch stark variieren. Eine solche Übersicht gibt also eher Hinweise darauf, wo sich ein Nachforschen lohnen könnte, als dass sie Erklärungen liefert. Außerdem ist Vorsicht bei der Interpretation geboten: Der Rückgang um über 130.000 Besuche von Suchmaschinen, wenn man die Besuchsdauer (< 3 Sekunden) berücksichtigt, sticht sofort ins Auge. Die folgende Tabelle, in der wir den prozentualen Rückgang berechnet haben, relativiert diesen Eindruck und rückt eher die Entwicklung bei den Websites in den Fokus:

Kanaltyp	Besuche gesamt	Besuche (Dauer >3 Sek.)	Rückgang absolut	Rückgang relativ
Suchmaschinen	140.223 (88%)	9.820 (65%)	130.403	93%
Direkte Zugriffe	16.573 (10%)	3.698 (24%)	12.875	78%
Websites	2.235 (1,4%)	1.448 (10%)	787	35%
Soziale Netzwerke	1.135 (0,7%)	196 (1,3)	939	83%

Wie bei den Werten der Besuchsdauer deutet auch diese Tabelle auf ein anderes Besucherverhalten bei den Besuchenden hin, die einem Link auf einer Website gefolgt sind. Auch hier kann man genauer betrachten, wo diese Besuchenden herkommen:

WEBSITE	▼ BESUCHE	AKTIONEN	AKTIONEN PRO BESUCH	DURCHSCHNITTZEIT AUF DER WEBSITE	ABSPRUNGRATE
www.variantengrammatik.net	1.007	10.018	9,9	7 Minuten 50s	0 %
variantengrammatik.net	141	995	7,1	6 Minuten 8s	0 %
www.canoonet.eu	75	616	8,2	8 Minuten 9s	0 %
www1.ids-mannheim.de	45	306	6,8	3 Minuten 40s	0 %
www.canoo.net	42	281	6,7	3 Minuten 33s	0 %
www.srf.ch	31	271	8,7	4 Minuten 17s	0 %
canoo.net	30	314	10,5	9 Minuten 4s	0 %
canoonet.eu	17	68	4	5 Minuten 46s	0 %
de.wiktionary.org	8	130	16,3	9 Minuten 20s	0 %
165512.homepagemodules.de	6	27	4,5	2 Minuten 50s	0 %

Es zeigt sich, dass eine Besonderheit der *Variantengrammatik* ursächlich für dieses Phänomen ist: Die Besuchenden kommen überwiegend von der Projektseite *www.variantengrammatik.net*, wo es ein

eigenes Suchfeld gibt, das dann auf die Sprachressource zugreift, die auf einem anderen Server<sup>16</sup> gehostet ist. Diese Besuche sind den direkten Zugriffen von der Art her sehr ähnlich, aber die Log-Files zeigen, dass die Besuchenden sich auf der Website anders verhalten (vgl. dazu Kap. 5.3.2.1).

Über die Log-Files ließ sich außerdem herausfinden, welche Artikel dabei wie oft besucht werden oder welche Aktionen (Seitenbesuche) hintereinander ausgeführt wurden. Es ist dabei stets zu bedenken, dass man es mit Durchschnittswerten zu tun hat und nicht weiß, wie Menschen tatsächlich vor dem Bildschirm agieren. Angenommen, jemand öffnet eine Website und nimmt sofort danach ein Telefonat an. Nach einer Viertelstunde am Telefon schließt die Person den Browser. In den Log-Files erscheint dann die geöffnete Seite mit einer 15-minütigen Nutzungsdauer. Gerade in Bezug auf die Nutzungsdauer einer Website beziehungsweise ihrer Einzelseiten lässt sich aufgrund dessen wenig ableiten. Um einen besseren Einblick zu bekommen, wurden deshalb die Besuchszahlen der Variantengrammatik in Relation zur Anzahl der Interaktionen und der Gesamtbesuchsdauer gestellt.

Interaktionen	2	3	4	5	6	7	8	9	10-14
Besuchsdauer									
0-10s	420	44	2	0	0	0	0	0	0
11-30s	476	195	48	19	7	2	2	0	0
31-60s	342	146	83	43	20	15	4	5	3
1-2 Min.	375	146	90	62	33	31	24	17	20
2-4 Min.	331	111	53	57	38	33	27	15	62
4-7 Min.	198	63	39	28	29	31	13	23	54
7-10 Min.	154	36	18	15	10	13	10	15	42
10-15 Min.	157	39	19	12	13	9	5	8	32
15-30 Min.	325	69	30	21	27	21	23	11	50
30+ Min.	2	18	10	2	13	5	12	4	35

Hier zeigt sich, dass es je nach Anzahl der Interaktionen ein Zeitfenster gibt, in welchem signifikant mehr Besuche stattfinden als in den kürzeren oder längeren Zeitfenstern. Gleichzeitig wird im Zeitfenster 15-30 Min. eine signifikante, aber nicht erklärbare Erhöhung der Besuchszahlen sichtbar. Es lässt sich aber gut erkennen: Die Besuchsdauer steigt mit der Zahl der Interaktionen, während die Besuchsdauer pro Seite leicht abnimmt (was man an der abflachenden Kurve sieht). Obwohl mit der Datenlage möglich wäre, eine durchschnittliche Nutzungsdauer pro Seite zu errechnen, spricht diese Visualisierung dagegen.

Für das Projekt „Digitale Sprachressourcen“ ergaben sich daraus zwei wichtige erste Erkenntnisse:

- 1) Über Suchmaschinen kommen 88% Prozent der Besuche. Selbst wenn man nur Besuche mit mehr als drei Sekunden Besuchsdauer betrachtet, kommen 65% der Besuchenden über Suchmaschinen. Das sind fast doppelt so viele wie die Zahl derer, die die Website direkt aufsuchen.
- 2) Es wirkt sich signifikant auf die Nutzungsweise der Sprachressource aus, ob jemand direkt die Website besucht oder über eine Suchanfrage bei einer Suchmaschine auf die Website gelangt.

## 3.2 Suchmaschinen verändern die Nutzung und den Nutzendenkreis

Das Potential der Wörterbuchbenutzungsforschung (Befragungen, Log-File-Analysen oder beispielsweise auch Eye-Tracking-Studien) ist groß und wird ausführlich und überblickend von Müller-Spitzer (2016a) beschrieben. Wichtige Forschungserkenntnisse, die bei der Konzeption und

<sup>16</sup> Die Daten und das Wiki der *Variantengrammatik* wurden ab einem bestimmten Zeitpunkt während des Untersuchungszeitraums auf einem Server des *Leibniz-Instituts für Deutsche Sprache* gehostet. Die ursprüngliche Seite *variantengrammatik.net* fungierte ab diesem Zeitpunkt nur noch als Verlinkung.

Ausarbeitung von Nachschlagewerken zu beachten sind, betreffen u.a. die Stichwortauswahl (bspw. Koplenig et al. 2014), die Artikelpräsentation (Koplenig/Müller-Spitzer 2014) und generell die Wünsche und Meinungen der Nutzenden (Klosa/Koplenig/Töpel 2014). In der Regel gehen solche Untersuchungen von einer impliziten Annahme aus: Man nimmt an, dass die Benutzenden von Sprachressourcen sich bewusst sind, dass sie ein Nachschlagewerk nutzen, also gezielt ein Wörterbuch oder Lexikon aus dem Regal genommen oder die Internetpräsenz einer digitalen Sprachressource besucht haben. Diese Annahme muss im Hinblick auf einen großen Anteil an über eine Suchmaschine akquirierten Benutzenden zumindest für den Moment des Erstkontakts in Frage gestellt werden. Denn dieser Erstkontakt findet nicht auf der Website der Online-Ressource statt, sondern in der Trefferliste einer Suchmaschine. Eine Suchanfrage bei *Google* nach „ekelig / eklig“ brachte beispielsweise folgende Trefferliste:<sup>17</sup>

#### Suchen | ekelig [selten] [eklig] - Duden

<https://www.duden.de> > [suchen](#) > [dudenonline](#) > [ekelig \[selten\] \[eklig\]](#) ▾

eklig, selten ekelig. Adjektiv - 1a. ekelerregend, widerwärtig, abscheulich; 1b. in seinem Auftreten unangenehm; gemein, ...2. sehr, ganz gehörig, tüchtig.

#### ekelig / eklig – Variantengrammatik des Standarddeutschen

[mediawiki.ids-mannheim.de](http://mediawiki.ids-mannheim.de) > [VarGra](#) > [index.php](#) > [Ekelig\\_ > \\_eklig](#) ▾

Die Adjektivvarianten ekelig und eklig werden mit oder ohne -e- in den Bedeutungen 'ekelerregend' und 'unangenehm, merkwürdig' nebeneinander verwendet.

#### Ekelig oder Eklig? (Schule, deutsch, Rechtschreibung) - Gutefrage

<https://www.gutefrage.net> > [frage](#) > [ekelig-oder-eklig](#) ▾

10 Antworten

16.11.2016 - Beide Varianten sind möglich, allerdings verwendet man eher eklig. Der Ausdruck ekelig ist veraltet.

#### DWDS – ekelig – Worterklärung, Grammatik, Etymologie u. v. m.

<https://www.dwds.de> > [ekelig](#) ▾

Ekel · **ekeln** · ekel · ekelhaft · **eklig**. Ekel m. 'Abscheu, heftiger Widerwille', im 16. Jh. vom Md. ausgehend (öfters bei Luther); wohl Substantivbildung zu **ekeln** (s.

#### eklig | Rechtschreibung - wie schreibt man eklig | [https://www ...](https://www.wissen.de)

<https://www.wissen.de> > [rechtschreibung](#) > [eklig](#) ▾

e k | lig < Adj. > ekelhaft; <auch> ekelig... ... ekelhaft, eklig, Ekel erregend, abscheulich,

Bei keinem der Einträge wird unmittelbar klar, dass es sich bei der verlinkten Seite um ein Online-Nachschlagewerk handelt. Selbst bei *duden.de* lässt sich das nur über das Allgemeinwissen zum *Duden-Verlag* erschließen, ein ausformulierter Hinweis findet sich nicht. Dementsprechend muss man davon ausgehen, dass Personen beim Betrachten der Suchergebnisse und auch bei einem möglichen anschließenden Besuch einer verlinkten Seite nicht sofort wissen, um welche Form einer Wissensressource es sich handelt. Zudem muss man – nicht zuletzt auch aufgrund eigener Erfahrungen mit *Google* – annehmen, dass für Personen, die beispielsweise herausfinden wollten, ob man „ekelig“ oder „eklig“ verwendet, die Informationen in der Ergebnisliste ausreichend sind: „eklig, selten ekelig“ (*duden.de*), „ekelig und eklig werden mit oder ohne -e- [...] nebeneinander verwendet“ (*Variantengrammatik*) oder „Beide Varianten sind möglich, allerdings verwendet man eher eklig“ (*gutefrage.net*). Wenn diese Informationen als Antwort für das Frageinteresse ausreichen, werden die Personen die Websites der Nachschlagewerke möglicherweise gar nicht besuchen. In einer Umfrage zur

<sup>17</sup> Die Trefferzusammenstellung beruht auf komplexen Algorithmen und ist daher bei jeder Anfrage unterschiedlich und nicht prinzipiell reproduzierbar.

Nutzung von digitalen Sprachressourcen (vgl. Adarve 2020) gaben 67% der Teilnehmenden an, dass ihre Suchanfragen immer oder meistens direkt auf einer Suchmaschinenseite beantwortet würden. Das ist eine neue Situation, und zwar eine sehr bedeutsame: **Potentielle Nutzerinnen und Nutzer bekommen von einem Nachschlagewerk eine Frage beantwortet, ohne es merklich zu benutzen.** Die Benutzung der Sprachressource bleibt unbemerkt und unerkannt und zwar sowohl für die suchende Person als auch für die Betreiberinnen und Betreiber der Website, die diese im Open-Access zur Verfügung stellen. Das heißt: Wenn man Benutzendenforschung nur auf der Basis von Besucherinnen- und Besucherzahlen beispielsweise der *Variantengrammatik* betreibt, dann bleiben diese Nutzerinnen und Nutzer, die schon mit der von *Google* präsentierten Information zufrieden waren, unbemerkt und unberücksichtigt. Sie erscheinen nicht in den Log-File-Statistiken. Startet man eine Umfrage zum Nachschlagen, ergibt sich ein ähnliches Bild: Personen, die eine Sprachressource nur über eine Suchmaschine nutzen, können gar keine Auskunft über die Benutzung geben, weil ihnen weder bewusst ist, dass sie Sprachressourcen benutzten, noch erinnern sie sich andernfalls daran, welche das waren. Selbst Untersuchungen wie von Lorentzen/Theilgaard (2012), die anhand der Log-Files die Eingaben in Suchmaschinen analysieren, haben nur jene Nutzenden im Fokus, die tatsächlich auch die Website der Sprachressource besucht haben.

Das Wichtigste am geschilderten Beispiel ist aber: Die digitalen Sprachressourcen, sofern sie wie im obigen Beispiel zur Klärung der Frage der Suchenden beitragen, erfüllen ihren Zweck. Oder um es deutlicher auszudrücken: **Digitale Sprachressourcen informieren (wahrscheinlich sehr häufig) über ein Sprachphänomen, ohne dass sie im Verständnis der Benutzendenforschung überhaupt benutzt wurden.**

### 3.3 Benutzungsforschung für digitale Sprachressourcen neu denken

Der Erforschung der Nutzung und der Nutzenden von digitalen Sprachressourcen sollte also ein Kapitel hinzugefügt werden, das den digital durchklickten Weg von einer Suchanfrage gegebenenfalls über verlinkte Zwischenseiten bis zur Sprachressource mit in den Blick nimmt. Einen Schritt weiter gedacht müsste man sogar attestieren, dass genau jene Sprachressource die benutzendenfreundlichste ist, die die Frage der Suchenden schon mit dem Ausschnitt in der Ergebnisliste einer Suchmaschine hinreichend und zufriedenstellend beantwortet. Und noch weiter gedacht müsste für die geschilderte und zahlenmäßig möglicherweise sehr häufige Nutzung von digitalen Sprachressourcen zur Debatte gestellt werden: Ist es überhaupt zutreffend, von Nutzerin, Nutzer und einer Benutzung zu sprechen, wenn diese Personen mit dem eigentlichen digitalen Werk gar nicht mehr in Kontakt kommen? Und wie kann man ein Werk, das zugespitzt gesagt nicht mehr benutzt wird, für die Benutzung optimieren?

Das Beispiel und die daran geknüpften Anregungen zeigen, dass die Digitalisierung das Nachschlagen und Auffinden von Sprachinformationen – und natürlich auch anderen Informationen – viel stärker verändert, als es sich durch die Analyse von Eingaben und Bewegungen in einer Online-Ressource erkennen lässt. Das sollte stets mitgedacht werden.

### 3.4 Datenbanken können mehr als statische Texte ausgeben

Untersucht man Daten zur tatsächlichen Benutzung einer digitalen Sprachressource, beispielsweise durch die Analyse von Log-Files, dann stößt man sehr schnell an die Grenze der validen Dateninterpretationen (vgl. die Infobox auf S. 15). Für die *Variantengrammatik* wurde analysiert, welche

Seitentypen (bspw. Suchseite, Artikelseite, Startseite) die Besuchenden nacheinander aufrufen. Die Frage war, ob sich typische Benutzungsmuster identifizieren lassen, für welche man die Strukturen der Online-Sprachressource optimieren könnte. Tendenziell kann man daraus zwar Ansätze für strukturelle Verbesserungen ableiten, wobei sich aber auch die Annahme bestätigt, dass die Interessen und Verhaltensweisen der Besucherinnen und Besucher sehr unterschiedlich gelagert sind. So wertvoll eine umfangreiche Untersuchung der Frageinteressen wäre, so sehr ist zu befürchten, dass sich nur sehr inhomogene Erkenntnisse und Handlungsoptionen für die Lexikographie daraus ableiten lassen. Außerdem besteht die in Kap. 2.2 erwähnte Gefahr, sich an Informationen aus einer Filterblase zu orientieren, wodurch man möglicherweise einen Großteil von Benutzenden nicht erreicht oder gar ausschließt. Deshalb lohnt es sich, in den Blick zu nehmen, dass digitale Sprachressourcen strukturell gar nicht aus „ganzen“ Artikeln bestehen müssen. Es sind in der Regel datenbankbasierte Werke mit mehr oder weniger kleinteilig gespeicherten Informationseinheiten, deren Systematik sich an den lexikographischen Textsegmenten (vgl. Wiegand 1989, 33) von Wörterbüchern orientiert. Es ist also technisch gesehen keinesfalls nötig, für die Anzeige der Inhalte einen fixen Artikel bzw. eine fixe Reihenfolge der einzelnen Informationen zu definieren.

Ein naheliegender Ansatz könnte daher in der Abkehr vom statischen Wörterbuchartikel liegen zugunsten einer an die Suchanfrage angepassten dynamischen Ergebnisdarstellung auf der Basis einer entsprechend konzipierten lexikographischen Datenbank. An anderer Stelle habe ich dies so dargelegt: „Das heißt, wenn sich die Parameter der Suchanfrage verändern, verändert sich auch das ausgegebene Ergebnis. Ein solches Ziel lässt sich nur mit einem dynamischen System verwirklichen, welches nicht vorgefertigte Antworten ausgibt (wie ein Wörterbuch), sondern die Antwort erst aufgrund der Parameter der Anfrage aus den vorstrukturierten Daten generiert“ (vgl. Dräger 2011, 225).<sup>18</sup> Diese Idee hat Storrer schon 2001 mit dem Verweis auf erste Prototypen unter dem Schlagwort „Wörterbuchartikel ‚on demand‘“ ausformuliert: „Es ist gerade die Stärke von Hypertextsoftware, aus ein und demselben Datenpool die Informationen herauszugreifen und in geeigneter Weise zu präsentieren, die für einen bestimmten Nutzungskontext typischerweise relevant sind“ (Storrer 2001, 64). Heute, gut 20 Jahre später, sollte man den Blick weiten und in Betracht ziehen, dass ein Großteil der sprachbezogenen Suchanfragen nicht direkt an Wörterbücher gerichtet wird, sondern an Suchmaschinen. Bei der *Varietengrammatik*, einem sehr jungen und noch wenig bekannten Nachschlagewerk, liegt der Anteil der von einer Suchmaschine kommenden Zugriffe derzeit bei rund 88%.<sup>19</sup> Fast alle dieser Besuchenden haben in der Suchmaschine nicht nach der *Varietengrammatik* gesucht, sondern nach Informationen zu einem konkreten, meist sprachlichen Phänomen. Es wurde also keine Sprachressource gesucht, sondern eine Information (vgl. dazu auch Kap. 5.1.1).

### 3.5 Kein Widerspruch: schneller Informationsgewinn und Qualität der Information

Wie muss ein Online-Nachschlagewerk gestaltet sein, damit es den Bedürfnissen der Nutzenden entspricht? Dies ist eines der Erkenntnisinteressen der Wörterbuchbenutzungsforschung. Dabei kommt es allerdings zu Diskrepanzen bei der Interpretation der Beobachtungen und Analysen. Wie Müller-

---

<sup>18</sup> So habe ich es in meiner Dissertation zu einem *Online-Lexikon zur diachronen Phraseologie* formuliert. Das Ergebnis eines sich daran anschließenden SNF-geförderten Forschungsprojekts ist unter [www.oldphras.net](http://www.oldphras.net) online.

<sup>19</sup> Bei etablierten und bekannteren Werken haben die direkten Zugriffe über die Website möglicherweise einen größeren Anteil, aber die mittelbaren Zugriffe über Suchmaschinen sind mit Sicherheit nicht vernachlässigbar.

Spitzer (2016a, 20) schreibt, ist unklar, ob ein benutzerfreundliches Wörterbuch vor allem ein flexibles (de Schryver 2003, 182) oder ein schnelles (Almind 2005, 39; Bergenholtz 2011) oder ein verlässliches und wertiges sein sollte. Diese Diskrepanz entsteht vor allem dann, wenn man Artikel als Texte mit festen, vordefinierten Strukturen vor Augen hat. In Bezug auf die einzelne Information muss man die genannten Ansichten jedoch keinesfalls als Widerspruch auffassen: **Was spricht dagegen, dass jeweils die relevanten wertigen und verlässlichen lexikographischen Informationen von den Nutzenden möglichst schnell gefunden werden können?** Meiner Ansicht nach weiß man über die Nutzenden von digitalen Nachschlagewerken zu wenig, um sich nach ihren Interessen zu richten. Hierfür gibt es einige Gründe: Mit Untersuchungen, wie wir sie auch selbst durchgeführt haben, erreicht man einen erheblichen Teil der Personen nicht, die über Suchmaschinen auf digitale Sprachressourcen stoßen. Von den Besuchenden, deren Verhalten man via Log-File-Analyse „beobachtet“, weiß man nur ungefähr, welches Erkenntnisinteresse sie hatten. Es bleibt also unklar, mit welcher Zufriedenheit eine Person die Website verlässt und ob sie noch andere Quellen konsultiert. Die Log-Files zeigen, dass 90,5% der Besuchenden die Website innerhalb von drei Sekunden wieder verlassen (vgl. die Infobox auf S. 15).

Am Beispiel dieser Besuche unter drei Sekunden lässt sich aufzeigen, weshalb es schwierig ist, auf der Basis bisheriger Erkenntnisse Entscheidungen zur Usability von digitalen Nachschlagewerken zu treffen. Es wurde meines Wissens bisher nicht untersucht, was diese Personen in den wenigen Sekunden ihres Besuchs wahrnehmen und warum sie nur so kurz verweilen. Eine Annahme könnte folgende sein: Die ursprüngliche Frage wurde schon durch den kleinen Websiteausschnitt im Suchergebnis bei *Google* beantwortet. Aber die Person möchte wissen, ob die Information auch verlässlich ist, weshalb sie den Link anklickt. Mit einem Blick, der drei Sekunden nicht überschreiten muss, sieht sie das Artikellayout, das Logo und die bibliographischen Angaben und ist zufrieden. Eine andere Annahme könnte sein, dass die Person mehr über das sprachliche Phänomen erfahren möchte, sich aber von dem längeren Artikeltext nicht angesprochen fühlt und deshalb weiterklickt. Dann wäre es doch gerade interessant, sich darum zu bemühen, diese Person für ein Weiterlesen zu gewinnen. Man kann also – ohne weiteres Wissen über die Benutzenden – diese Unter-drei-Sekunden-Besuche als irrelevant aus der Erforschung des Benutzendenverhaltens ausklammern. Genauso gut kann man sich für das Gegenteil entscheiden und sie in den Fokus der Usability-Bemühungen rücken (zur Usability vgl. v.a. Kap. 5.2.2).

Es hat sich im bisherigen Text angedeutet, dass man Sprachressourcen in klassischer Tradition als Werke mit Artikeln betrachten kann und dass im Digitalen einzelne Informationen losgelöst vom Werk betrachtet werden können. Diese beiden unterschiedlichen Ansichten wirken sich erheblich auf die Nutzung, Gestaltung und auch Publikation digitaler Sprachressourcen aus. Dies wird im nächsten Kapitel aufgezeigt.

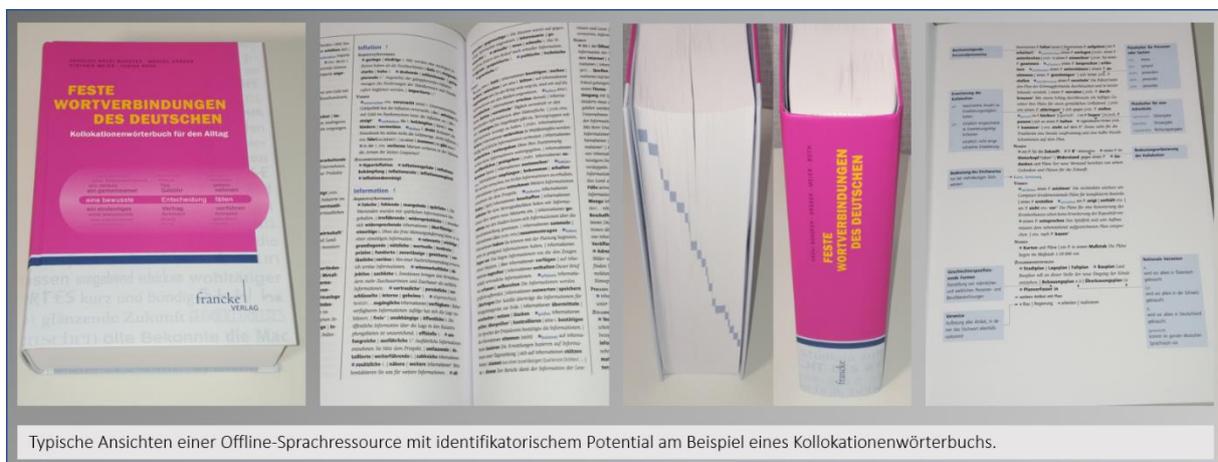
# 4 Werk oder Informationsspeicher? – Ein neuer theoretischer Ansatz für die digitale Lexikographie

## 4.1 Vom Werk zur digitalen „Ressource“

### 4.1.1 Nachschlagen im Wörterbuch hat die Lexikographie geprägt

Der direkte Online-Zugriff auf Informationen aus einer digitalen Sprachressource in der Suchmaschine ist eine neue<sup>20</sup> Situation, die weiterer Überlegungen bedarf. In den Fokus rücken sollte, dass das eigentliche Werk nur mittelbar benutzt wird. Bei einem gedruckten und gebundenen Nachschlagewerk – selbst bei einer CD-ROM-Ausgabe dieses Werks – ist kaum vorstellbar, dass Nutzende Informationen entnehmen, ohne zu erkennen, dass es sich um ein Nachschlagewerk handelt.

Die folgenden Abbildungen zeigen, dass ein gedrucktes Wörterbuch zahlreiche Signale enthält, die es auch für Fremde sofort als ein Nachschlagewerk erkennbar werden lassen. Seine Identität als ein Werk zum Nachschlagen wird beim ersten Kontakt klar, selbst dann, wenn es im Regal steht. Dafür sorgen unter anderem der Buchrücken, die Buchdicke, der Titel, die Namen des Verfassendenkollektivs, der Verlag, die Umschlaggestaltung, der Klappentext, der Hinweis auf eine innere Ordnung an den Seiten und, sofern man es aufschlägt, die Benutzungshinweise im Einband sowie nicht zuletzt eine klar erkennbare Artikel-Mikrostruktur.



Typische Ansichten einer Offline-Sprachressource mit identifikatorischem Potential am Beispiel eines Kollokationenwörterbuchs.

All diese Hinweise deuten darauf hin, dass es sich um ein Nachschlagewerk handelt. Das heißt: Beim Nachschlagen im Buch wird eine Sprachressource zwangsläufig zuerst und maßgeblich als Werk wahrgenommen. Entsprechend dieser Wahrnehmung nutzen Personen ihre Kompetenzen, die für das erfolgreiche Auffinden einer Information in der Makro- und Mikrostruktur eines Werkes relevant sind.<sup>21</sup> Die nötigen Kompetenzen für ein erfolgreiches Nachschlagen reichen dabei weit über das Blättern oder Klicken hinaus. Ausgehend von einem Erkenntnisinteresse (beispielsweise: Gibt es alternative Formulierungen für den Ausdruck „Informationen weitergeben“?) gilt es, eine zielführende Frage zu

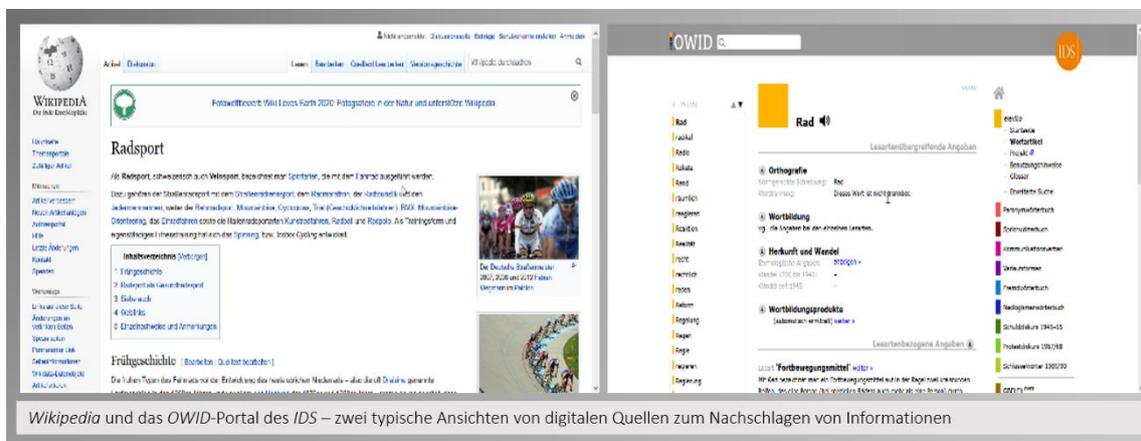
<sup>20</sup> Neu ist die Situation weniger in zeitlicher Hinsicht, schließlich gibt es digitale lexikographische Projekte schon länger. Neu ist sie deshalb, weil man sich bisher kaum damit beschäftigt hat.

<sup>21</sup> Für diesen Nachschlageprozess ist die Optimierung der sowohl digitalen als auch analogen Strukturen mithilfe der Erkenntnisse der Wörterbuchbenutzungsforschung notwendig.

formulieren, ein entsprechendes Themenfeld zu definieren (hier: Stilistik oder Wortverbindungen), ein zu diesem Thema passendes Werk zu identifizieren (hier ein Kollokationenwörterbuch), dieses digital oder physisch zu konsultieren und die relevante Information zu finden, um daraus im günstigen Fall eine zufriedenstellende Antwort abzuleiten.



Übertragen auf die Nutzung von Online-Ressourcen ist der Vorgang vergleichbar: Auch digitale Sprachressourcen haben solche Merkmale, durch welche auf den ersten Blick die Werk-Identität offensichtlich wird. Man besucht die Website eines bekannten Werkes wie *duden.de* oder *Wikipedia* und wendet dann die dort angebotenen Recherchestrategien an. Doch welche Merkmale übernehmen auf den Websites digitaler Sprachressourcen die identifikatorischen und anderen Aufgaben von Buchtitel, Hardcover, Buchdicke und Co.? Dies müsste man separat untersuchen. Auf der Basis eines cursorischen Vergleichs bekannter lexikographischer Ressourcen lässt sich die Hypothese formulieren, dass das *Wikipedia* oder *OWID* typische Online-Wissensressourcen sind, die damit auch die Identifikationsmerkmale solcher prägen (vgl. die Infobox auf S. 56).



Wikipedia und das OWID-Portal des IDS – zwei typische Ansichten von digitalen Quellen zum Nachschlagen von Informationen

Auch auf den Startseiten von digitalen Nachschlagewerken finden sich Werk-Merkmale wie Titel oder Informationen zur Erstellung, zu Autorinnen und Autoren oder auch Hinweise auf die Inhalte. Doch wer regelmäßig gedruckte Sprachressourcen nutzt(e), weiß, dass auch die Optik und Haptik des Einbandes, der Hinweis auf den (wissenschaftlichen) Verlag oder auch die Dicke (= Seitenzahl) zur Beurteilung der zu erwartenden Qualität und Verlässlichkeit der enthaltenen Informationen beitragen. Manche dieser Bewertungskriterien fehlen oder sind nicht sofort zu erkennen, wenn man die Startseite einer digitalen Sprachressource anschaut. Durch die Startseite ist beispielsweise nicht ersichtlich, welchen Umfang das digitale Werk hat.

### **INFO: Recherchekompetenz durch Forschendes Lernen**

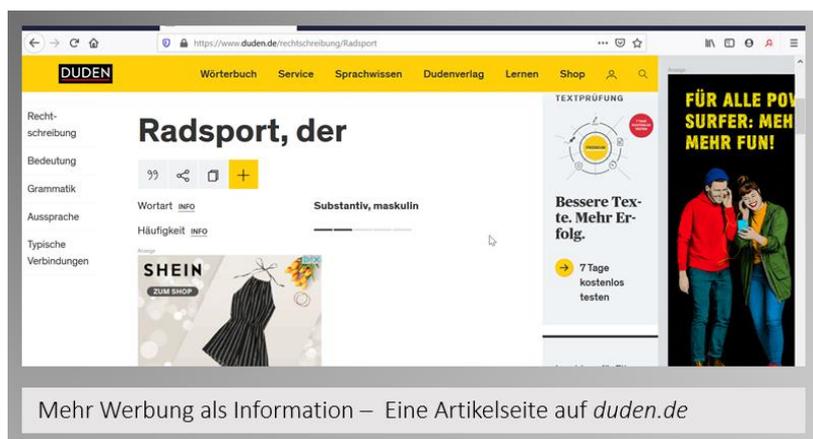
Im Projekt „Digitale Sprachressourcen“ wurde eine Publikation erarbeitet, in der Dürscheid/Sutter (2020) aufzeigen, wie „Forschendes Lernen mit digitalen Sprachressourcen“ – so der Titel – zu einer allgemeinen Ausbildung von Recherchekompetenzen führen kann. Die Autorinnen (ebd., 20f.) stellen heraus, dass es Bestandteil des schulischen Unterrichts sein sollte, Recherchekompetenzen zu erwerben. Das bedeutet, Schülerinnen und Schüler sollen lernen und üben, wie sie Informationen finden, über die sie selbst nicht verfügen. Kesper (2018, 259) folgend gehen Dürscheid/Sutter davon aus, dass *Google* von einer Mehrheit der Schülerinnen und Schüler zu privaten wie schulischen Zwecken und daher auch bei sprachbezogenen Fragen genutzt wird. Sie zeigen am Beispiel „Portemonnaie“, dass die Website-Ausschnitte in den Suchergebnissen (= Snippets) eine wichtige Rolle spielen, da sie oftmals die Frage schon beantworten (vgl. auch Kap. 3.2 des vorliegenden Textes). Das hat zur Folge, dass eine Vielzahl an Sprachressourcen, die „ihre Informationen auf anschauliche und gut verständliche Weise darstellen“ (Dürscheid/Sutter 2020, 14), unberücksichtigt bleiben. Ein Ziel der Autorinnen ist deshalb zu zeigen, wie man „Schülerinnen und Schüler auf verlässliche und für die eigene Arbeit nützliche Quellen aufmerksam macht und ihnen zeigt, welchen Mehrwert deren Nutzung hat.“ (ebd., 14) Sie stellen exemplarisch einige nützliche Sprachressourcen vor, die Schülerinnen und Schüler kennen und nutzen sollten und zeigen am Beispiel des Unterrichtsthemas „Sprachvarietäten und ihre gesellschaftliche Bedeutung, wie mit solchen Sprachressourcen Forschendes Lernen gelingen kann. Ein solches angeleitetes und explizites Erlernen einer Informations- und Recherchekompetenz ist nötig, weil die aktuelle Situation (dezentrale Verfügbarkeit der digitalen Sprachressourcen; Schnelligkeit, mit der *Google* Ergebnisse liefert) nicht automatisch zu korrekten, vollständigen und überprüften Antworten auf eine Fragestellung führt. Es geht also darum, dass Schülerinnen und Schüler die Informationen, die sie finden, auch einordnen und bewerten können und dass sie in der schulischen Ausbildung Kriterien an die Hand bekommen, um Informationen in Bezug auf die Qualität und Zuverlässigkeit beurteilen zu können (vgl. Dürscheid/Sutter 2020, 21; Beißwenger/Storrer 2011, 120).

In Bezug auf die Recherche- und Informationskompetenz ist auch zu fragen, inwiefern die Qualität eines Nachschlagewerkes für Erstnutzende überhaupt erkennbar wird. So zeigen etwa Müller-Spitzer/Koplenig (2014) in einer Online-Studie von 2010 und aktuell Adarve (2020), dass Vertrauen in die Richtigkeit und Verlässlichkeit der Informationen ein zentraler Anspruch an (digitale) Sprachressourcen ist. Welche werkbezogenen Informationen, welche Hinweise und welche Merkmale lassen den digitalen Auftritt einer Sprachressource vertrauenswürdig erscheinen und wodurch wird möglicherweise das Gegenteil bewirkt? Welche Signale lassen einerseits die Informationen in klassischen Nachschlagewerken, andererseits in digitalen Sprachressourcen verlässlich erscheinen? Da digitale Sprachressourcen im Internet neben allerlei anderer Informationsangebote auftauchen, darf die Frage nach Signalen der Verlässlichkeit, Vertrauenswürdigkeit und Wertigkeit möglicherweise nicht getrennt von diesen anderen Informationswebsites betrachtet werden (vgl. Kap. 6).

## 4.1.2 Die Finanzierung lexikographischer Online-Angebote mit Werbung

Ein zentrales Thema in diesem Zusammenhang ist die Bannerwerbung auf den Websites digitaler Sprachressourcen, die einerseits einen kostenlosen Zugang zu Informationen ermöglicht, sich andererseits aber auch auf die Wahrnehmung, Bewertung und Benutzung der Angebote auswirkt. Beispielsweise enthält *duden.de* zahlreiche Werbebanner, welche – so müsste man annehmen – die

Wahrnehmung eines Werkes im klassischen Verständnis behindern. Der Screenshot macht deutlich, dass der Informationsgehalt im ersten sichtbaren Ausschnitt beim Aufrufen eines Artikels sehr gering ist. Das zeigen auch Untersuchungsergebnisse von Dziemianko (2018), wonach sich die Benutzungsdauer von Sprachressourcen aufgrund von Werbebannern verlängert.



Mit einem Nachschlagewerk, insbesondere mit einem gedruckten, ist ein Qualitätsversprechen verbunden, das nicht zuletzt auf kommerzieller Unabhängigkeit gründet. Gerade vor dem Hintergrund von digitalen Angeboten, die sich augenscheinlich durch mehr Werbung als Information auszeichnen, stellt sich die Frage, wie und ob das Werk Vertrauen in die präsentierten Inhalte erzeugen kann oder ob – im Falle von *duden.de* – die Website vom Ruf und Renommée des traditionellen gedruckten Werkes zehrt. Dabei ist auch zu bedenken, dass es zu zufälligen oder bewussten Interferenzen zwischen Werbeanzeige und lexikographischer Information kommen kann. Das wäre beispielsweise der Fall, wenn im Artikel „Fahrrad“ Werbung für einen Radhersteller erscheint. Andererseits ist es in Bezug auf die Wahrnehmung denkbar, dass Werbebanner zum Alltag der Internetnutzenden gehören und ihr Einfluss daher vernachlässigt werden kann.<sup>22</sup> Obwohl die Akzeptanz von Werbung in digitalen Sprachressourcen ein kontroverses Thema zu sein scheint, gibt doch eine Mehrheit (64%) der Befragten in der Studie von Adarve (2020) an, dass sie Werbung akzeptieren würden. Auf jeden Fall sollten diese Aspekte im Rahmen künftiger Publikationslösungen für digitale Sprachressourcen diskutiert und evaluiert werden, damit man mehr über die Einflüsse von Werbung auf die Wahrnehmung, Bewertung, Rezeption und den Gebrauch der Nachschlagewerke erfährt und dies bei der Gestaltung und Publikation berücksichtigen kann.

## 4.1.3 Signale der Vertrauenswürdigkeit

Bei digitalen Sprachressourcen ist es möglich – und über Suchmaschinen oder Verlinkungen in Social-Media-Angeboten sogar üblich –, direkt und ohne den Umweg über eine Startseite zu einem Artikel zu gelangen. Der Werkcharakter ist dabei nicht erkennbar. Das gilt auch oftmals für die Ergebnisauflistung

<sup>22</sup> Das zumindest legt eine Studie von Burke et al. (2005) nahe. Die Studie stammt aus einer Zeit, als der Bannergebrauch noch moderater war, und stellt trotzdem fest, dass Banner häufig der Aufmerksamkeit von Betrachtenden entgehen.

bei *Google*: Als Beispiel hierfür kann der *DWDS*-Eintrag in der schon weiter oben zitierten Trefferliste zu „ekelig vs. eklig“ gelten:

Daran wird erkennbar, dass es sich bei der verlinkten Seite um einen digitalen Speicherort handelt, an dem für das Wortfeld „Ekel · ekeln · ekel · ekelhaft · eklig“

[DWDS – ekelig – Worterklärung, Grammatik, Etymologie u. v. m.](https://www.dwds.de)  
<https://www.dwds.de> > ekelig ▾  
Ekel · ekeln · ekel · ekelhaft · eklig. Ekel m. 'Abscheu, heftiger Widerwille', im 16. Jh. vom Md. ausgehend (öfters bei Luther); wohl Substantivbildung zu ekeln (s.

Informationen zur „Worterklärung, Grammatik, Etymologie, u.v.m.“ gespeichert sind.

Man nimmt die Quelle zunächst als undefinierten Informationsspeicher wahr, auch deshalb, weil explizite Hinweise auf ein Wörterbuch fehlen – außer man weiß, wofür das Akronym *DWDS* steht. Dahingegen gibt bei *duden.de* neben der Bekanntheit des Verlags die numerisch-alphabetische Aufzählung, die ein klassisches Merkmal von Artikelmikrostrukturen ist, einen schwachen Hinweis auf einen typischen gedruckten Wörterbucheintrag.

#### Suchen | ekelig [selten] [eklig] - Duden

<https://www.duden.de> > suchen > dudenonline > ekelig [selten] [eklig] ▾  
eklig, selten ekelig. Adjektiv - 1a. ekelregend, widerwärtig, abscheulich; 1b. in seinem Auftreten unangenehm; gemein, ...2. sehr, ganz gehörig, tüchtig.

Aus einer Auflistung solcher Suchergebnisse ist die

vertrauenserweckende und verlässliche Information bzw. Informationsquelle auszuwählen. Welche Kriterien hierbei ausschlaggebend sind und – so der aus lexikographischer Sicht erwünschte Fall – zu einem Besuch der digitalen Sprachressource führen, lässt sich derzeit nur mutmaßen. Auf Basis der Analyse der Log-Files der *Varietengrammatik* ist zu vermuten, dass einerseits eine obere, aber nicht prinzipiell die oberste Position in der Ergebnisliste relevant ist. Das betrifft die prinzipielle Auffindbarkeit von digitalen Sprachressourcen und ihren Inhalten (vgl. Kap. 5.1). Darüber hinaus spielt die Sichtbarkeit (vgl. Kap. 5.2)<sup>23</sup> der Quelle eine große Rolle, denn man muss erkennen können, was sich hinter dem Suchergebnis verbirgt, welche möglicherweise weitergehenden Informationen zu erwarten sind und sicher auch, wie verlässlich und vertrauenswürdig die Informationsquelle ist.

#### 4.1.4 Auffindbarkeit und Sichtbarkeit

Die Auffindbarkeit ist ein Maß dafür, wie gut ein Informationsangebot über Suchmaschinen gefunden wird.<sup>24</sup> Durch die Sichtbarkeit im Netz wird eine digitale Markenbildung erzielt bzw. die Bekanntheit einer Sprachressource erhöht. So wird ein Angebot allein dadurch bekannter, dass es in den Suchergebnislisten regelmäßig zu einem Thema sichtbar ist. Das gilt selbst dann, wenn der Link nicht jedes Mal oder möglicherweise gar nicht angeklickt wird. *gutefrage.net* ist eine solche Seite, die viele nur deshalb kennen, weil sie sehr präsent in Suchergebnissen ist. Das folgende Beispiel illustriert den Unterschied zwischen Auffindbarkeit und Sichtbarkeit. Mit der Suchanfrage „Regionale Varianten des Deutschen“ findet man beispielsweise einen FAZ-Artikel über die *Varietengrammatik*.

[www.faz.net](http://www.faz.net) > ... > Karriere & Hochschule ▾  
**Sprachvarianten des Deutschen: Wägen fahren nur südlich ...**  
17.01.2019 - Wer den grammatischen Variantenreichtum der deutschen Schriftsprache ...  
Karten zeigen, in welchen Gebieten die verschiedenen Varianten ...

<sup>23</sup> Gelegentlich werden Auffindbarkeit und Sichtbarkeit im Internet gleichbedeutend verwendet, wohl unter der Annahme, dass das, was sichtbar ist, auch auffindbar ist. Dem schließe ich mich nicht an, weil ich gerade in der Unterscheidung (vgl. den weiteren Text) einen großen Nutzen für die Publikation von digitalen Sprachressourcen sehe.

<sup>24</sup> Neben dem Finden über die algorithmische Suche gibt es dabei weitere Möglichkeiten, wie beispielsweise finanzierte Suchmaschinenwerbung (bspw. *Google Ads*), die hier außer Acht gelassen werden, vgl. hierzu Alpar et al. (2015, 11).

Informationen über die *Variantengrammatik* und sogar die im Artikel erwähnte und verlinkte Website sind damit zumindest auffindbar, aber nichts davon ist im Suchergebnis sichtbar. Diese Differenzierung zwischen auffindbar und sichtbar ist gerade in Bezug auf Sprachressourcen sehr wichtig, weil – so meine These – etwas anderes gesucht und gefunden wird, als letztlich sichtbar werden soll. Das gilt es nun weiter auszuführen: Im DL-Projekt wurden die Rahmenbedingungen der meisten aufgesuchten Artikel der *Variantengrammatik* genauer betrachtet. Dabei hat sich durch diverse Testsuchanfragen bei *Google* gezeigt, dass diese Artikel einerseits gut in der Ergebnisliste positioniert sind. Andererseits weist aber das in der Ergebnisliste Sichtbare einen Informationsgehalt auf, der in hohem Maße mit dem Suchinteresse der Suchenden zu korrelieren scheint. Das zeigt sich am folgenden Beispiel, wo das Zitat aus der *Variantengrammatik* eine hochkomprimierte und treffende Erläuterung zum fraglichen Phänomen ist.

[ekelig / eklig – Variantengrammatik des Standarddeutschen](#)  
[mediawiki.ids-mannheim.de](#) › [VarGra](#) › [index.php](#) › [Ekelig\\_](#) › [\\_eklig](#) ▼  
Die Adjektivvarianten **ekelig** und **eklig** werden mit oder ohne -e- in den Bedeutungen 'ekelerregend' und 'unangenehm, merkwürdig' nebeneinander verwendet.

Dieses Snippet macht einen sehr gut passenden inhaltlichen Ausschnitt aus der *Variantengrammatik* sichtbar, aber beispielsweise nicht das Werk als solches. Ergänzend hierzu muss erwähnt werden, dass Snippets in den Suchergebnissen algorithmisch bestimmt werden und daher nicht von den Website-Betreibenden gesteuert werden können. Wie die Sichtbarkeit gestaltet werden kann und welche Einflussmöglichkeiten es dazu gibt, wird in Kapitel 5.2 genauer erläutert. Im Folgenden wird vorrangig die Frage thematisiert: Was soll auffindbar sein und was soll sichtbar werden? Hierfür sollen nochmals die theoretischen Überlegungen zur Werk-Identität eines Nachschlagewerks aufgenommen werden.

## 4.2 Zwei Identitäten einer digitalen Sprachressource

### 4.2.1 Im Internet werden Informationen direkt gesucht und gefunden

Im zuvor geschilderten Beispiel zu „eklig / ekelig“ werden relevante Informationen aus der *Variantengrammatik* aufgefunden und sichtbar – wahrscheinlich werden sie auch zur Beantwortung der Frage der Suchenden genutzt. Wer die obige Ergebnisdarstellung betrachtet und die Information verwertet, dem begegnet die Quelle eher in Form eines Informationsspeichers denn eines Werkes. Kriterien, die ein Werk letztlich ausmachen, spielen beim Erstkontakt keine Rolle und werden auch nur bedingt sichtbar. Das führt mich dazu, anstelle von Werk-Identität in Bezug auf digitale Angebote von einer Informationsspeicher-Identität zu sprechen. Die Informationsspeicher-Identität ist der Werk-Identität gegenüber unspezifisch – die Quelle kann beispielsweise ein Blog, eine Zeitung oder eine Werbe- oder Informationsseite sein, aber auch eine pädagogische oder wissenschaftliche Wissenssammlung oder eben ein Nachschlagewerk. Beim Erstkontakt mit der Quelle wird diese Identität nicht spezifisch geklärt. Anders ist es, wie oben dargelegt, bei einem Printwörterbuch, dessen Einband man zwangsläufig wahrnehmen muss, wenn man es aus dem Regal nimmt. Suchende bekommen die Information aus einer unspezifizierten Quelle, einem Informationsspeicher.

Das Nachschlagen wandelt sich damit grundlegend. Folgendes Gedankenspiel macht den Wandel deutlich: Stellen Sie sich vor, Sie stehen vor einem Regal mit Nachschlagewerken, von welchen Sie lediglich die Informationen auf dem Buchrücken sehen. Außerdem erhalten Sie aus allen Büchern, in welchen etwas zu Ihrer Fragestellung auffindbar ist, einen kleinen relevanten Auszug. Diese Information hängt entsprechend davon ab, wie Sie ihre Frage formuliert haben. Sie erhalten die Information, bevor

Sie sich für ein Buch aus dem Regal entscheiden müssen. Die Entscheidung für ein Werk treffen Sie damit informationsabhängig und weniger vom Thema und Titel abhängig. Im Vergleich zur oben abgebildeten Visualisierung der Recherche in Werken (Kap. 4.1) stellt sich die Recherche via Suchmaschine in Informationsspeichern folgendermaßen dar.

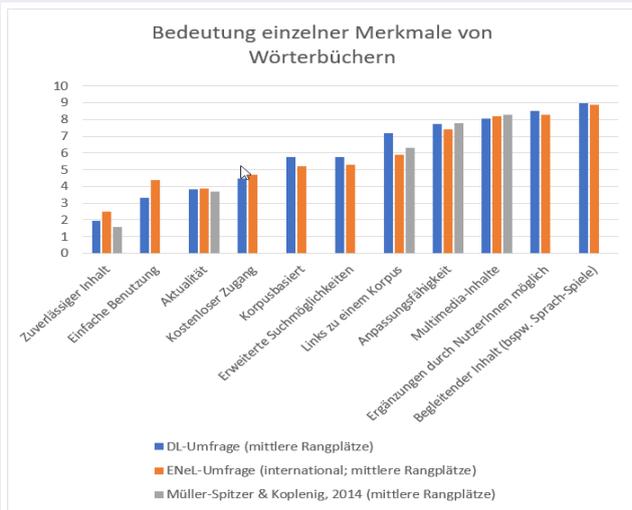


Im Vergleich der Grafiken wird deutlich: Die Werk-Identität einer Quelle ist nicht mehr für das Auffinden, sondern für das Auswählen einer Information relevant. In Bezug auf die Auffindbarkeit und Sichtbarkeit von digitalen Sprachressourcen bedeutet das: Die passende Information muss auffindbar und möglichst im Snippet sichtbar werden. Der digitale Buchrücken – letztlich die Überschrift und die URL des Werkes in der Ergebnisliste – muss möglichst relevante Informationen über das Werk sichtbar werden lassen. Das folgende Beispiel zeigt, wie dies derzeit für die *Variantengrammatik* gelöst ist.

[ekelig / eklig – Variantengrammatik des Standarddeutschen](#)  
[mediawiki.ids-mannheim.de](#) > [VarGra](#) > [index.php](#) > [Ekelig\\_](#) > [\\_eklig](#) ▾

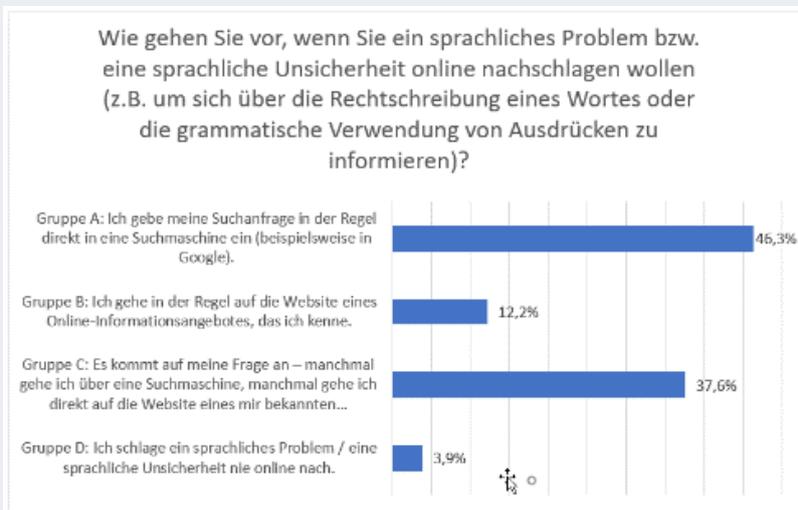
### **INFO: Umfrage zu digitalen Sprachressourcen**

Teil des Projekts „Digitale Sprachressourcen“ war eine Online-Umfrage mit 255 Teilnehmenden zur Einstellung gegenüber digitalen Sprachressourcen, zu Erwartungen, Bewertungen und Nutzungsweisen (vgl. Adarve 2020). Die Umfrage ergänzte eine frühere europaweite Umfrage des *European Network of e-Lexicography (ENeL)* (vgl. Kosem et al. 2018) sowie eine auf Deutschland bezogene Studie (vgl. Müller-Spitzer et al. 2018). Gefragt wurde u.a., ob es eine Bereitschaft gibt, für ein Nachschlagewerk etwas zu bezahlen. Während 74% der Teilnehmenden auch heute noch dazu bereit wären, für ein gedrucktes Wörterbuch Geld auszugeben, lag diese Bereitschaft bei Online-Wörterbüchern lediglich bei 38% (im Mittel 6,30 CHF). Ein weiteres Ergebnis war das folgende: Die Mehrheit der Personen (56%) konsultieren Wörterbücher in bevorzugter Weise auf dem PC, immerhin 18% bevorzugen das Smartphone zum Nachschlagen und inzwischen nur noch 16% das Printformat. Die Bereitschaft, für eine App etwas (im Schnitt 4,50 CHF) zu zahlen, liegt allerdings mit 50% höher als bei den Websites. Interessant ist auch, dass fast 64% der Befragten bereit wären, Werbung in Online-Sprachressourcen zu akzeptieren. Das bedeutet aber im Gegenzug, dass noch über ein Drittel Werbung kritisch gegenübersteht und Werbung bei diesen Personen zu einer negativen Bewertung der Qualität oder Brauchbarkeit eines Nachschlagewerkes führen kann. Leider lässt sich nichts darüber aussagen, inwiefern diese, der Werbung kritisch gegenüberstehende Gruppe sich mit der traditionell angesprochenen Zielgruppe, dem Fachpublikum und den interessierten Laien, deckt.

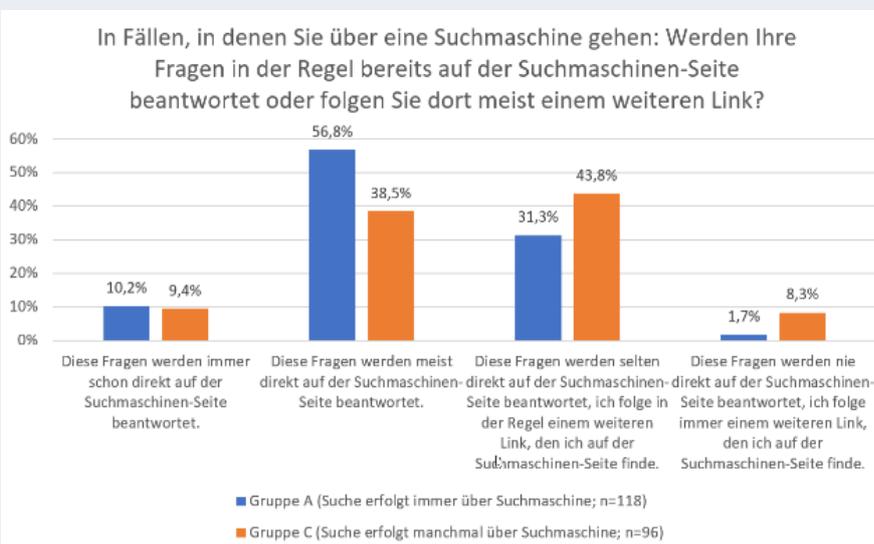


In Bezug auf die Bedeutung, die einzelnen Merkmalen von Wörterbüchern zugeschrieben werden, zeigt sich eine Konstanz im Vergleich zu der Umfrage von Müller-Spitzer/Koplenig (2014): Zuverlässigkeit und Aktualität sind weiterhin wichtige Werte von Sprachressourcen. Mit unserer und der ENEt-Umfrage wurden zudem Merkmale abgefragt, die mit hohen Zustimmungswerten die Tendenz zur Digitalisierung bestätigen: Das sind vor allem die einfache Benutzung, ein kostenloser Zugang und auch Korpusbasiertheit sowie erweiterte Suchmöglichkeiten.

Zusätzlich zu dem Teil der Umfrage, der sich an der Vorgängererhebung der ENEt orientiert hat, wurde versucht, mehr über das Online-Nutzungsverhalten herauszufinden. Dabei hatten wir die Auffindbarkeit und Sichtbarkeit der Sprachressource im Fokus und vermuteten, dass einige Personen ihre Informationen aus einer Suchmaschine und nicht von den Websites beziehen.



Diese Vermutung bestätigte sich in der Umfrage eindrucksvoll: 46% geben ihre Frage direkt in eine Suchmaschine ein, und weitere 37% tun dies, wenn es der Fragestellung angemessen ist. Und auch die anfänglich im Projekt noch als These formulierte Vermutung, dass häufig die Information im Snippet einer Suchmaschine schon ausreicht und in Folge dessen die dahinterstehende Sprachressource gar



nicht mehr konsultiert wird, bestätigte sich in der Umfrage. Man kann aus nebenstehender Grafik zu den Umfrageergebnissen vorsichtig ableiten, dass fast die Hälfte der Suchenden eine Antwort aus unterschiedlichen Quellen erhalten, ohne diese Quelle zu besuchen. Rund 10% der Befragten gibt an, dass die Informationen auf der Suchmaschinen-Seite für die Beantwortung ihrer Frage immer ausreichend seien.

Informationen auf der Suchmaschinen-Seite für die Beantwortung ihrer Frage immer ausreichend seien.

## 4.2.2 Ist das Werk noch wichtig?

Man könnte nach alledem zum Schluss kommen, dass es vielen Personen ausreicht, wenn ihnen eine zufriedenstellende Information zu ihrer Suchanfrage auf einer Suchmaschinen-Seite präsentiert wird. Aus welchem Werk bzw. aus welcher Quelle diese Information stammt, wäre dann nebensächlich oder gar irrelevant. Doch die bereits erwähnte Umfrage zu den Erwartungen an Nachschlagewerke zeigt auch, dass es durchaus über die reine Information hinausreichende Ansprüche an Sprachressourcen gibt. Das ist neben der Vertrauenswürdigkeit und Verlässlichkeit auch die Zitierfähigkeit. Während all diese Ansprüche mit dem Griff in ein Bücherregal, spätestens mit der Wahrnehmung des Einbandes, als geprüft gelten konnten, müssen sie beim digitalen Nachschlagen via Suchmaschine noch geprüft oder verifiziert werden. Und wie man beim obigen Gedankenspiel bei gleichwertigen Informationen wohl zu dem Buch greifen würde, dessen Einband am ehesten verspricht, die Erwartungen zu erfüllen, klicken Suchende im Internet wohl auf den Eintrag in der Suchliste, der ihren Interessen und Ansprüchen am ehesten zu entsprechen scheint.

Zu der Frage, wie Suchende die Darstellung und die Informationen in der Ergebnisliste beurteilen, darüber gibt es noch keine Erkenntnisse: Wie viele und welche Merkmale sind beispielsweise nötig, damit digitale Sprachressourcen als Werke und nicht nur als anonyme Informationsspeicher wahrgenommen werden? Und wie gut lässt sich die Werk-Identität – und die damit verbundene Vertrauenswürdigkeit und Zuverlässigkeit – schon in der Ergebnislisten-Darstellung einer Suchmaschine zum Ausdruck bringen? Diese Fragen rücken im Kontext von Sichtbarkeit und Usability einer digitalen Sprachressource in den Fokus. Denn mit der Verfügbarkeit als Informationsspeicher im Internet verändert sich das potentielle Publikum. Das bedeutet nicht, dass der klassische Nutzer oder die klassische Nutzerin wegfallen, sie sind gewissermaßen die Stammbesuchenden, die direkt die Website anwählen und auf das Werk zugreifen. Aber ihr Anteil an der Gesamtnutzung der digitalen Sprachressource wird kleiner aufgrund derer, die nur dann das Werk besuchen, wenn es in der Suchmaschine erstens erscheint und sich zweitens besuchswürdig präsentiert. Unter Marketing-Gesichtspunkten steckt in diesen Einmal-Besuchenden das Potential, dass diese zu Stammbesuchenden werden. Dazu müssen sie aber bei ihrem einmaligen – oder zufälligerweise zweiten, dritten oder vielfachen – Besuch entdecken, dass es sich bei der Quelle um ein (nützliches) Werk handelt. Und sie müssen erkennen, welche Informationen dieses Werk liefert. Sie müssen die digitale Sprachressource also nicht mehr nur als Informationsspeicher, sondern als Werk wahrnehmen. **Über die Informationen wird die Sprachressource aufgefunden, über die Werk-Identität wird sie bekannt und etabliert.**

## 4.2.3 Der Nutzen dieser Unterscheidung

Die Unterscheidung in Informationsspeicher-Identität und Werk-Identität ist deshalb sinnvoll, weil es das Anliegen jedes lexikographischen Teams ist, ein Werk zu schaffen. Das Ideal der blätternden und stöbernden Fachnutzende und der „interessierten Laien“ baut auf der Vorstellung des Nachschlagewerks als einem „lesbaren“ Gesamtwerk auf, ein Gesamtwerk, das sich durch ein durchdachtes Kompositionsprinzip und zahlreiche Querverweise auszeichnet. Doch wahrscheinlich muss man sich um diese Gruppe von kundigen Nutzerinnen und Nutzern weniger intensiv kümmern, weil sie sich des Wertes eines durchdachten und wissenschaftlich erarbeiteten Nachschlagewerkes bewusst ist. Deshalb rücke ich die Gruppe der über Suchmaschinen und gegebenenfalls andere digitale Kanäle angesprochenen und interessierten Nutzerinnen und Nutzer stärker in den Fokus. Für diese

Gruppe existiert in der Lexikographie bislang wenig Bewusstsein, obwohl sie ein wachsendes Potential für die Nutzung und Bekanntheit von digitalen Sprachressourcen bietet. Das gilt, wie eingangs erwähnt, in besonderem Maße für Sprachressourcen mit Fokus auf speziellere Fachthemen. Potentielle Stammesbesuchende, die beim Erstkontakt mit der Sprachressource hinter einem Link oder einer Website lediglich einen Informationsspeicher unbekannter Qualität vermuten, brauchen während ihrer Erstnutzung prägende Hinweise auf den Werkcharakter und auf die Verlässlichkeit und Wertigkeit der Quelle. Hier steckt das größte Potential für alle Anbietenden von digitalen Sprachressourcen, die sich mehr interessierte Besucherinnen und Besucher wünschen.

#### 4.2.4 Lexikographie-kompatibles Wissenschaftsmarketing

Unter Wissenschaftlerinnen, Wissenschaftlern und Institutionen gibt es begründete und nachvollziehbare Argumente gegen Suchmaschinen(-Optimierungen) und werbefinanzierte Angebote. Meist sieht man die wissenschaftliche Unabhängigkeit in Gefahr und befürchtet, dass ein lexikographisches Werk in unerwünschter Weise verändert werden müsse. Ich möchte im Folgenden zeigen, dass Suchmaschinen und Optimierungsideen von digitalen Sprachressourcen nicht einem plakativen Marketing-Denken folgen müssen. Vielmehr geht es darum, Wege zu finden, um die wertigen und wichtigen Inhalte, die in oftmals vielen Jahren lexikographischer Arbeit erstellt wurden, auffindbar zu machen. Und es geht darum, das Werk so gut wie möglich und akzeptiert im Internet wahrnehmbar und sichtbar zu machen. Es geht darum, dass Suchende geprüfte, empirisch fundierte und verlässliche Antworten auf ihre Fragen bekommen und dabei erkennen, welches Werk dahintersteckt.

Das dient letztlich der Vermarktung des Werkes: Das Werk – letztlich eine Marke – soll für alle Personen, die es finden (deshalb: Auffindbarkeit erhöhen), sichtbar werden (deshalb: Sichtbarkeit des Werks erhöhen) und einfach zu nutzen sein (deshalb: Usability verbessern). Im Folgenden wird ein Ansatz zum Wissenschaftsmarketing skizziert, der mit den Ansprüchen und Bedingungen der wissenschaftlichen Lexikographie gut verträglich ist und das vor allem nicht den Erkenntnissen über die Wörterbuchbenutzung und -optimierung zuwiderläuft.

##### ***INFO: Auffindbarkeit und Sichtbarkeit***

Die Verwendung der Ausdrücke „Auffindbarkeit“ und „Sichtbarkeit“ im Zuge der SEO ist nicht einheitlich. Strenggenommen meint Auffindbarkeit, dass eine Website technisch so programmiert ist, dass Suchmaschinen die Seite indizieren können. Wenn die Seite auf dem Index steht, ist sie prinzipiell auffindbar. Das bedeutet, bei einer Suche mit einem indizierten Keyword wird die Seite in den Suchergebnissen angezeigt. Damit die Seite sichtbar wird, muss sie besonders weit oben in den Suchergebnissen erscheinen, also einen guten PageRank haben. Für die Verbesserung des PageRank ist die Suchmaschinenoptimierung zuständig.

In den folgenden Kapiteln werden alle Bestrebungen für eine gute Positionierung der Seite unter dem Stichwort „Auffindbarkeit“ zusammengefasst, denn faktisch findet niemand einen Eintrag, der auf Seite 10 der *Google*-Suchergebnisse aufgeführt ist. Mit „Sichtbarkeit“ wird von der SEO-Terminologie abweichend der Informationsgehalt der Suchergebnisse in Bezug auf die verlinkte Seite beschrieben. Denn nur wenn in den Ergebnislisten der Suchmaschinen etwas Aussagekräftiges steht, wird sichtbar, welche Seite sich hinter dem Link befindet.

## 5 Auffindbarkeit, Sichtbarkeit und Usability digitaler Sprachressourcen

Wie angedeutet, hat sich durch die Digitalisierung ein zweiter, anderen Bedingungen und Verhaltensweisen folgender Zugang zu Sprachressourcen eröffnet: Das Auffinden eines passenden Eintrags im Nachschlagewerk wurde an eine externe Suchmaschine ausgelagert, die direkt in den Artikel verlinkt, ohne auf die werkimmanenten Strukturen und Zugriffsoptionen einer digitalen Sprachressource zurückzugreifen. Dieser neue Zugang zu Nachschlagewerken sollte nicht als Alternative oder Ersatz für die klassische Nutzung gesehen werden, sondern als Ergänzung. Genau in diesem Sinne sind die folgenden Fragen, Ideen und Vorschläge formuliert. Sie orientieren sich an den drei Schritten: Auffinden einer Quelle für die gesuchte Information, diese Quelle sehen, erkennen und auswählen und schließlich dort die gewünschten Antworten auf eine Frage finden.

### 5.1 Auffindbarkeit – Digitale Werke und Informationen in Suchmaschinen

Eine digitale Sprachressource steht in keinem Bibliothekskatalog und wird normalerweise von keinem Verlag beworben – dennoch soll sie gefunden und genutzt werden. Die Auffindbarkeit von Websites wird direkt mit Suchmaschinenoptimierung (SEO) in Verbindung gebracht. Doch oft werden mit SEO sinnlose Texte verknüpft, die mit sich wiederholenden Schlagwörtern (Keywords) überladen sind. Darauf begründen sich viele Abneigungen gegenüber SEO-Maßnahmen. Diese Vorstellung allerdings ist nicht mehr aktuell<sup>25</sup>, denn seit 2013 verwendet die Suchmaschine *Google* ein Update des Such-Algorithmus, mit dem die Häufigkeit der Keywords weniger stark gewichtet wird und stattdessen semantische Informationen ausgewertet werden.<sup>26</sup> Dieses Update mit dem Namen *Hummingbird*<sup>27</sup> arbeitet nach dem Prinzip der semantischen Suche, bei der der Algorithmus versucht zu berücksichtigen, ob die Wörter in einem Satz einen sinnhaften Zusammenhang haben. Thematisch zusammenhängende Texte gelten daher bei der Suchmaschinenoptimierung als gut passend. Aus diesem Grund sind Wörterbuchartikel prinzipiell nicht ungeeignet für die semantische Suche mit dem *Hummingbird-Update*. Mit einem weiteren Update des Algorithmus, dem *Panda-Update*, favorisiert die Suchmaschine Inhalte von hoher Qualität und Individualität.<sup>28</sup> Auch das kommt wissenschaftlichen und spezialisierten Sprachressourcen prinzipiell entgegen und hebt sie von Informationsangeboten ab, die keine eigene Redaktion haben, sondern fremde Inhalte kopieren. Außerdem ist es vorteilhaft für das Ranking, wenn eine Seite bei bestimmten Suchanfragen häufig besucht wird. Auch diese Ranking-Parameter sind prinzipiell kein Nachteil für digitale Sprachressourcen. Das zeigt sich beispielsweise bei der *Variantengrammatik* daran, dass sie mit solchen Artikeln, die gut zu einer Suchanfrage passen, auf den ersten drei Positionen gerankt ist. Sie ist gut auffindbar.

---

<sup>25</sup> Man vergleiche den Text zum Information-Retrieval von A. Henrich aus dem Jahr 2008, in dem minderwertige Inhalte, Manipulationsmöglichkeiten und sinnlose Verlinkungen als zentrale Kritikpunkte am PageRank angeführt werden (Henrich 2008, 392).

<sup>26</sup> Das sogenannte Pinguin-Update der *Google*-Suchmaschine bewertet es als Verstoß gegen die Guidelines, wenn Texte übermäßig mit bestimmten Schlagworten angereichert werden (sog. Keyword-Stuffing).

<sup>27</sup> Vgl. [https://en.wikipedia.org/wiki/Google\\_Hummingbird](https://en.wikipedia.org/wiki/Google_Hummingbird).

<sup>28</sup> Vgl. [https://de.wikipedia.org/wiki/Google\\_Panda](https://de.wikipedia.org/wiki/Google_Panda); <https://googleblog.blogspot.com/2011/02/finding-more-high-quality-sites-in.html>.

Obwohl die Bedingungen für die Auffindbarkeit digitaler Sprachressourcen in *Google*-Suchen eher günstig als nachteilig sind, ist eine professionelle und fortlaufende Suchmaschinenoptimierung dennoch teuer. Es ist auch fraglich, inwiefern sich diese professionelle Optimierung für ein Nachschlagewerk lohnt, weil man jeden Artikel spezifisch optimieren müsste. Doch es bietet sich – insbesondere bei der Neukonzeption oder Überarbeitung von Sprachressourcen – an, einige der im weiteren Verlauf aufgeführten Aspekte zu berücksichtigen. Mit der Frage, ob und wie eine Website von *Google* erfasst wird und wie sie in den Suchergebnissen gerankt wird, befasst sich der für den Einstieg empfehlenswerte von *Google* publizierte *Startleitfaden für die Suchmaschinenoptimierung (SEO)*.<sup>29</sup> Die erste dort formulierte Frage ist: Was sollen potentielle Nutzerinnen und Nutzer finden? Hier wird die weiter oben thematisierte Unterscheidung relevant: Soll ein Artikel aus dem Informationsspeicher gefunden werden oder soll das Werk anhand des Titels und thematischer Schwerpunkte gefunden werden? Weiter oben wurde argumentiert, dass der primäre Zugang zu einer digitalen Sprachressource über die enthaltenen Informationen lohnend erscheint – insbesondere in Bezug auf einen Nutzendenkreis, der das Werk bislang nicht kennt. Diese Fokussierung auf die Artikelinhalte wird zumindest für die *Variantengrammatik* gestützt, weil sich herausstellte, dass für das Werk relevante Suchbegriffe keine hohen Suchvolumina haben.

### 5.1.1 Das Werk finden

Die *Variantengrammatik* ist daher ein sehr gutes Beispiel, um aufzuzeigen, welche Überlegungen nötig sind, wenn man erreichen möchte, dass das Werk – also die rahmenden und übergreifenden Seiten einer digitalen Sprachressource – gefunden wird. Im Kern geht es darum, herauszufinden, mit welchen Formulierungen nach sprachlichen Phänomenen gesucht wird, zu welchen die *Variantengrammatik* generell etwas beitragen kann. Um sich über die Suchvolumina einzelner Suchbegriffe informieren zu können, stellt *Google* einen *Keyword-Planer*<sup>30</sup> zur Verfügung. Mit diesem Tool lässt sich die durchschnittliche monatliche Anzahl an Suchanfragen für bestimmte Formulierungen und Ausdrücke ermitteln. Der *Keyword-Planer* bietet auch die Möglichkeit, eine bestehende Website auf hilfreiche Keywords untersuchen zu lassen. Für die *Variantengrammatik* lieferte diese Analyse der bestehenden Website allerdings keine Vorschläge. Das heißt, auf den aktuellen Seiten der Sprachressource wurde kein Keyword mit einem relevanten Suchvolumen identifiziert.

Keyword (nach Relevanz) ↓	Durchschnittl. Suchanfragen pro Monat	Wettbewerb	Anteil an mi Anzeigenimpre	Gebot für obere Positionen (unterer Bereich)	Gebot für obere Positionen (oberer Bereich)	Kontostatus
Von Ihnen eingegebene Begriffe						
<input type="checkbox"/> österreichischer dialekt	320	Gering	–	–	–	
Keyword-Ideen						
<input type="checkbox"/> österreichische wörter dialekt	90	Gering	–	–	–	
<input type="checkbox"/> österreichische dialekte	140	Gering	–	–	–	
<input type="checkbox"/> österreichischer dialekt und mundart	20	Gering	–	–	–	

<sup>29</sup> Vgl. <https://support.google.com/webmasters/answer/7451184>.

<sup>30</sup> Vgl. <https://ads.google.com/home/tools/keyword-planner/>.

### 5.1.1.1 Keywords – der Schlüssel zur Auffindbarkeit

Also wurden Überlegungen angestellt, welche Ausdrücke das Themenfeld der *Variantengrammatik* am treffendsten beschreiben. Zu diesen Ausdrücken haben wir überprüft, wie hoch die jeweilige monatliche Anzahl an Suchanfragen ist. Ein Ausdruck wie „Grammatik“ hat zwar mit rund 27.000 Suchanfragen pro Monat ein vergleichsweise hohes Volumen, allerdings ist davon auszugehen, dass die meisten Suchanfragen nach „Grammatik“ nicht zum Angebot der *Variantengrammatik* passen. Im Gegensatz dazu sind für eine gut passende Formulierung wie „Variation Sprache“ nur zehn Suchanfragen pro Monat verzeichnet, „Regionalsprache“ hat immerhin 90. Es zeigt sich für dieses konkrete Nachschlagewerk, dass zum Werk passende thematische Anfragen insgesamt selten sind.

Wenn man das Thema auf für das Werk linguistisch nicht mehr zutreffende Ausdrücke wie „Dialekt“ oder „Mundart“ erweitert, dann kommen beispielsweise „österreichischer Dialekt“ mit 320 Suchanfragen pro Monat oder „österreichische Wörter Dialekt“ mit 90 Suchanfragen pro Monat in Frage. Inwiefern das Angebot der *Variantengrammatik* zum Interesse von Personen passt, die diese Suchanfragen formulieren, lässt sich allerdings ad hoc nicht klären. Weitere denkbare Keywords wären beispielsweise auch:

Schweizerdeutsch (9.900 Suchanfragen/Monat); Schwyzerdütsch (990); Duden schweizerdeutsch (90); Schweizer Dialekt (720); schweizerdeutsch Dialekt (70); österreichische Wörter Dialekt (90); österreichische Wörter Mundart (140); österreichischer Dialekt (480); typisch österreichisch (320); österreichische Mundart (70) auf schweizerdeutsch (50); auf österreichisch (50)

Weiterhin muss man bedenken, dass es nicht um Klickzahlen geht, sondern darum, Besuchende mit einem relevanten Interesse an dem Informationsangebot zu gewinnen. Es ist also nicht die quantitative, sondern eine qualitative Optimierung der Keywords anzustreben. Wenn die Menge an relevanten Keywords klein ist, bedeutet das letztlich auch, dass nach dem konkreten Thema mit den gängigen Bezeichnungen niemand sucht.

### 5.1.1.2 Textgestaltung für SEO

Da die Keywords zentral für die Suchmaschinensuche sind, müssten diese auch auf der Website erscheinen. Spätestens an diesem Punkt wird es problematisch: In der *Variantengrammatik* geht es schließlich gerade nicht darum, schweizerische Dialekte mit österreichischen zu vergleichen, sondern die standardsprachlichen Varianten aus den unterschiedlichen Regionen darzustellen (vgl. Dürscheid et al. 2019). Es entfällt also die SEO-taktisch vielleicht vielversprechende Option, auf der Website über Dialekt, Mundart oder Umgangssprache zu schreiben, um damit mehr Interessierte zu gewinnen. Bestenfalls ließen sich solche Ausdrücke mit rhetorischen Kniffen einfügen. So könnte man schreiben: „In der *Variantengrammatik* werden standardsprachliche Varianten aus dem deutschsprachigen Raum beschrieben. Es geht hier ausdrücklich nicht um Dialekte oder Mundart.“

Insgesamt zeigt die Häufigkeit der relevanten Keywords aber auch, dass nach einem Werk mit den Inhalten der *Variantengrammatik* kaum jemand sucht, was angesichts des spezialisierten und eher wissenschaftlich motivierten Themas nicht verwundert. Für eine Vermarktung über Suchmaschinen ist dieses Werk eher ungeeignet – ganz anders als beispielsweise ein Deutsch-Englisch-Wörterbuch mit einem begrifflich klar definierten und entsprechend gesuchten Thema. Die Tatsache allerdings, dass Menschen nicht nach einem Thema suchen, das sie nicht gut benennen können, oder nach einem

Nachschlagewerk, von dessen Existenz sie nichts wissen, heißt nicht, dass sie nicht nach konkreten Phänomenen dieses Themenbereichs oder nach Inhalten dieser Sprachressource suchen. Interessanter als die Suchmaschinenpositionierung des Gesamtwerks ist daher die Optimierung der Auffindbarkeit der einzelnen Artikel-Seiten.

### 5.1.2 Informationen finden

Damit lexikographische Artikel von einer Suchmaschine gefunden werden, ist es wichtig, dass die sprachliche Darstellung eines Phänomens zu den Formulierungen passt, mit welchen nach diesem Phänomen gesucht wird. Damit wird eine anwendungsorientierte Ausrichtung von Nachschlagewerken gestärkt und Redaktionen müssen sich vermehrt Gedanken über die Situation machen, in welcher ein Nachschlagebedürfnis entsteht (vgl. Müller-Spitzer 2016a, 335f.). Wieder stellen zweisprachige Wörterbücher den Idealfall dar: Eine Person kennt ein Wort in einer Ausgangssprache (bspw. „Fahrrad“) und möchte wissen, wie dieses Wort in die Zielsprache bzw. Zielvarietät übersetzt wird (bspw. ins Schweizerhochdeutsche). Also fragt sie sich beispielsweise: „Wie heißt ‚Fahrrad‘ im Schweizerdeutschen?“<sup>31</sup> Das für diese Person relevante Wort „Velo“ kommt logischerweise nicht in der Frage vor. Ein „zweisprachiges“<sup>32</sup> Wörterbuch ist hier für die Auffindbarkeit der Informationen prädestiniert, weil es im Artikel sowohl „Fahrrad“ als auch „Velo“ enthält. Ein einsprachiges Wörterbuch zum Schweizer Standarddeutsch hingegen würde möglicherweise nur „Velo“ enthalten und damit wäre der Artikel mit der Suchanfrage nicht auffindbar. Für die Texte in der Sprachressource gilt also, dass möglichst die Formulierungen enthalten sind, mit welchen nach dem Phänomen gesucht wird.

Die beispielhaft formulierte Suchanfrage enthält zwei solcher typischer Fragebestandteile, die für die Suche wichtig sein können. Erstens verweist „Wie heißt ...“ darauf, dass Sprachinformationen gesucht werden und nicht beispielsweise ein Geschäft für Fahrräder. Da es bei Suchmaschinen nicht nur um das Auffinden, sondern auch um die Position in der Ergebnisliste geht, liegt in solchen Zusätzen eine Konkretisierungsmöglichkeit für Suchende, die eben keine Fahrradgeschäfte suchen. Das kann man selbstverständlich auch bei der Gestaltung von Sprachressourcen nutzen, indem man typische Formulierungen, mit welchen nach Sprachinformationen gesucht wird, in die Seite(n) integriert. Zweitens verweist der Ausdruck „Schweizerdeutsch“ in der Suchanfrage auf die Zielsprache. Und das ist ein wichtiger Aspekt, in dem sich n von Informationsspeichern unterscheiden. Bei einem zweisprachigen Wörterbuch stehen im Titel die beiden Sprachen, und damit wissen Nutzerinnen und Nutzer, dass die Inhalte des gesamten Buches sich auf diese beiden Sprachen beziehen. Der Einband des Nachschlagewerkes wie auch seine Umtexte (bspw. Abkürzungsverzeichnisse oder Zeichenerklärungen) liefern Rahmeninformationen, die sich auf den gesamten lexikographischen Inhalt beziehen. Bei digitalen Sprachressourcen finden sich solche Rahmeninformationen auf der Startseite oder auf spezifischen Unterseiten der Website, aber sie umschließen die Inhalte nicht wie ein Bucheinband. Diese Rahmeninformationen befinden sich möglicherweise auch nicht auf der Artikelseite, auf die eine Suchmaschine verlinkt. Wenn also auf der Artikelseite weder steht, dass „Velo“ eine Variante aus der Schweiz ist, noch die Variante „Fahrrad“ genannt wird, dann fehlen für den Sucherfolg wichtige Angaben. Im Kontext der Analysen von Suchanfragen, die zur *Varietätsgrammatik* passen, wurde

---

<sup>31</sup> Beim aufmerksamen Lesen werden weitere sprachliche Kniffe offensichtlich: Während für Expertinnen und Experten eine korrekte Bezeichnung des schweizerischen Standards beispielsweise „Schweizerhochdeutsch“ ist, würden Suchende etwa den Ausdruck „Schweizerdeutsch“ benutzen. Dieser wiederum verweist aus fachlicher Sicht auf die Dialekte der Schweiz.

<sup>32</sup> Das Beispiel „Fahrrad – Velo“ ist selbstverständlich kein Beispiel, das zwei Sprachen, sondern nur zwei Varietäten betrifft. Daher steht „zweisprachig“ hier in Anführungszeichen; vgl. zu der Frage Sprache oder Dialekt/Varietät: <https://www.linguistik.uzh.ch/de/easyling/faq/kolmer-schweizerdeutsch.html>.

festgestellt, dass die Website *www.gutefrage.net* sehr häufig in den Suchergebnissen erscheint. Eine Erklärung könnte sein, dass durch das Konzept dieser Website Frage und Antwort auf ein und dergleichen Seite erscheinen. Dieses Prinzip, Frage und Antwort parallel zu nennen, findet sich nach unseren Recherchen bislang auf keiner spezifisch lexikographischen Seite. Doch es erscheint aufgrund der Spezifik der Suchalgorithmen vielversprechend, wenn möglichst viele Aspekte der Suchanfrage auch auf der Website ausformuliert sind. Das soll in den folgenden Überlegungen berücksichtigt werden.

### 5.1.2.1 Optimierung der Artikelseiten

Optimierung heißt, wie zuvor beschrieben: Die sprachliche Darstellung der Informationen passt zu den Formulierungen, mit welchen nach diesen Informationen gesucht wird. Das betrifft einerseits die sichtbaren Informationen auf der Artikelseite, insbesondere den lexikographischen Artikel, aber auch Grafiken, Bilder, Logos, Verlinkungen oder Überschriften. Außerdem betrifft das die sogenannten Meta-Informationen, die im HTML-Code der Website notiert werden und vor allem für Suchmaschinen relevante Informationsquellen sind. Der folgende Auszug zeigt den Kopfbereich des für Websitebesuchende normalerweise nicht sichtbaren HTML-Codes der *Varietengrammatik*, also alles, was nicht zum Seiteninhalt gehört.

```
2 <html class="client-nojs" lang="de" dir="ltr">
3 <head>
4 <meta charset="UTF-8"/>
5 <title>Varietengrammatik des Standarddeutschen</title>
6 <script>document.documentElement.className = document.documentElement.className.replace( /(^|\s)client-nojs(\s|$)/, "$1client-js$2" );</script>
7 <script>(window.RLQ=window.RLQ|| []).push(function() {mw.config.set({"wgCanonicalNamespace":"","wgCanonicalSpecialPageName":false,"wgNamespacesNumber":0,"wgPag
8 mw.user.tokens.set({"editToken":"+\\","patrolToken":"+\\","watchToken":"+\\","csrfToken":"+\\"});/*@nomin*/;
9
10 });mw.loader.load(["mediawiki.page.startup","mmv.head"]);</script>
11 <link rel="stylesheet" href="/VarGra/load.php?debug=false&lang=de&modules=ext.bootstrap.styles&only=styles&skin=chameleon"/>
12 <script async="" src="/VarGra/load.php?debug=false&lang=de&modules=startup&only=scripts&skin=chameleon"></script>
13 <meta name="ResourceLoaderDynamicStyles" content="" />
14 <link rel="stylesheet" href="/VarGra/load.php?debug=false&lang=de&modules=site.styles&only=styles&skin=chameleon"/>
15 <meta name="generator" content="MediaWiki 1.28.2"/>
16 <meta name="description" content="Die Varietengrammatik informiert auf wissenschaftlicher Basis über Unterschiede in der Grammatik der deutschen Sprache in
17 <meta name="viewport" content="width=device-width, initial-scale=1.0"/>
18 <link rel="shortcut icon" href="/VarGra/resources/assets/favicon-vg.ico"/>
19 <link rel="search" type="application/opensearchdescription+xml" href="/VarGra/opensearch_desc.php" title="Varietengrammatik des Standarddeutschen (de)"/>
20 <link rel="EditURI" type="application/rsd+xml" href="http://mediawiki.ids-mannheim.de/VarGra/api.php?action=rsd"/>
21 <link rel="copyright" href="https://creativecommons.org/licenses/by-sa/3.0/deed.de"/>
22 <link rel="alternate" type="application/atom+xml" title="Atom-Feed für „Varietengrammatik des Standarddeutschen“" href="/VarGra/index.php?title=Spezial:Let
23 </head>
```

Anders als bei den Einträgen im HTML-Code, die auf einer Website nicht sichtbar sind, ist die Veränderung von Artikeltexten schwieriger. Wie stark kann und will man in einen fachlich präzisen Text und in traditionelle und sich als brauchbar erwiesene Strukturen eingrenzen, um die Auffindbarkeit von digitalen Sprachressourcen zu erhöhen? Das sind Fragen, die in jedem Projekt diskutiert werden müssen und sollten. Da gibt es selbstverständlich und berechtigterweise Reibungen, die aber zu Lösungen führen können, die nicht nur „freundlicher“ in Bezug auf Suchmaschinen, sondern auch freundlicher in Bezug auf ein heterogener werdendes Publikum sind. Ich sehe in einer solchen Diskussion die Chance, aus neuen Perspektiven auf alte Strukturen zu schauen und dadurch auch sinnvoll innovativ werden zu können.

### 5.1.2.2 Lexikographische Artikel auffindbar gestalten

Im Folgenden sind Beispiele und Erkenntnisse der Recherchen zu den Artikeln der *Varietengrammatik* dargestellt. Dabei steht im Fokus, wie gut oder schlecht die Artikel potentiell auffindbar sind.

#### Welche Artikel der *Varietengrammatik* werden gefunden?

Diese Frage impliziert mit dem Blick auf die Log-File-Daten zu den am häufigsten besuchten Artikeln der *Varietengrammatik* folgende Unterfrage: Wo passt die Lemmaformulierung am besten zur

Suchanfrage? Inwiefern die Suchmaschinen Lemmata identifizieren und damit einen strukturierten Zugang zu Sprachressourcen bieten können, ist unklar. Aber der Algorithmus scheint aufgrund von Formatierungen der Überschriften und aufgrund von häufigeren Nennungen des Lemmas im Artikel zu erkennen, dass das Lemma ein zentraler Ausdruck der jeweiligen Seite ist. Neben der Passgenauigkeit von Lemma und Suchanfrage war es für die gute Auffindbarkeit eines Artikels relevant, wie viele alternative Treffer es gab und ob diese von der Suchmaschine als relevant bewertet wurden. Dabei konnten wir drei Einflussfaktoren identifizieren, die zu einem guten Ranking in der Trefferliste führen, und damit zu einer guten Auffindbarkeit:

### 1) Seltene Suchwörter:

Selten meint in diesem Kontext, dass es wenige andere Seiten gibt, die sich mit diesem Wort in einer für Suchalgorithmen relevanten Intensität beschäftigen. Das heißt konkret: Es gibt keine Shops, keine Interessengruppen, keine aktuellen Themen und auch sonst nichts, bei welchen das Wort im Zentrum steht. „Ellbogen“ ist hierfür ein gutes Beispiel, weil die alternative Variante „Ellenbogen“ zu keinen lexikographischen Einträgen führt, sondern ausschließlich zu medizinischen Seiten (die wahrscheinlich ein besseres SEO haben).

Weitere Wörter, die aufgrund ihrer geringen Häufigkeit in für Suchmaschinen relevanten Kontexten im Internet zum Auffinden der *Variantengrammatik* führen, sind: „betreffend“, „entgegen“, „Schippe“ und „terminisieren“.

### 2) Mehrgliedrige Ausdrücke: Phrasen

Bei den mehrgliedrigen Ausdrücken wird das Trefferfeld stark eingegrenzt, weil alle Seiten herausfallen, die nur einen Teil des Gesamtausdrucks enthalten oder auf welchen die Bestandteile des Ausdrucks nicht kontinuierlich, sondern über die Seite verteilt vorkommen.

Ein gutes Beispiel ist „Hinzu kommt, dass“. „Hinzu kommt“ ist zum einen ein häufig gesuchtes sprachliches Phänomen, zum anderen wird es nur von grammatischen und nicht lexikalischen Nachschlagewerken behandelt. Möglicherweise interessiert sich die Mehrheit der Suchenden aber eher für die Rechtschreibung, nicht für die Grammatik: Denn die beiden orthographischen Varianten „hinzukommt“ oder „hinzu kommt“ haben ebenfalls eine relevante Suchhäufigkeit.

Ein zweites Beispiel ist „Verbzweitstellung nach weil“. Auch hier kommen zwei Aspekte zusammen, die diesen Suchausdruck so erfolgreich machen: Erstens ist dieses Phänomen Teil eines aktuellen – wenn auch abklingenden – Diskurses. Und zweitens hat sich, eben aufgrund dieses Diskurses, der feststehende, mehrgliedrige Ausdruck herausgebildet, der das besagte Phänomen beschreibt. Wahrscheinlich gibt es in der *Variantengrammatik* nicht so viele vergleichbare mehrteilige Ausdrücke, bei welchen die skizzierten positiven Rahmenbedingungen zusammenkommen. Dennoch liegt eine Chance in diesen exponierten Beispielen, denn sie führen mit hoher Wahrscheinlichkeit zu Besuchen der Sprachressource. Und diese Besuche sind von einer gewissen Passgenauigkeit und Qualität, weil sich die Besucherinnen und Besucher tatsächlich genau für dieses Phänomen interessieren und damit auch für den Gegenstand der *Variantengrammatik*. Oder etwas konkreter formuliert: Artikel, die mit einer wahrscheinlich grammatischen Fragestellung häufiger besucht werden (wie „betreffend“, „entsprechend“, „Hinzu kommt, dass“ oder „Verbzweitstellung nach weil“) bieten sich vorrangig für das Marketing des Nachschlagewerkes an (vgl. dazu Kap. 5.2).

### 3) Mehrgliedrige Ausdrücke: Variantenpaare

Die dritte, auffällig häufige Gruppe von Anfragen sind Variantenpaare wie „oft oder öfters“. Weitere Beispiele hierfür sind „Banker/Bänker“, „neidig neidisch“, „ekelig vs. eklig“, „brenzlich vs. brenzlig“ oder „Ellbogen und Ellenbogen“.

In den Beispielen sind unterschiedliche Muster dieser Variantenpaare aufgeführt. Wir konnten mit unseren Tests bei Suchmaschinen keine signifikanten Auswirkungen der verschiedenen Formulierungsalternativen auf die Suchergebnisse feststellen. Auch die Reihenfolge der beiden variierten Autosemantika schien keine Rolle zu spielen. Entscheidend für den Sucherfolg war vielmehr die hohe Übereinstimmung der Suchanfrage mit der konkreten Formulierung im Artikel.

Diese Beispiele zeigen, dass in der *Varietengrammatik* fast ausschließlich jene Artikel über Suchmaschinen gefunden und auch besucht werden, in welchen das grammatische Phänomen lexikalisiert ist, also sich in zwei oder mehr Schreibvarianten ausdrückt. In den Log-Files ließen sich jedoch keine verallgemeinerbaren Hinweise finden, auf welche Weise sich grammatische Phänomene beschreiben lassen, die keine eindeutige lexikalische Ausprägung haben. Auffallend dabei war, dass Fachtermini oder Umschreibungen der Phänomene bei den erfolgreichen Suchanfragen keine Rolle spielen. Das führt zu der nächsten Frage:

#### Welche Artikel werden nicht gefunden – und warum?

Diese Frage lässt sich nicht generell beantworten, auch weil das von Lemma zu Lemma und Wörterbuchgegenstand zu Wörterbuchgegenstand variiert. Am Beispiel des Eintrags „sein + am / beim / im + Infinitiv“<sup>33</sup> sei aber exemplarisch aufgezeigt, mit welchen Fragen wir versucht haben, herauszufinden, wo die Gestaltung und Struktur der Artikelseiten den Suchalgorithmen zuwiderlaufen. Das Lemma des untenstehend abgebildeten Artikels lautet wie geschrieben „sein + am / beim / im + Infinitiv“. Eine Suchanfrage „sein + am“ brachte den Artikel auf der zweiten Suchergebnisseite als Treffer, davor befanden sich thematisch fremde, aber auch passende Artikel (z.B. aus *wikipedia.de*, *pons.de*, *dict.cc*, *linguee.com*).<sup>34</sup> Die Suche nach „am + Infinitiv“ führt zu einem dritten Platz in den

Suchergebnissen. Es zeigt sich also, dass der Artikel durchaus auffindbar ist. Ein Abgleich im Keyword-Planer von *Google* prognostiziert jedoch nur zehn Treffer pro Monat für „am + sein“ und keinen Treffer für „am + Infinitiv“. Faktisch bedeutet das, dass der Artikel nicht gefunden wird (Seite drei für „am + sein“ reicht nicht aus)<sup>35</sup> oder mit einer potentiell passenden Suchformulierung nicht gesucht wird.

#### sein + am / beim / im + Infinitiv

**Allgemeines zum Variantenphänomen**

Tritt das Verb sein in Verbindung mit am, beim oder im und einem Infinitiv auf, wird damit eine so genannte Verlaufsform ausgedrückt. Natürlich soll auch ständig etwas **am Blüten sein**. (Der Tagesspiegel) Das Trio war **beim Spielen**, als der Hund zuschnappte. (Mitteldeutsche Zeitung) Ein Radparcours **ist im Entstehen**. (Nürnberger Nachrichten) Der Infinitiv ist ein Tätigkeitsverb (z. B. arbeiten, tanzen) oder ein Vorgangsverb (z. B. wachsen, schlafen) und wird in solchen Konstruktionen substantiviert, also großgeschrieben. Ausgedrückt wird mit dieser Konstruktion zum einen ein Vorgang oder ein Zustand im zeitlichen Verlauf, also als andauernd. Zum anderen kann sie eine wiederholte Tätigkeit oder einen wiederholten Vorgang bezeichnen: Die Leute sind bei Wind und Wetter **am Arbeiten**. (Neue Westfälische)

**Beispielebeleg**

- **sein + am + Infinitiv**
  - Sie **ist** wahnsinnig **am Wanken**, sie verliert beinahe alles, was sie hat. (Nürnberger Nachrichten)
  - **Erfreulicher Hintergrund ist** [...], dass der Verein dank **Zulaufs junger Sportler derzeit am Wachsen sei**. (Märkische Online Zeitung)
  - **Bereits jetzt schon sei** die Stadt **derart am Sparen**, dass kaum Luft zum Atmen bliebe. (Thüringer Allgemeine)
  - **Wir sind am Proben für eine abwechslungsreiche Show**. (Berliner Zeitung)
  - Die 34-jährige Fahrerin [...] **übersah dabei den heranfahrenden Motorroller, der gerade am Überholen war**. (inFranken.de)
  - Die **waren** bereits stark **am Qualmen** und wurden von den Feuerwehrleuten ins Freie gebracht. (Südkurier)
- **sein + beim + Infinitiv**
  - In der Ortschaft [...] **fuhr ein Pkw eine Frau um, die gerade beim Schneeschleichen war**. (Nordkurier)
  - **Derzeit sei man beim Ausforschen** eines Zeugen, der die notwendigen Auskünfte geben und Unterlagen zu dem mutmaßlichen Sozialbetrug vorlegen könne [...]. (Kronen Zeitung, Steiermark und Kärnten)
  - Der LKW **war schon beim Abbiegen**, als von hinten ein 30-jähriger Mann mit seinem Opel rechts am Lkw vorbeifahren wollte. (Thüringer Nachrichten)
  - **Der Junge, der gerade beim Spielen war**, wurde dabei mittelschwer verletzt. (Südtirol Online)
  - **Ihr Mann war beim Velofahren**. (1815 – Das Oberwalliser Nachrichtenportal)
- **sein + im + Infinitiv**
  - Die Infektionen **seien** aber schon wieder **im Abklingen**, so ein Sprecher des AKH. (Neue Westfälische)
  - Als sich die Tram Türen schlossen und die Bahn gerade **im Anfahren war**, sei die Zehnjährige auf die Fahrbahn gelaufen. (1815 – Das Oberwalliser Nachrichtenportal)
  - **Dass Gesellschaftstanz wieder im Kommen ist**, zeigt die Tatsache, dass sich vermehrt wieder junge und jungegebliebene Menschen für diese Art des Sports begeistern können. (Nürnberger Nachrichten)
  - **Jene 20 Maßnahmen [...] sind demnach überwiegend schon umgesetzt oder im Werden**. (Kronen Zeitung, Steiermark und Kärnten)
  - Die **Nitratwerte sind im Steigen**. (Kleine Zeitung, Steiermark und Kärnten)

Der Beginn des Artikels „sein + am / beim / im + Infinitiv“ der *Varietengrammatik*

<sup>33</sup> Vgl. [http://mediawiki.ids-mannheim.de/VarGra/index.php/Sein\\_%2B\\_am/\\_beim/\\_im\\_%2B\\_Infinitiv](http://mediawiki.ids-mannheim.de/VarGra/index.php/Sein_%2B_am/_beim/_im_%2B_Infinitiv)

<sup>34</sup> Diese Ergebnisse lassen sich nicht reproduzieren, weil die Suchalgorithmen die einzelnen Faktoren jeweils anders gewichten. Zwar haben wir mit anonymisierter IP gesucht, doch es ist nicht auszuschließen, dass beispielsweise ein erhöhtes Suchvolumen nach einem bestimmten Phänomen auch das Ranking beeinflusst.

<sup>35</sup> Wenn der Artikel nicht mindestens auf der dritten Seite der Suchergebnisse erscheint, gehen wir davon aus, dass er nicht gefunden wird.

Dahingegen werden beispielsweise für „beim Spielen“ ein Suchvolumen von 390 Anfragen pro Monat prognostiziert und für „im Kommen“ 110 Anfragen. Solche Suchvolumina könnten ein Hinweis darauf sein, dass bei der Suche nach Erklärungen für grammatische Phänomene konkrete sprachliche Beispiele genutzt werden. Das wäre naheliegend und nachvollziehbar, da man für eine Beschreibung und Benennung des Phänomens das entsprechende Fachvokabular kennen muss. Beispiele wie „beim Spielen“ finden sich auch im *Variantengrammatik*-Artikel, allerdings zeigen einige Test-Suchanfragen, dass der Artikel nicht über die Beispiele aufzufinden ist. Hier ist zu überlegen, mit welchen Maßnahmen die konkreten Beispiele für ein grammatisches Phänomen besser auffindbar werden. Das betrifft auch die Benennung des Phänomens mit „Verlaufsform“. Damit wird es von den Suchmaschinen nicht gefunden. Das Wort „Verlaufsform“ kommt in diesem Artikel nur zweimal vor (einmal im Text, einmal in einer Literaturangabe). Der Artikel „Verlaufsform“ ist dahingegen nur ein Verweisartikel ohne Text und damit sehr ungeeignet für Suchmaschinen. Mit der terminologischen Alternative „Progressiv“<sup>36</sup> findet man das Phänomen in den Suchmaschinen hingegen wesentlich besser. Dass „Verlaufsform“ kein geeignetes Suchwort ist, liegt auch daran, dass der Ausdruck sehr stark auf die Verbflexion ausgerichtet ist, und vor allem daran, dass es auch die Verlaufsform von Krankheiten gibt, die sich in den Suchergebnissen mit den grammatischen Verlaufsformen vermischt. Hier wird klar, dass bei der Wahl einer Terminologie zu fachlichen Überlegungen auch strategische hinzutreten können. Ohne großen Aufwand könnte man die fachsprachliche Alternative beispielsweise im ersten Satz des Artikels durch einen Zusatz einbringen: „damit wird eine sogenannte Verlaufsform (fachspr. auch „Progressiv“) ausgedrückt“.

Doch solche Überlegungen zu Textformulierungen stoßen an Grenzen, denn schließlich müsste man für alle Ausdrücke recherchieren, wie ihre Auffindbarkeit ist. Zwei Punkte lassen sich aber festhalten:

- 1) Beispiele spielen möglicherweise eine wichtige Rolle beim Formulieren von Suchanfragen, daher sollten Beispiele, die für das Phänomen relevant sind, unbedingt für den Artikel ausgewählt werden.
- 2) Eine terminologische Vielfalt bringt in Bezug auf die Suchmaschinenoptimierung im Zweifelsfall eher Vor- als Nachteile mit sich.

Die Texte alleine mit vielversprechendem Wortmaterial zu füllen, reicht aber nicht aus. Wir konnten beispielsweise auch mit dem seltenen Beispiel „war beim Velofahren“ den Artikel bei *Google* nicht finden. Das Beispiel wird wahrscheinlich durch den *Hummingbird-Algorithmus* abgewertet, weil es sich semantisch nicht in den Kontext fügt und insbesondere nicht zu den Überschriften passt.

### 5.1.2.3 Keywords müssen als wichtig erkennbar sein – die Überschriften

Wenn man Schlüsselwörter identifiziert hat, die ein Phänomen treffend beschreiben und auch ein zählbares Suchvolumen aufweisen, dann sollten diese an wichtigen Stellen im Text platziert werden. Als wichtige Stellen gelten die Überschriften, die von den Suchmaschinen als solche erkannt und ausgewertet werden. Dabei folgt die Gewichtung der Überschriften der Logik eines Text-Dokuments: Es gibt eine zentrale Überschrift und dann eine Struktur aus hierarchisch gewichteten Unterüberschriften. Die Überschrift sollte, wie in einem Text auch, den Inhalt des folgenden Abschnitts wiedergeben.

---

<sup>36</sup> Auch beispielsweise „am-Progressiv“ oder „beim-Progressiv“.

<h1>SEITENÜBERSCHRIFT</h1> (trägt den <h1>-Tag)

<h2>Hauptüberschrift</h2> (trägt den <h2>-Tag)

<h3>Unterüberschrift</h3> (trägt den <h3>-Tag)

Im SEO-Leitfaden von *Google* wird ausdrücklich darauf hingewiesen, dass es wichtig ist, korrekte und aussagekräftige Überschriften zu erstellen, was als Hinweis gewertet werden kann, wie wichtig die Titel für die Suchalgorithmen sind. Für einen lexikographischen Artikel bietet es sich an, das Lemma in die Seitenüberschrift (<h1>-Tag) zu schreiben. Das ist auch in der *Varietengrammatik* so gelöst. Die Hauptüberschriften „Allgemeines zum Variantenphänomen“ und „Beispielbelege“ (vgl. den vorherigen Screenshot) verschenken aber die Möglichkeit, der Suchmaschine hier mehr Relevantes über den folgenden Abschnitt mitzuteilen. Man erkennt, dass sich hier zwei Interessen gegenüberstehen: einerseits das lexikographische Interesse, wonach die Überschriften die Artikelstruktur sichtbar machen sollen, andererseits das Suchmaschinenoptimierungsbestreben, wonach Überschriften möglichst konkrete Informationen über den folgenden Abschnitt enthalten sollen. Auch hierfür lassen sich Lösungen finden, die beides kombinieren (bspw. graphische Lösungen oder Formatierungen, die keine Überschriften sind).

Abschließend soll nochmals die Seitenüberschrift beziehungsweise das Lemma des Artikels betrachtet werden. Diese Hauptüberschrift scheint eine größere Rolle für die Auffindbarkeit einer Seite zu spielen. Daher sollte man unbedingt überlegen, ob im Falle der *Varietengrammatik* neben dem Phänomen nicht auch dessen Bezeichnung in den Titel aufgenommen werden könnte, beispielsweise so: „sein + am / beim / im + Infinitiv (Progressiv bzw. Verlaufsform)“.

Es geht gerade in der Überschrift auch darum, einerseits das Phänomen genau zu benennen und andererseits, denkbare Suchanfragen möglichst gut abzudecken. Hierzu zwei weitere Beispiele aus der *Varietengrammatik*: Der Artikel „froh“ betrifft vor allem die mit diesem Adjektiv einhergehenden Präpositionen. Daher wäre das Phänomen mit „froh über / froh um / froh mit“ besser beschrieben und auch besser auffindbar. Außerdem ist davon auszugehen, dass Suchende beim Formulieren einer Frage auch die Präpositionen nennen, weil die alleinige Suche nach „froh“ kaum sinnvolle Treffer liefert.

The screenshot shows a dictionary entry for the word "froh". The entry includes the following text:

**froh** Adj. Präposition: **über** ÜBERALL; **um** A-MITTE, A-WEST, CH, D-SÜD, LIE, STIR

Beim Adjektiv *froh* können verschiedene Präpositionen stehen. *Froh* wird im ganzen deutschsprachigen Raum mit der Präposition *über* verbunden. *Froh über* ist in D-NORD, D-MITTE und im (Süd-)Osten von A die fast ausnahmslos verwendete Variante: *Und daher ist die Heimatstube über jede Geldspende froh.* (Volksstimme) **Adjektive und ihre Ergänzungen** *fn froh um* auf ein Gebiet von D-süd bis nach STIR konzentriert. Dabei ist *froh um* insbesondere in CH und – weniger häufig – in A-WEST gebräuchlich. In den umliegenden Regionen D-süd, A-MITTE und STIR sowie in LIE kommt *froh um* selten vor. *Zur Überbrückung wären wir daher froh um einen Zustupf von einer halben Million Franken.* (St. Galler Tagblatt).

Gelegentlich wird die Bedeutung von *froh über / um* durch *froh mit* ausgedrückt: *Das Elternpaar ist sehr froh mit den Angeboten für junge Familien in Beeskow.* (Märkische Online Zeitung).

[Frequenzangaben und Arealkürzel anzeigen]

Der Artikel „froh“ der *Varietengrammatik*

Ähnliches lässt sich beim Artikel „Iran“ beobachten, in dem es maßgeblich um die Frage des adäquaten Artikelgebrauchs geht. Dieses Phänomen wird aber mit der Seitenüberschrift „Iran“ nicht abgebildet, und der Artikel wird bei *Google* auch nicht gefunden, obwohl die laut Keyword-Planer in Suchanfragen verwendeten Formulierungen wie „Iran mit oder ohne Artikel“ bzw. „im Iran oder in Iran“ im Artikeltext durchaus enthalten sind.

## Iran

**Iran** Artikelgebrauch: **ohne Artikel** A, CH, D; **mit Artikel** ÜBERALL; –(-s)–(-s), kein Plural

[Frequenzangaben und Arealkürzel anzeigen]

Der Eigenname *Iran* wird mit oder ohne Artikel gebraucht (vgl. Artikelgebrauch bei Länder- und Landschaftsnamen). Mehrheitlich wird ÜBERALL die Variante mit Artikel verwendet: *Selbst in **den Iran** wurden 20 alte gelbe Bahnen geliefert.* (Rheinische Post). In CH ist daneben auch die artikellose Variante gebräuchlich, im übrigen deutschsprachigen Raum kommt diese Variante nur selten vor: *Unter anderem mit dem Roten Kreuz in **Iran**, im Irak und in der Türkei, auf Kurdisch Hesya Sore Kurdistan genannt.* (Aargauer Zeitung). *Die Bibliothekarin ist bekennende Bahai, eine Religion, die ursprünglich aus **Iran** stammt.* (St. Galler Tagblatt).

### Beispielbelege

#### • Iran ohne Artikel

- Die Oberwalliser Gruppe "Katze Steffan und das Lügenorkestar" überzeugt mit viel Charme und Musik von Istanbul **bis Iran**. (1815 – Das Oberwalliser Nachrichtenportal).
- Doch der 28-Jährige, der seit sechs Jahren in der Schweiz lebt, will länger reisen – **nach Iran** und nach Australien [...]. (St. Galler Tagblatt).
- Der Fingerhut ist aus Tschechien, die Maulbeeren **aus Iran**, Estragon aus Frankreich. (Der Tagesspiegel).
- Die Meldungen über das Erdbeben **in Iran** vom Sonntag machen betroffen. (Basellandschaftliche Zeitung).
- Wenn es zu keinen Zwischenfällen kommt, findet ihr nächstes Weltcuprennen am letzten Augustwochenende **in Iran** statt. (St. Galler Tagblatt).

#### • Iran mit Artikel

- [...] deren Graphic Novel *Persepolis* über ihre Kindheit **im Iran** zu einem Bestseller wurde. (Der Standard).
- Nach ihrer Flucht aus Afghanistan kamen sie über **den Iran** in die Türkei. (Oberbayerisches Volksblatt).
- Im Jahr 2009 entsteht **im Iran** die "grüne Bewegung", [...]. (Stuttgarter Zeitung).
- [...] auch eine Geigerin aus **dem Iran** sowie hoffnungsvolle Musiker aus Israel [...]. (Thüringer Allgemeine).
- Die beiden Frauen aus **dem Iran** und aus Sri Lanka werden sich am 10. November gemeinsam die Kochschürze umbinden. (Berner Zeitung).

Der Artikel „Iran“ der *Variantengrammatik*

Hier sind aus Sicht der Suchmaschinenoptimierung folgende Veränderungen zu überlegen, die sich auch positiv auf die Usability auswirken können:

- 1) Statt „Iran“ in der Überschrift bspw. „Iran, mit oder ohne Artikel“
- 2) Statt „Beispielbelege“ bspw. „Belege für *Iran* mit und ohne Artikel“. Die Aufzählungspunkte „*Iran* ohne Artikel“ und „*Iran* mit Artikel“ könnten ebenfalls als Unterüberschrift (h3-Tag) formatiert werden und würden damit mehr Relevanz für die Suchalgorithmen bekommen. Dort könnten zudem die wichtigsten Formulierungen (bspw. „im Iran, in den Iran, aus dem Iran“) aufgegriffen werden.

Diese Veränderungen können so gestaltet werden, dass sie sich einfach und teilweise automatisiert umsetzen lassen (bspw. mit Textbausteinen wie „Belege für ...“). Außerdem würde sich ohne Aufwand die Usability erhöhen, weil durch diese Veränderungen der Informationsgehalt der Überschriften steigt.

### 5.1.2.4 Der <title>-Tag informiert über den Inhalt der Seite

Der <title>-Tag wirkt sich sowohl auf die Auffindbarkeit als auch auf die Sichtbarkeit aus: Er wird von den Suchalgorithmen



berücksichtigt und gibt gewissermaßen das Thema der Seite vor. Zudem wird der Titel in den Suchmaschinenergebnissen dargestellt, und er wird bei geöffneter Seite im Browser als Seitentitel (in dem Screenshot oben links) in einem Tab angezeigt. In der *Variantengrammatik* hat der <title>-Tag das Format: <title>Seitenüberschrift – Variantengrammatik des Standarddeutschen</title>. Das folgende Beispiel zeigt allerdings, dass gerade bei der Tab-Darstellung in Browsern nur ein Teil des Seitentitels dargestellt wird: Das gilt es gegebenenfalls zu berücksichtigen, wobei stets sein doppelter Nutzen für Suchmaschinen und für die Seitendarstellung beachtet werden sollte.

### 5.1.2.5 Bilder für die Auffindbarkeit nutzen

Die Artikel der *Varietengrammatik* zeichnen sich insbesondere durch die Visualisierungen der regionalen Varianten aus, wie hier am Beispiel „Vermögenssteuer / Vermögensteuer“ erkennbar ist. Solche Bilder werden von den Suchmaschinen ebenfalls indiziert und ausgewertet. Dafür müssen sie eindeutig benannt sein. Der SEO-Leitfaden von *Google* empfiehlt, einen „beschreibenden Dateinamen“ zu vergeben und ein „alt“-Attribut anzulegen. Das „alt“-Attribut ist eine im HTML-Code (also auf der Website nicht sichtbar) gespeicherte Information zum Bild, die nur dann angezeigt wird, wenn das Bild nicht darstellbar ist.<sup>37</sup> Aus dem „alt“-Attribut kann ein Suchalgorithmus weitere inhaltliche Informationen über ein Bild gewinnen und es entsprechend besser in der Bildersuche berücksichtigen. Mit

**Vermögenssteuer / Vermögensteuer**

Vermögenssteuer die, -, -n UBERALL; Vermögensteuer die, -, -n D

Die Zusammensetzungen Vermögenssteuer und Vermögensteuer stehen mit oder ohne s-Fuge ohne Bedeutungsunterschied nebeneinander. Die Variante Vermögensteuer ist insgesamt häufiger, sie wird im gesamten deutschsprachigen Raum mehrheitlich bis fast ausschließlich verwendet. Nachdrücklich forderte eine Vermögenssteuer, um die Kluft zwischen Arm und Reich zu glätten. (Kurier, Oberösterreich). Etwas weniger häufig tritt in D, A-west und A-ost aber die Variante ohne s-Fuge auf. Neben einem höheren Spitzensteuersatz planen beide Parteien auch, die Vermögenssteuer wieder einzuführen. (Der Tagesspiegel)

**Beispielbelege**

- **Vermögenssteuer**
  - Die Ressortchefin plädierte einmal mehr für Erbschafts- und Vermögenssteuern. (Wiener Zeitung).
  - Beim Kindererbzug, bei der Einkommenssteuer und der Vermögenssteuer beschloss eine Mehrheit der Kommission jedoch weiter reichende Erleichterungen. (Aargauer Zeitung).
  - Zudem habe die Stadt weniger Einnahmen bei der Einkommens- und Vermögenssteuer verbucht und eine geringere Rückvergütung für Thurgauer Bürger erhalten. (St. Galler Tagblatt).
  - Aber auch eine Vermögenssteuer für Reiche sowie gleiche Löhne für gleiche Arbeit, eine Verbesserung des Schulsystems und die Sozialverträglichkeit der Energiepreise seien wichtige Themen gewesen. (Nürnberger Nachrichten).
  - Wenn das alles geschafft ist, dann sollte einer Vermögenssteuer nichts mehr im Wege stehen. (Kronen Zeitung, Steiermark und Kärnten).
- **Vermögensteuer**
  - Nach Jahren zinkender Spitzensteuersätze und der Abschaffung von Vermögensteuern kommt im Eurogebiet das Thema höhere Steuern für Bezieher höherer Einkommen und Vermöge zunehmend auf die politische Agenda [...]. (Der Tagesspiegel).
  - Vor dem Hintergrund einer "ehrlichen Art der Staatsfinanzierung" sollte daher die Vermögensteuer wieder eingeführt und der Spitzensteuersatz erhöht werden, forderte der Minister. (Rheinische Post).
  - Vermögenssteuer und Finanztransaktionssteuer könnten insgesamt bis zu 100 Milliarden Euro einbringen. (Thüringer Allgemeine).
  - Die Vermögensteuer, die allein den Ländern zusteht, soll jährlich 11,5 Milliarden Euro in die Kassen spülen. (Stuttgarter Zeitung).
  - Strikt abgelehnt wird von der ÖHv die Einführung einer Vermögensteuer. (Troler Tageszeitung).

**Karten**

Relative Auftretenshäufigkeit von Vermögenssteuer / Vermögensteuer innerhalb der einzelnen Areale

Relative Auftretenshäufigkeit von Vermögenssteuer innerhalb der einzelnen Areale im Verhältnis zu Vermögensteuer (in %)

Relative Auftretenshäufigkeit von Vermögensteuer innerhalb der einzelnen Areale im Verhältnis zu Vermögenssteuer (in %)

Visualisierung der regionalen Variation im Artikel „Vermögenssteuer“

eben dieser Bildersuche erschien bei unseren Testaufrufen bei *Google* eine Grafik aus dem Artikel „Vermögenssteuer / Vermögensteuer“, als wir nach „Vermögensteuer“ suchten.

Lässt man sich als Suchergebnis Bilder anzeigen, dann werden die Ergebnisse dargestellt wie im Beispiel zu „anderthalb“, und es

Google vermögenssteuer

Umgefahr 333.000 Ergebnisse (0,54 Sekunden)

de.wikipedia.org • wiki • Vermögensteuer •  
**Vermögensteuer – Wikipedia**  
 Die Vermögensteuer ist eine Steuer auf das Gesamtvermögen eines Steuerpflichtigen. Ihre Bemessungsgrundlage umfasst in der Regel das nach Abzug der ...  
 Historische Entwicklung · Vermögensteuer in ... · Vermögenbezogene ...

**Nutzer fragen auch**

- Wie viel Prozent Vermögensteuer?
- Wie hoch ist die Vermögensteuer in der Schweiz?
- Wie hoch ist die Vermögensteuer in Deutschland?
- Warum wurde die Vermögensteuer abgeschafft?

Ein Bild aus der *Varietengrammatik* erscheint in der Ergebnisdarstellung bei *Google*.

zeigt sich, dass die *Varietengrammatik* hier eine sehr gute Positionierung hat.

Die Bildersuche ist für digitale Sprachressourcen eine gute Möglichkeit, um die Auffindbarkeit zu erhöhen, gerade wenn entsprechende Grafiken sowieso verfügbar sind. Das (Um-)Benennen der Dateinamen und das Anlegen eines suchmaschinenoptimierten „alt“-Attributs sind ein vergleichsweise kleiner Aufwand, der sich gegebenenfalls sogar

Google anderthalb

anderthalb jahren anderthalb stunden marinetaucher stirbt sharon frucht eineinhalb thailändischer marinetaucher royalty free apfel dreamstime

anderthalb / eineinhalb - V...  
 anderthalb / eineinhalb - V...  
 anderthalb / eineinhalb - V...  
 ANDERTHALB Synonym L...  
 61 Anderthalb Apfel Fotos - Kostenlose un...  
 anderthalb | LinkedIn  
 Anderthalb Balthargam...  
 Weren anderthalb 1,5 ist...  
 Anderthalb Jahre | Julie Fahrheit  
 Wörter unter der Lupe  
 anderthalb Sharon Frucht stockfoto. Bild von ...  
 Anderthalb - its internationale filmschule köln

Die Karten zur Variantenverteilung von „anderthalb“ in der *Google*-Bildersuche.

<sup>37</sup> Das „alt“-Attribut spielt auch in Bezug auf die Usability eine Rolle, weil damit bei der Benutzung eines Screenreaders – also wenn man sich eine Website automatisiert vorlesen lässt – das Bild umschrieben werden kann.



entsprechen können. Zudem werden gut aufgebaute Verlinkungen die Usability erhöhen. Das gilt sowohl für Textlinks als auch für Verlinkungen von Bildern.

### 5.1.2.7 Das Top-Ranking als hervorgehobenes Snippet

Neben der üblichen Auflistung der Suchmaschinenergebnisse gibt es bei *Google* auch die Darstellung als hervorgehobenes Snippet. In diesem Fall wird das Suchergebnis mit den Bilddateien graphisch herausgehoben in einem Kasten über allen anderen Suchergebnissen dargestellt. Diese Position kann nicht käuflich erworben werden und sie lässt sich auch nur indirekt beeinflussen. Voraussetzung hierfür ist eine hohe Klickrate: Eine Website muss signifikant häufig angeklickt werden, damit sie in der Ergebnisliste zu einer Suchanfrage erscheint. Dann wählt der Algorithmus die Website für eine hervorgehobene Darstellung aus. Besonders wichtig ist dabei, dass ein thematisch aussagekräftiger Textausschnitt auf der Seite vorhanden ist. Die Darstellung dieses Snippets wird im Kapitel zur Sichtbarkeit (Kap. 5.2) nochmals thematisiert.



### Zusammenfassung: Auffindbarkeit erhöhen und Usability steigern

Es wurde aufgezeigt, dass für Maßnahmen zur besseren Auffindbarkeit in Suchmaschinen keinesfalls teure Kampagnen nötig sind und dass es sich lohnt, die Auffindbarkeit von Informationen in den Fokus zu rücken. Meiner Einschätzung nach steht dieser lohnenden Veränderung gerade die traditionelle und oftmals unbewusste Orientierung auf das Werk im Weg. Die daraus resultierende Schlussfolgerung war, weniger die Startseite und die Seiten mit Umtexten zu optimieren und stattdessen die Strukturen und Formulierungen auf den Artikelseiten in den Blick zu nehmen. Viele sinnvolle Optimierungen – wie beispielsweise die Überschriften – betreffen strukturelle Positionen der Artikel-Mikrostruktur und können damit zumindest halbautomatisiert vorgenommen werden. Auch Bilder zu integrieren kann auf ökonomische Weise bewerkstelligt werden. Neu entstehende Sprachressourcen sollten auf jeden Fall die Ansprüche der Algorithmen bei der Konzeption der Mikrostruktur berücksichtigen. Hierfür ist schon alleine der erwähnte SEO-Leitfaden von *Google* eine gute Grundlage.

Suchmaschinenoptimierte Artikel müssen in keiner Weise nachteilige Auswirkungen mit sich bringen: Entsprechende Darstellungen von Strukturen und bewusst gewählte Formulierungsveränderungen können durchaus Vorteile für die Nutzenden mit sich bringen, und sie stellen den begrifflichen Diskurs zu einem Thema möglicherweise sogar besser dar. Insgesamt kann in vielen Fällen die Verbesserung der Auffindbarkeit mit der Verbesserung der Usability in Einklang gebracht werden. Und es zeigt sich deutlich, dass lexikographische Artikel aufgrund ihrer informativen Grundstruktur, ihrer thematischen Konzentration und ihrer logisch-immanenten Vernetzung bestens den Ansprüchen von aktuellen Suchmaschinen-Algorithmen entsprechen.

Gerade wenn die lexikographischen Artikel alleinstehende Informationen liefern, seltene Phänomene betreffen oder mit feststehenden Ausdrücken adressiert werden können, sind sie bestens dafür geeignet, potentiell Interessierte für das Angebot einer digitalen Sprachressource zu gewinnen. Wie

dargestellt, werden einzelne, aus welchen Gründen auch immer häufig besuchte Artikel zum Hauptzugang eines Werks. Der klassische Weg, über eine Startseite ein Werk zu besuchen, verliert damit an Bedeutung, weil, so wurde an der *Variantengrammatik* gezeigt, für werkrelevante Schlagworte nur selten passende *Google*-Suchanfragen zu verzeichnen sind.

## 5.2 Sichtbarkeit – Das Werk hinter den Informationen wahrnehmen

Neben der Auffindbarkeit ist es das Ziel von Optimierungen, dass die Besucherinnen und Besucher künftig direkt die Seiten der Sprachressource ansteuern. Auf jeden Fall aber sollten Suchende auf den „richtigen“ Eintrag in der Ergebnisliste von Suchmaschinen klicken. Dazu muss das, was dort sichtbar wird, so gestaltet sein, dass die Qualität und der Nutzen der verlinkten Sprachressource erkennbar wird. Die besondere Chance für digitale Sprachressourcen liegt darin, die vielfältigen qualitativen Informationen aus ihren Artikeln zu nutzen, um Menschen mit ganz unterschiedlichen Fragestellungen für das Werk zu gewinnen und zu begeistern.

### 5.2.1 Was soll im Suchergebnis dargestellt werden?

Am sinnvollsten erscheint es, diese Frage aus der Situation der Suchenden heraus zu betrachten: Jemand formuliert eine sprachbezogene Suchmaschinenanfrage und erhält eine Auswahl an Seiten, die potentiell eine Antwort geben können. An dieser Stelle kann man zwei mögliche Situationen unterscheiden:

- 1) Das Snippet einer oder mehrerer Seiten gibt eine ausreichend zufriedenstellende Antwort.
- 2) Aus den Snippets lässt sich für die Person keine zufriedenstellende Antwort herauslesen.

Welche Gründe könnte es geben, dass eine Person in der ersten Situation trotzdem auf eine der Seiten klickt? Naheliegend ist Folgendes:

- Die Person möchte mehr erfahren, vielleicht auch, weil das Snippet mitten in einem Satz abgeschnitten wurde.

Aus den bereits erwähnten Studien zu den Erwartungen an Nachschlagewerken lassen sich weitere Gründe ableiten:

- Die Person möchte schauen, ob Information und Quelle aktuell, verlässlich und vertrauenswürdig sind.
- Die Person möchte die Information zitieren und benötigt genauere Angaben.
- Die Person möchte überprüfen, ob die Information ihrem Anliegen passt.

Wenn die Situation eintritt, dass das Snippet keine zufriedenstellende Antwort liefert, hängt der weitere Verlauf der Suche wahrscheinlich stark von der Konkurrenzsituation ab: Wenn eine andere Seite relevante Informationen im Snippet bietet, ist diese Seite im Vorteil. Liefert keines der Suchergebnisse eine zufriedenstellende Antwort, dann müssen Suchende anhand der verfügbaren Informationen in der Titelzeile, der URL und dem Snippet entscheiden, welche Quelle geeignet erscheint. Für diesen Fall sollte die Darstellung der Ressource in der Suchergebnisaufstellung den Eindruck vermitteln, dass die betreffende Quelle eine für die Fragestellung geeignete ist.

Es gibt bislang keine Erkenntnisse darüber, welche Signale hier wirksam sind. Hier sei aber eine offene Liste mit Anforderungen an die Darstellung in der Ergebnisliste formuliert:

- Nennung des Namens der Sprachressource (zwecks Bekanntheit und Markenbildung)
- Information über den Gegenstand der Sprachressource (früher: Wörterbuchgegenstand)
- Signale zur Verlässlichkeit der Quelle (zwecks Vertrauensbildung) und keinesfalls gegenteilige

Der dritte Punkt ist vage, weil unklar ist, welche Signale vertrauensbildend wirken. Für wissenschaftlich erarbeitete Sprachressourcen bietet es sich natürlich an, dieses Qualitätsmerkmal zu betonen. Wenn feststeht, welcher Name, welcher Gegenstand und welche Verlässlichkeitssignale benannt werden sollen, kann man überlegen, wie diese Angaben sich auf die zwei oder drei Bereiche des Suchergebniseintrags verteilen lassen. Dabei ist zu bedenken, dass sich der Inhalt des Snippets nicht verlässlich bestimmen lässt. Für die Darstellung der *Variantengrammatik* wurde beispielsweise nebenstehende Darstellung im Team diskutiert.

[öfter / öfters | grammatische Varianten in D, A, CH, ...](https://www.variantengrammatik.net/artikel/öfter_öfters)  
[https://www.variantengrammatik.net/artikel/öfter\\_öfters](https://www.variantengrammatik.net/artikel/öfter_öfters)  
Das Adverb *öfter* sowie die mit -s abgeleitete Variante *öfters* werden ohne Bedeutungsunterschied nebeneinander gebraucht. Im gesamten deutschsprachigen ...

Hier wurde die Entscheidung getroffen, im Titel primär den Gegenstand der Sprachressource zu beschreiben, insbesondere weil der Ausdruck „Variantengrammatik“ mehrdeutig und allgemeinsprachlich auch nicht klar ist. Außerdem bringt „Standarddeutsch“ für potentielle Besuchende nicht zum Ausdruck, dass das gesamte deutschsprachige Gebiet gemeint ist. Der Werk-Titel hingegen wurde auf einen klar formulierten Link ausgelagert: Damit wird – wie für eine digitale Quelle passend – die Website „www.variantengrammatik.net“ in den Vordergrund gerückt. Gleichzeitig wurde die Verzeichnisstruktur so verändert, dass sie einem Wörterbuch ähnelt: Es gibt das Werk, dieses enthält „Artikel“ und der hier betroffene Artikel heißt „öfter\_öfters“. Was in dieser Darstellung aufgrund der begrenzten Fläche fehlt, ist ein Hinweis auf den wissenschaftlichen Hintergrund.

Mit Sicherheit kann man unter Abwägung aller Interessen zu unterschiedlichen Darstellungen kommen. Wichtig ist, dass die Darstellung mit dem Fokus auf potentiell Besuchende und auf die Ziele der Website-Betreibenden entwickelt wird und dass sie sich nicht aus einer Gepflogenheit oder beispielsweise einer unreflektierten Verzeichnisstruktur ergibt. Auch sollte jeweils der gesamte Eintrag betrachtet werden, sodass sich beispielsweise die URL und der Titel ergänzen können und nicht wie im nächsten Screenshot dreimal untereinander derselbe Ausdruck vorkommt. Letztlich ist das Ziel, dass Nutzende erkennen und entscheiden können, ob die Seite und die Inhalte der Seite für sie relevant sind. Nicht außer Acht gelassen werden sollte, dass bei einer adäquaten Darstellung der Seiteninhalte im Snippet die Wahrscheinlichkeit steigt, dass die Suchenden die Seite gar nicht mehr besuchen. Für diesen Fall wäre es nachhaltig, wenn sie zumindest unterbewusst mitnehmen, welche Quelle die für sie nützliche Information geliefert hat.

## 5.2.2 Sichtbarkeit in der Suchmaschine

Welche Informationen zu bzw. aus einem digitalen Werk werden in der klassischen Darstellung in der *Google*-Ergebnisliste sichtbar? Betrachten wir wieder das nebenstehende Beispiel:

[öfter / öfters – Variantengrammatik des Standarddeutschen](#)  
[mediawiki.ids-mannheim.de](#) > [VarGra](#) > [index.php](#) > [Öfter\\_ > \\_öfters](#) ▾  
Das Adverb *öfter* sowie die mit -s abgeleitete Variante *öfters* werden ohne Bedeutungsunterschied nebeneinander gebraucht. Im gesamten deutschsprachigen ...

Es gibt hier drei Bereiche – Titel, URL, Snippet –, auf die Betreibende einer Website einwirken können. Das ist relativ wenig Platz, der möglicherweise nicht ausreicht, um alle gewünschten Informationen über das verlinkte Werk und die Seite unterzubringen. Es gilt also genau zu überlegen, was man über die digitale Sprachressource mitteilen möchte. Im dargestellten Beispiel finden wir das Lemma des Artikels und den Titel der Sprachressource (blau), den Link zum Artikel (grün) und einen Auszug aus der Website, das Snippet (schwarz).

Die blaue Überschrift wird direkt aus dem schon weiter oben beschriebenen <title>-Tag übernommen. Entsprechend scheint es – in Bezug auf die Sichtbarkeit – die wichtigste Funktion des <title>-Tags zu sein, hier eine sprechende und informative Angabe zu liefern. Wie auch im gezeigten Beispiel setzt sich der <title>-Tag aus einer automatisiert übernommenen Angabe der Seite (hier der Seitenüberschrift mit der Lemmaangabe) und einem statischen Zusatz (hier „–Variantengrammatik des Standarddeutschen“) zusammen. Dieser Eintrag lässt sich vollumfänglich von den Betreibende der Website gestalten.

Die grüne URL ist die Adresse der Seite, auf die verlinkt wird. Diese Adresse zeigt die Domain (hier: „mediawiki.ids-mannheim.de“) und dann die Ordnerstruktur des Servers: Die Daten liegen also in einem Verzeichnis „VarGra“ mit einer Startseite „index.php“ und der Unterseite mit dem Namen „öfter\_/\_öfters“<sup>38</sup>. Es wird deutlich, dass diese URL-Struktur sich nur durch Umstrukturierungen und Umbenennungen der Ordnerstrukturen verändern lässt. Noch aufwändiger ist es, die Domain zu ändern, weshalb es sinnvoll ist, diese Struktur schon von Anfang an zu planen und keine kryptischen Abkürzungen wie „VarGra“ als Verzeichnisname zu wählen, sondern beispielsweise „Variantengrammatik“.

Das Snippet, d.h. der schwarze, zweizeilige Text, liefert den relevanten Inhalt der Website. Es werden algorithmisch – also nicht nachvollziehbar – Ausschnitte ausgewählt, die eine bestmögliche Antwort auf die Suchanfrage geben sollen. Das zeigt sich daran, dass sich bei einer veränderten Suchanfrage auch dieser Ausschnitt ändert. Unsere Testanfragen via *Google* haben gezeigt, dass die Algorithmen aus den nicht suchmaschinenoptimierten Artikeln der *Variantengrammatik* in der Regel gut passende und sinnhafte Ausschnitte auswählen. Solch wenig aussagekräftige Snippets, wie hier von der dwds.de-Seite

entnommen, ließen sich für die *Variantengrammatik* nicht beobachten. Das unterstreicht, dass ausformulierte lexikographische Texte für die Suchmaschinendarstellung eher vorteilhaft sind.

[www.dwds.de](http://www.dwds.de) › [zensieren](#) ▼

[zensieren](#) – [Schreibung](#), [Definition](#), [Bedeutung](#), [Etymologie](#) ...

[zensieren](#) – [Schreibung](#), [Definition](#), [Bedeutung](#), [Etymologie](#), [Synonyme](#), ... ↗ [unterbinden](#) · ↗ [untersagen](#) · ↗ [verboten](#) · [zensieren](#) ○ [zensurieren](#) österr.

Während bei den Artikelseiten jeweils gute Textausschnitte präsentiert wurden, hat die Untersuchung der Startseite der *Variantengrammatik* die Problematik der Snippets vor Augen geführt:

Hier war kein adäquater Text vorhanden und entsprechend problematisch ist der Ausschnitt. Eine solche Beschreibung gibt

## Variantengrammatik des Standarddeutschen

[mediawiki.ids-mannheim.de](http://mediawiki.ids-mannheim.de) › [VarGra](#) › [index.php](#) › [Start](#) ▼

Bis Ende November 2019 werden weitere Artikel ergänzt. Auch Änderungen an den bestehenden Artikeln sind bis dahin möglich. Zur Zitierweise siehe hier.

<sup>38</sup> Der Slash im Seitennamen „/“ wird hier von Google fälschlicherweise als „>“ aufgelöst, wie es bei der Angabe der URLs üblich ist. So sieht es aus, als ob eine Unterseite mehr existiert, als es tatsächlich der Fall ist. Auch das müsste man bei der Vergabe von Seitennamen berücksichtigen.

weder Auskunft über die Seiteninhalte noch repräsentiert sie das Werk in einer Form, die zu einem Besuch der Website einladen würde. Für solche Fälle, bei welchen der Algorithmus einen passenden zitierbaren Inhalt auf der Seite identifizieren kann, ist das „description“-Meta-Tag vorgesehen. Diese im HTML-Code versteckte Beschreibung einer Website war auf der Startseite der *Varietengrammatik* nicht angelegt. Daher wurde die auf der Website einzige Information angezeigt, und diese war unpassend, nicht informativ und veraltet. Im „description“-Bereich des HTML-Codes kann das Thema der betreffenden Seite mit ein oder zwei Sätzen beschrieben werden. Diese Beschreibung wird einerseits bei der Suche als Auskunft über die Themen der Seite berücksichtigt (Stichwort „Auffindbarkeit“) und es wird als Snippet ausgewählt, wenn sich keine andere Textstelle auf der Seite anbietet. Im SEO-Leitfaden von *Google* wird dringend empfohlen, ein aussagekräftiges „description“-Meta-Tag für jede Seite anzulegen, und für weiterführende Informationen zur Gestaltung von Snippets wird auf den Webmaster-Blog<sup>39</sup> verwiesen. Das Snippet der *Varietengrammatik* wurde durch die Überarbeitung des „description“-Meta-Tags aussagekräftiger:

[www.variantengrammatik.net](http://www.variantengrammatik.net) ▾

## Varietengrammatik des Standarddeutschen

Die **Varietengrammatik** informiert auf wissenschaftlicher Basis über Unterschiede in der Grammatik der deutschen Sprache in Deutschland, Österreich, der ...

Gerade bei Seiten mit Bezug auf das Werk (bspw. die Startseite) scheint das „description“-Meta-Tag sehr wichtig zu sein, da man damit konzentriert die Aussage einer Seite oder des ganzen Werkes darstellen kann. Bei den Artikelseiten ist es möglicherweise weniger wichtig – weil mehr unterschiedlich relevante Informationen im Artikel stecken – und es ist wegen des konzentrierten Inhalts auch aufwändiger zu erstellen.

### 5.3 Usability – Bieten, was versprochen ist

In Kap. 3.5 wurde aufgezeigt, dass unterschiedliche Ansprüche an Nachschlagewerke (vgl. Müller-Spitzer 2016a, 320) nicht unvereinbar in der Umsetzung sein müssen. Vielleicht muss man die Zielsetzungen hierarchisiert betrachten. So dürfte es Konsens sein, dass die Nutzenden eine Antwort auf ihre Frage finden sollten und dass dies besser schnell als langsam geschehen sollte. An diese Antwort sind verschiedene Ansprüche gerichtet, unter anderem die von Müller-Spitzer genannten wie Verständlichkeit, Aktualität und Verlässlichkeit. Auch dazu gibt es möglicherweise keine größeren Meinungsverschiedenheiten. Es gehört also in den Aufgabenbereich der Usability, dass eine digitale Sprachressource diese qualitativen Werte einhält, sie faktisch unter Beweis stellt und sie auch für die Nutzenden einfach prüfbar macht. Das gilt insbesondere, wenn diese Werte schon im Eintrag bei einer Suchmaschine versprochen werden. In den Log-Files der *Varietengrammatik* ist ein Nutzungsmuster auffällig: Via *Google* gelangt jemand direkt auf einen Artikel, verweilt dort eine gewisse Zeit und klickt dann auf die Startseite oder auf die Seite „Über uns“. Ich deute das als den Versuch, mehr über das Projekt zu erfahren, beispielsweise um herauszufinden, wie verlässlich die angebotenen Informationen sind. Ein naheliegender Schritt der Usability ist daher, Artikelseiten entsprechend zu gestalten und zu ermöglichen, dass sie ausgehend von anderen Artikelseiten oder Umtexten gut erreichbar sind.

<sup>39</sup> <https://webmasters.googleblog.com/2007/09/improve-snippets-with-meta-description.html>

Doch benutzendenfreundlich könnte auch sein, wenn möglichst viele der in den Umtexten versteckten Qualitätsmerkmale direkt auf den Artikelseiten benannt und entsprechend positioniert werden. Das würde den Umweg auf eine weitere Seite ersparen und damit die Möglichkeit eröffnen, aus dem Artikel heraus eine weiterführende Information (bspw. einen Kategorie-Artikel) zu verfolgen (vgl. auch das Kap. 5.3.2.2). Es wurde bereits angedeutet, dass manche mit der inhaltlichen Antwort im Snippet schon zufrieden ist und die Seite nur noch besuchen, um die Qualität der Inhalte zu beurteilen. Um den benannten Ansprüchen gerecht zu werden, könnte überlegt werden, Quellen und Methoden zu nennen (Verlässlichkeit), Aktualisierungsdaten oder Versionierungen anzuzeigen (Aktualität) und in verständliche Darstellungen und Texte zu investieren.

Es gibt meines Wissens keine Erkenntnisse dazu, wie und woran Nutzende von Nachschlagewerken deren Qualität, Aktualität oder andere Werte, die ihnen wichtig sind, beurteilen. Man kennt beispielsweise die Hinweise auf *Wikipedia*, wenn ein Artikel als überarbeitungsbedürftig markiert ist. Auf *dwds.de* gibt es einen ähnlichen Hinweis, der letztlich auch ein Aktualitätsversprechen impliziert und den redaktionellen Prozess transparent macht.

Ähnliches gilt für Einträge wie „maschinell ausgesucht aus dem DWDS-Korpus“<sup>40</sup>, die offenlegen, dass Informationen nicht redaktionell, sondern automatisiert generiert wurden. Diese Transparenz scheint sehr wichtig, wenn es darum geht, die Vertrauenswürdigkeit einer Ressource zu stärken.

 Dieser Artikel ist nicht auf dem aktuellen Stand. Er wird im Rahmen des DWDS-Projekts von der Redaktion überarbeitet.

## Anfrage, die

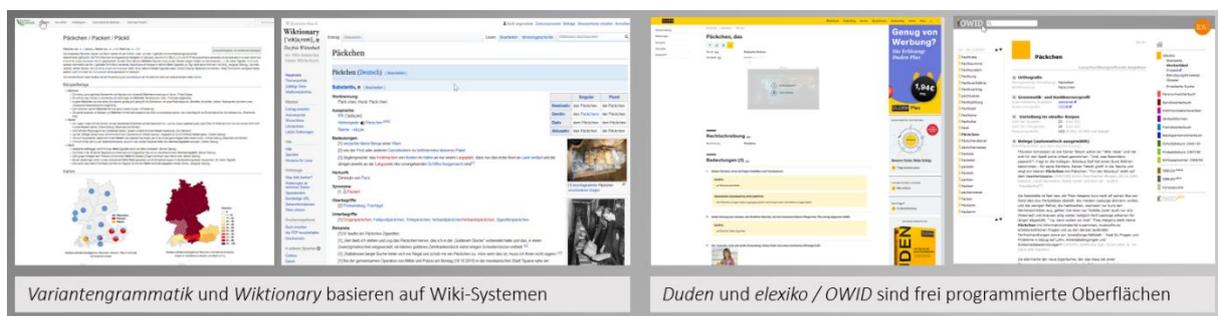
Welche auch immer diese wichtigen Signale sind, es reicht – wie gesagt – nicht aus, sie auf der Startseite – quasi dem Einband des digitalen Werks – zu platzieren. Wenn ein Großteil der Besuchenden von den Suchmaschinen direkt auf die relevanten Artikelseiten geführt wird, dann müssen solche relevanten Angaben dort zu finden sein. Auf den Artikelseiten der *Varietengrammatik* beispielsweise lässt sich von Personen, die mit den Informationen im Logo links oben nichts anfangen können, nichts über den Wörterbuchgegenstand herausfinden. Die Autorenschaft ist am Artikelende genannt, was möglicherweise aber eine schlecht wahrnehmbare Position ist. Dass die Informationen einem wissenschaftlichen Projekt entstammen, ist auf der Artikelseite nicht ersichtlich. Ein Änderungsdatum des Artikels wird zur Überprüfung der Aktualität ebenfalls am Ende des Artikels angegeben. Orientiert man sich an den genannten Umfragen, wonach Aktualität und Verlässlichkeit wesentliche Erwartungen an eine digitale Sprachressource darstellen, dann würde man die genannten Informationen möglicherweise prägnanter auf der Artikelseite positionieren.

Man muss also in Betracht ziehen, dass Besuchenden ihre dringendste inhaltsbezogene Frage schon (teilweise) im Snippet beantwortet wurden. Möglicherweise hat der Grund für den Besuch der Sprachressource auch mit deren Inhalten gar nichts zu tun. Um diese Diversität aufzufangen, kann es hilfreich sein, die weiter oben vorgestellte Unterscheidung der Wahrnehmung der Sprachressource als Informationsspeicher und erst im zweiten Schritt als Werk zu berücksichtigen. Vor allem aber eröffnet sich ein ganz neues Feld der Usability von digitalen Sprachressourcen, wenn man den Marketing-Aspekt mit in den Blick nimmt. Die Leitfrage hier ist: Wie lassen sich Fragende, die über eine Suchmaschine auf eine konkrete Information im Artikel gelenkt wurden, für weitere Informationsangebote der digitalen Sprachressource gewinnen? Um diese neuen Anforderungen an die Usability geht es im Folgenden.

<sup>40</sup> [www.dwds.de/wb](http://www.dwds.de/wb) beispielsweise bei der Angabe „Wortprofil“.

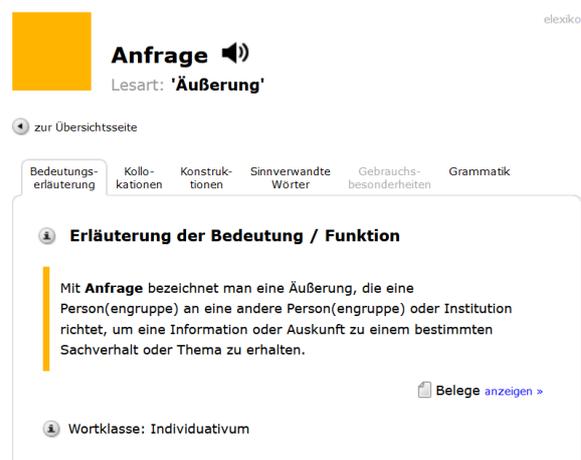
### 5.3.1 Neue Anforderungen an die Usability

Der vorherige Abschnitt hat in den Blick gerückt, dass in Nachschlagewerken wesentlich mehr Informationen relevant sind als nur die inhaltsbezogenen. Die Kompetenz, ein Werk anhand der Haptik, Optik und einiger Informationen auf dem Bucheinband hinsichtlich der Qualität und Verlässlichkeit zu bewerten, ist im Digitalen nicht wirksam. Neben der Ergebnisdarstellung einer Suchmaschine bekommt damit jede einzelne Artikelseite eine zentrale Rolle, die meines Wissens bislang nicht erforscht ist. Aus der Analyse verschiedener Start- und Artikelseiten digitaler Sprachressourcen kristallisieren sich zwei charakteristische Darstellungsvarianten heraus: Das ist zum einen die Wiki-Struktur, wie man sie bei *wikipedia.de*, *wiktionary.de* oder auch *variantengrammatik.net* findet. Zum anderen gibt es Anordnungen von Informationskategorien in einer frei programmierten Oberfläche, die man von *dwds.de*, *duden.de* und beispielsweise von den Sprachressourcen auf *owid.de* kennt.



Vielleicht kann man davon sprechen, dass diese beiden Grundstrukturen einen identifikatorischen Charakter für Sprachressourcen haben, wobei die eher tabellarische und mit kürzeren Texten versehene Struktur näher am klassischen Wörterbucheintrag ist und die Wikis eine gewisse Ähnlichkeit mit Blogs aufweisen. Entsprechend sind Wikis tendenziell auch textlastiger, wodurch das gezielte Auffinden von Informationen gegebenenfalls erschwert wird. Ebenfalls eine geschickte und für Computernutzende bekannte Darstellung findet man auf *owid.de*, wo unterschiedliche Kategorien in Form von Reitern (bzw. Tabs) dargestellt sind.

Diese Darstellung mit Reitern wie in einem Ordersystem ist insofern interessant, als sie auch bildhaft aufdeckt, dass es sich bei den digitalen Sprachressourcen zumindest technisch nicht um Sammlungen von Wörterbuchtexten handelt, sondern um Datenbanken. Den Darstellungsmöglichkeiten sind sehr wenige Grenzen gesetzt, insbesondere in Bezug auf multimediale Inhalte (bspw. die Audiodatei zur Aussprache im obigen *owid.de*-Eintrag „Anfrage“). Inwiefern die Nutzenden das Angebot wahrnehmen und letztlich den erwünschten Informationsgewinn aus einer digitalen Sprachressource ziehen, muss weiter getestet und evaluiert werden.



### 5.3.2 Nutzende auf der Seite halten

Dabei kommt die Frage auf, wie sich Erstbesuchende dafür begeistern lassen, das Angebot einer digitalen Sprachressource nicht nur für den reinen Informationsgewinn zu nutzen. Hier wäre es hilfreich,

die Interessen der Besuchenden einer digitalen Sprachressource besser zu kennen. Die Analyse des Nutzungsverhaltens der Nutzenden der *Varietengrammatik* legt nahe, dass dies sehr unterschiedlich gelagerte Interessen sind.

### 5.3.2.1 Das Nutzungsverhalten der Varietengrammatik

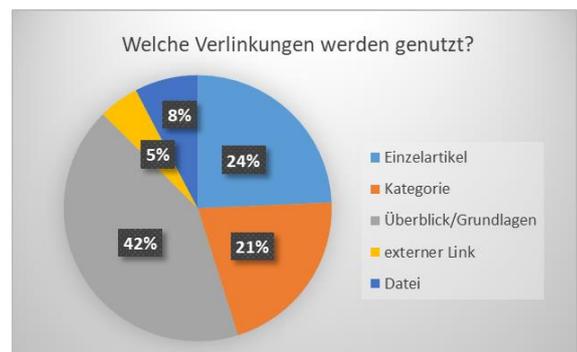
Generell besuchten im Untersuchungszeitraum von einem halben Jahr 8.020 Personen jeweils eine einzige Artikelseite für mindestens 2 Sekunden. Über diese Personen weiß man so gut wie nichts,



lediglich an der Besuchsdauer könnte man vorsichtig abschätzen, wie intensiv sie einen Artikel wahrgenommen haben. Solche Vermutungen hätten allerdings einen hohen spekulativen Anteil. Das heißt, Anhaltspunkte darüber, wie man diese Personen für eine intensivere Beschäftigung mit der Sprachressource gewinnen kann, muss man bei denen suchen, die zwei oder mehr Seiten aufrufen. Dies wurde im DL-Projekt genauer analysiert. Dabei zeigte sich: Ausgehend

von der ersten besuchten Artikelseite<sup>41</sup> klickten 20 % der Besuchenden auf einen Umtext, 27 % suchten nach einem anderen Artikel und 53 % folgten einer Verlinkung. Gewertet werden dabei nur die Aktionen, bei welchen die zweite besuchte Seite für mindestens 5 Sekunden aufgerufen wurde oder zu einer weiteren Aktion geführt hat.

27 % haben nach dem ersten Artikel nach weiteren Artikeln gesucht. 70 % dieser Gruppe benutzten dafür die Suchfunktion, 30 % wählten die Übersicht „Alle Artikel“ zur Auswahl eines zweiten Artikels. Von der mit 53 % größten Gruppe, die einem Link folgten, wählten die meisten (42 %) einen Überblicks- oder Grundlagenartikel. 21 % schauten sich die Kategorie an, in welche das Lemma eingeordnet ist, und 24% klickten auf einen Einzelartikel.



Auf verschiedenen Wegen finden die Personen also zu einer zweiten Informationsseite (also einer Seite aus den Umtexten oder mit einem Artikel). Nach dem vorherigen Schema wurde auch untersucht, welche Aktion ausgehend von dieser zweiten Informationsseite erfolgte. Eine vorsichtige Klassifizierung der Nutzenden auf der Basis ihres Klickverhaltens könnte man folgendermaßen vornehmen.<sup>42</sup> In diese Klassifizierung sind noch weitere Werte (wie die Besuchsdauer einer Seite) eingeflossen:

Nutzende mit ...	phänomenbezogenem Interesse	themenbezogenem Interesse	werkbezogenem Interesse
sind erkennbar an ...	- tendenziell weniger (1 bis 3) Seitenaufrufen	- drei oder mehr Seitenaufrufen - systematischen, eingrenzenden oder	- der Navigation über die Seite „alle Artikel“ - der Wahl von Artikeln im vorderen

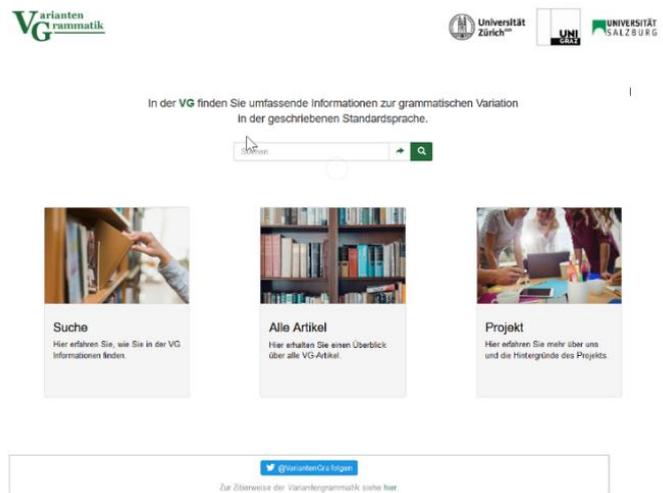
<sup>41</sup> Der Besuch der ersten Artikelseite muss nicht der ersten Aktion entsprechen, beispielsweise wenn jemand über Startseite > Suchfunktion auf den Artikel gekommen ist. Dennoch ist es sinnvoll, vom ersten Besuch einer Artikelseite auszugehen, weil anzunehmen ist, dass damit auch das erste Informationsbedürfnis bedient wurde.

<sup>42</sup> Es wurden die Aktionen von Nutzenden mit 2, 3, 5, 9 und 13 Aktionen im Laufe eines halben Jahres ausgewertet. Die Analyse beruht auf einer qualitativen Betrachtung aller Bewegungen, wobei die einzelnen besuchten Seiten (bspw. Suchseite, Startseite, Artikelseite, Kategorie-Seite, etc.) unterschiedlich eingefärbt waren. Es muss aber auch betont werden, dass der Großteil der Aktionen keinen nachvollziehbaren Mustern folgt, sodass die vorgeschlagene Klassifizierung lediglich für eine Thesenbildung herangezogen werden konnte. Für diese würde es sich lohnen, sie an anderen Nachschlagewerken zu überprüfen und weiterzuentwickeln.

	- einem klaren Fokus auf einen oder zwei Artikel	erweiternden Suchanfragen - der Nutzung von Kategorien und Überblicksartikeln für die Suche	Abschnitt des Alphabets - der Suche nach bekannten regionalen Varianten oder Dialektbeispielen
kommen von ...	- Suchmaschinen	- direkt	- Websites
suchen überwiegend nach ...	- Wortformen - Schreibvarianten	- grammatischen Termini - Konstruktionen	- typischen Beispielen
besuchen anschließend ...	- verlinkte Artikel - Seiten mit Umtexten	- übergeordnete Artikel - den Download für pdf-Dateien - die Suchfunktion	- verschiedene Seiten nach ähnlichen Auswahlmustern wie zuvor
verweilen ...	- mit abnehmender Dauer pro Seite	- mit einer mit dem Artikelumfang korrelierenden Dauer	- un stet, eher kürzer

### 5.3.2.2 Qualitätsprüfung

Man kann davon ausgehen, dass einige Personen nach dem Besuch eines Artikels – den sie über eine Suchmaschine gefunden haben – die Startseite oder die „Über uns“-Seite aufrufen, um sich über Hintergründe zu der Informationsquelle kundig zu machen. Das kann nur aus Interesse sein, es kann aber der Wunsch dahinterstecken, die Wertigkeit und Verlässlichkeit der Informationen zu prüfen bzw. die Vertrauenswürdigkeit der Quelle. Diese Personen versuchen, ihr Wissen über das Werk zu ergänzen, nachdem sie die nötigen Informationen zu ihrer Fragestellung erhalten haben – oder auch nicht erhalten haben. Es zeigt sich, dass die Startseite der *Variantengrammatik* hierfür wenig Informationen direkt bereithält, allerdings beispielsweise einen Link zu Informationen über das Projekt anbietet. Ebenso sind die Logos der beteiligten Universität markant platziert.



### 5.3.2.3 Verlinkung und Vernetzung

Die *Variantengrammatik* bietet am Ende jedes Artikels unter der Überschrift „Siehe auch“ ein systematisches und kategorisiertes Angebot an Verlinkungen. Wie man an dem Screenshot sieht, gibt es dabei übergeordnete Artikel (Grundlagenartikel, Überblicksartikel) und nebengeordnete Einzelartikel mit vergleichbaren Phänomenen. Außerdem wird dargestellt, zu welchen Kategorien der aufgerufene Artikel zählt. Oftmals wird auch

#### Siehe auch

- Grundlagenartikel: [Wortbildung](#) (oder direkt Unterkapitel [Verbderivation](#))
- Überblicksartikel: [Verben auf -isier\(en\)](#)
- Einzelartikel: [schulblädieren](#)

Kategorien: [Verben auf -isier\(en\)](#) | [Einzelartikel](#)

im Text selbst durch Verlinkungen auf thematisch relevante Seiten hingewiesen (vgl. den nächsten Screenshot). Selten gibt es auch Verlinkungen auf externe Seiten.

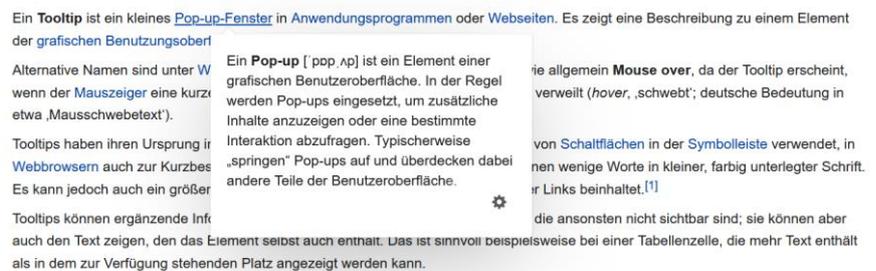
Vom Substantiv *Termin* können die Verben *terminieren* und *terminisieren* abgeleitet werden. Die beiden Varianten werden ohne Bedeutungsunterschied nebeneinander gebraucht (vgl. auch [Verben auf -isier\(en\)](#)). Das Verb *terminieren* ist ÜBERALL

Die Log-File-Betrachtungen zeigten, dass 53 % der über den ersten Artikelbesuch hinausgehenden Aktionen auf der Website durch solche Verlinkungen motiviert sind. Das spricht dafür, Verlinkungen prominenter, attraktiver und gegebenenfalls häufiger in die Artikel einzubauen. Die Position am Artikelende entgeht leicht der Aufmerksamkeit, worauf etwa das folgende Verhalten hinweisen könnte: Eine Person hat nach der 59 Sekunden dauernden Ansicht des Artikels *terminieren* / *terminisieren* die Suchanfrage „fundieren“ und „amortisieren“ durchgeführt. Offensichtlich lag das Interesse dieser Person also auf Verben mit der Endung *-ieren* oder *-isieren*. Dazu wäre ein entsprechender Überblicksartikel am Ende und im Text (vgl. die beiden Screenshots oben) als auch das weitere Beispiel *schubladisieren* verlinkt gewesen. Offensichtlich wurde dieses Angebot allerdings nicht genutzt. An den weiteren Suchanfragen nach *fundieren* und *amortisieren*, die als zweite Aktion beim Besuch ausgeführt wurden, zeigt sich, dass die prinzipiell vorhandenen Verlinkungen ebenfalls interessant gewesen wären.

Entscheidender ist allerdings die Feststellung, dass die Besuchenden thematisch relevante Verlinkungen nutzen (42 % besuchen einen passenden Grundlagen- oder Überblicksartikel). Das zeigt, dass sie also mehr über das für sie interessante Phänomen erfahren wollen. Eine Verbesserung des Nutzungserlebnisses müsste also darauf abzielen, die Informationen in den einzelnen Artikeln untereinander intensiver, sichtbarer und nachvollziehbar zu vernetzen. Das digitale Potential des traditionellen lexikographischen Verweisens und Verknüpfens ist sicher noch nicht ausgereizt. Visuelle

Varianten wie die in Kap. 5.1.2 erwähnte Wortwolke sollten ebenso evaluiert werden wie graphische Hervorhebungen von Textlinks. Auch sogenannte „Tooltips“, kleine Info-Fenster, die sich öffnen, wenn man den Mauszeiger über einen Link führt, sollte man evaluieren.

## Tooltip



Ein **Tooltip** ist ein kleines [Pop-up-Fenster](#) in [Anwendungsprogrammen](#) oder [Webseiten](#). Es zeigt eine Beschreibung zu einem Element der [grafischen Benutzeroberfläche](#).

Alternative Namen sind unter [W](#) wenn der [Mauszeiger](#) eine kurzzeitige [Mausschwebetext](#)).

Tooltips haben ihren Ursprung in [Webbrowsern](#) auch zur [Kurzbes](#). Es kann jedoch auch ein größerer

Tooltips können ergänzende Info auch den Text zeigen, den das Element selbst auch enthält. Das ist sinnvoll beispielsweise bei einer Tabellenzelle, die mehr Text enthält als in dem zur Verfügung stehenden Platz angezeigt werden kann.

Ein **Pop-up** [ˈpɒp.ʊp] ist ein Element einer grafischen Benutzeroberfläche. In der Regel werden Pop-ups eingesetzt, um zusätzliche Inhalte anzuzeigen oder eine bestimmte Interaktion abzufragen. Typischerweise „springen“ Pop-ups auf und überdecken dabei andere Teile der Benutzeroberfläche.

ie allgemein **Mouse over**, da der Tooltip erscheint, verweilt (*hover*, ‚schwebt‘; deutsche Bedeutung in [von Schaltflächen](#) in der [Symboleiste](#) verwendet, in denen wenige Worte in kleiner, farbig unterlegter Schrift für Links beinhaltet.<sup>[1]</sup> die ansonsten nicht sichtbar sind; sie können aber

Solche Verlinkungen zu weiterführenden und die gesuchte Information ergänzenden Informationen führen dazu, dass die Besuchenden sich länger mit dem Angebot des Nachschlagewerks auseinandersetzen. Dadurch steigt der Informations- und Erkenntnisgewinn und – so die Annahme – auch die Zufriedenheit des Besuchs. Verlinkungen lassen sich prinzipiell automatisiert generieren, wobei natürlich ein im redaktionellen Prozess gesetzter, sehr spezifisch zu einer passenden Information führender Link einen Mehrwert besitzt. Daher bietet es sich gerade für neue Sprachressourcen an, die Vernetzungsoptionen im Vorfeld strategisch zu planen. Für die Suchmaschinenoptimierung sind auch Links auf externe Seiten und Backlinks von diesen Seiten wichtig (vgl. Kap. 5.1.2.6). So sind beispielsweise in der *Varietengrammatik* grammatische Fachtermini mit einer Quelle verlinkt, in welcher der entsprechende Ausdruck verlässlich und nachvollziehbar erläutert wird.<sup>43</sup> Wesentlich schwieriger dürfte es angesichts des für die *Varietengrammatik* beobachteten Nutzungsverhaltens

<sup>43</sup> Bspw. <https://grammis.ids-mannheim.de/>

werden, andere, thematisch nicht anschließende Artikel in den Fokus von Besuchenden zu rücken. Dafür, dass es ein signifikantes Potential an Nutzenden gibt, die interessiert und regelmäßig durch unterschiedliche Artikel einer digitalen Sprachressource schmökern, finden sich in den Log-Files der *Varietengrammatik* keine Anhaltspunkte.<sup>44</sup>

### 5.3.2.4 Informationen vernetzen, nicht Artikel

Auch bei der Vernetzung von digitalen Sprachressourcen bringt der Fokus auf Informationen (anstelle von Artikeln) Vorteile:

- Die Vernetzung von Informationen, also ganzen Abschnitten innerhalb der Artikel, ist zielführender und präziser als die Vernetzung von Artikeln, die eine Vielzahl von Informationen – auch irrelevanten – bündeln.
- Ein Netzwerk von Informationen ist von jedem Knotenpunkt aus zugänglich, wodurch die Abhängigkeit von klassischen Zugriffsstrukturen über Lemmata an Bedeutung verliert.
- Nutzende suchen keine Artikel, sondern Informationen. Artikel sind ein aus dem klassischen Wörterbuch herrührendes Mittel zum Zweck der thematischen Bündelung von Inhalten (Stichwort: Zugriffsstruktur). Die *Google*-Suchergebnisse zeigen die Möglichkeiten der Digitalisierung auf: Informationen (bspw. eine passende Bedeutung oder eine erklärende Grafik) werden ohne den Umweg über die Zugriffsstruktur adressiert und angezeigt, und sie haben das Potential, die Fragestellungen ohne den Artikelkontext zu klären.

In diese Richtung weiterzudenken würde bedeuten, dass der ersten Information nicht der ganze Artikel folgt, sondern eine Auswahl an vernetzten Informationen. Das scheint gerade für die in der Benutzungsforschung als zu ausführlich kritisierten Artikel interessant (vgl. Tarp 2015, 18; Gouws/Tarp 2017). Statt solche Artikel unter dem Verlust von Informationen zu kürzen, könnten durch die Fokussierung auf Informationsabschnitte genau die Passagen aufgezeigt werden, die für die suchende Person relevant sind. Davon ausgehend könnten dann weitere Informationen angeboten und erschlossen werden. Darin sehe ich das Potential für die digitale Weiterentwicklung von Sprachressourcen, die inhaltlich gut ausgearbeitet sind, bisher aber den wachsenden Ansprüchen an Häppchenkost und am schnellen Informationsgewinn in der bestehenden Form nicht gerecht werden. Eine solche Darstellung gelingt nicht mit einer einmaligen Suchanfrage, sie braucht vielmehr eine beständige Interaktion der Nutzenden mit dem Online-Angebot. Die Sprachressource bietet den Nutzenden fortwährend passende Informationsangebote zur Auswahl. Diese Informationspräsentation wird ständig verbessert durch einen Algorithmus, der beispielsweise auf der Basis der von Nutzenden getroffenen Auswahl lernt.



Digitale Sprachressourcen haben alles, was es dafür braucht. Sie haben die exklusiven Inhalte – die sich (noch) nicht automatisch generieren lassen – und sie sammeln im laufenden Betrieb Nutzendendaten<sup>45</sup>,

<sup>44</sup> Zwar gibt es die Besuchenden, die mehrere Artikel konsultieren (Gruppe 3 in der Tabelle auf S. 51), allerdings gehen wir davon aus, dass das überwiegend Erstbesuche und Testbesuche waren, die auch keine weiteren Besuche mit der gleichen IP-Adresse zur Folge hatten.

<sup>45</sup> Ich denke hier nicht an personenbezogene Daten, sondern an Verhaltensmuster, die personenunabhängig ausgewertet werden können.

über welche sich Zusammenhänge zwischen den nachgeschlagenen Informationen entdecken lassen. *Amazon* beispielsweise schlägt auf der Basis von Nutzendendaten Produkte vor, die andere Personen zu einem Ausgangsprodukt dazugekauft oder zumindest angesehen haben. Was spricht dagegen, in digitalen Sprachressourcen Informationsangebote, die gerne genutzt wurden, auch anderen zu empfehlen?

Im nächsten Kapitel richtet sich der Fokus auf die Zukunft von digitalen Sprachressourcen. Bislang orientierten wir uns an der *Varietengrammatik*, also an einem in der Tradition der digitalen Lexikographie stehenden Nachschlagewerk. Doch neben diesen lexikographischen Angeboten gibt es viele weitere Informationsseiten im Internet, die das Recherchieren und Informieren prägen. Das abschließende Kapitel soll nun zeigen, welche im Internet schon etablierten Konzepte und Features die künftige Gestaltung von digitalen Sprachressourcen prägen könnten.

## 6 Die Zukunft digitaler Sprachressourcen

### 6.1 Usability aus nicht-lexikographischen Umgebungen importieren

Digitale Sprachressourcen haben ein großes Potential für neue Features. Am IDS-Mannheim wurden solche Innovationen immer wieder getestet. Caroline Müller-Spitzer (2016a, 323) hält dazu fest: „Unsere Daten zeigen damit, dass es lohnend ist, innovative Features in Internetwörterbücher zu integrieren, dass sich die Anbieter solcher Wörterbücher aber auch gleichzeitig klar darüber sein müssen, dass die Nutzerinnen nur nach und nach von dem Nutzen dieser neuen Features überzeugt werden können.“ Mit der Innovation einher geht quasi das Anlernen dieser Features. So erwähnt Müller-Spitzer (ebd.), dass der Nutzen neuer Features höher bewertet wird, wenn die Nutzenden vorher über die Neuerungen informiert wurden. Das klingt nach einem schwierigen Unterfangen, bedingt dadurch, dass die Möglichkeiten der Einwirkung auf das Publikum begrenzt sind, und bedingt durch die Tatsache, dass viele Besuchende gar nicht die Umtexte wahrnehmen, in denen man sie über solche Neuerungen informieren könnte. Unsere Überlegung war deshalb: **Digitale Sprachressourcen können Konzepte und Features adaptieren, deren Funktionsweise und Nützlichkeit durch andere Informationsangebote im Internet schon bekannt und akzeptiert sind.** Diese Überlegung resultiert u.a. aus der Tatsache, dass man über einen Großteil der Nutzenden eines online verfügbaren wenig weiß. Das Einzige, das man annehmen kann, ist, dass diese Personen mit der Nutzung der üblichen Angebote im Internet vertraut sind.

Im DL-Projekt sollte deshalb verglichen werden, wie sich das Informieren in digitalen Sprachressourcen einerseits und das Online-Informieren andererseits zueinander verhalten. Hierfür wurde die Präsenz von verschiedenen lexikographischen Internetangeboten (jeweils Startseite und eine Artikelseite) untersucht. Ziel war herauszufinden, wie auf diesen Seiten Informationen zugänglich und verfügbar gemacht werden. Außerdem wurden ausgewählte Websites aus der Liste der 20 am meisten aufgerufenen Websites der Schweiz analysiert.<sup>46</sup> Dabei interessierte, wie man jenseits digitaler Sprachressourcen nach Produkten oder Informationen sucht und wie die Suchergebnisse jeweils dargestellt werden. An welche Recherche- und Auswahlstrategien, Entscheidungshilfen und

---

<sup>46</sup> Die Liste wurde abgerufen von Wikipedia.ch am 06.02.2020. *Booking.com* ist bei einem aktuellen Aufruf der Liste als Konsequenz aus der Corona-Pandemie nicht mehr in der Liste zu finden; [https://de.wikipedia.org/wiki/Liste\\_der\\_meistaufgerufenen\\_Websites#In\\_der\\_Schweiz](https://de.wikipedia.org/wiki/Liste_der_meistaufgerufenen_Websites#In_der_Schweiz).

Ergebnispräsentationen sind Internetnutzende bereits gewöhnt? Daraus leitet sich die These ab, dass genau diese technischen Lösungen und Strategien, die generell beim Informieren und Auswählen genutzt werden, auch Einzug in die Lexikographie finden sollten. Usability würde in dieser Hinsicht bedeuten, dass bestehende Gewohnheiten von der Wissenschaft adaptiert werden. Dahinter steht die folgende Überlegung: Websites werden heute mit bestimmten Bausteinen konstruiert, weshalb es naheliegend erscheint, dass diese auch für die Lexikographie genutzt werden (könnten). Derzeit sind Wiki-Systeme und selbst programmierte Oberflächen üblich. Doch sind auch andere Baukästen (bspw. solche für Shop-Systeme oder für Bilder- bzw. Video-Portale) denkbar?

### **INFO: Analyse zum Online-Recherchieren und Informieren – Websites und Sprachressourcen im Vergleich**

Für das Recherchieren und Informieren im Internet sind digitale Grundkompetenzen nötig. Gleichzeitig werden diese Kompetenzen durch die Nutzung des Internets geprägt. Es liegt also nahe, dass man bei der Konzeption, Struktur und Darstellung digitaler Sprachressourcen auf solche Grundkompetenzen zurückgreift.

#### AUFBAU UND DURCHFÜHRUNG DER UNTERSUCHUNG

Aus der Liste der meistbesuchten Websites wurden Seiten unterschiedlicher Genres (bspw. eine Zeitung, ein Webshop, eine Suchmaschine, etc.) ausgewählt. Die Seiten sind:

- *Google* <https://www.google.ch> (als Suchmaschine)
- *YouTube* <https://www.youtube.com> (als Videoportal)
- *Facebook* <https://www.facebook.com> (als Social-Media-Seite)
- *20min.ch* <https://www.20min.ch> (als meistbesuchte Zeitungsseite der Schweiz)
- *Amazon* <https://www.amazon.de> (als Marktplatz)
- *Zalando* [www.zalando.ch](http://www.zalando.ch) (als Onlineshop)
- *Booking.com* [www.booking.com](http://www.booking.com) (als Reiseportal)
- *Ricardo* [www.ricardo.ch](http://www.ricardo.ch) (als Auktionsportal)

Weiter wurden die folgenden Sprachressourcen ausgewählt:

- *DWB „Grimm online“* [http://woerterbuchnetz.de/cgi-bin/WBNetz/wbgui\\_py?sigle=DWB](http://woerterbuchnetz.de/cgi-bin/WBNetz/wbgui_py?sigle=DWB)  
(als Beispiel aus dem Wörterbuchportal mit digitalisierten Printwerken)
- *Duden Wörterbuch* [www.duden.de/woerterbuch](http://www.duden.de/woerterbuch)
- *elexiko – ein Online-Wörterbuch zur deutschen Gegenwartssprache*  
[www.owid.de/docs/elex/start.jsp](http://www.owid.de/docs/elex/start.jsp) (als Stellvertreter für das OWID-Portal)
- *DWDS-Portal* [www.dwds.de](http://www.dwds.de) (als Portal)
- *Oxford English Dictionary* [www.oed.com](http://www.oed.com)
- *Varietengrammatik des Deutschen* [www.variantengrammatik.net](http://www.variantengrammatik.net)
- *Wikipedia* [www.wikipedia.org](http://www.wikipedia.org) (im engeren Sinn keine Sprachressource, sondern ein Lexikon)
- *Wiktionary* [www.wiktionary.org](http://www.wiktionary.org)
- *Wortschatz „Uni Leipzig“* <https://wortschatz.uni-leipzig.de/de>  
(als lexikographisch nutzbares Korpus-Angebot)

Bei allen Seiten wurde, sofern möglich, eine festgelegte Abfolge an Navigationsschritten durchgeführt und die einzelnen Schritte mit Screenshots dokumentiert. Als Screenshot wurde jeweils die Ansicht bei 100% auf einem 17“-Bildschirm aufgenommen und dann (sofern dies durch Zoomen möglich war) die gesamte Seite.

1. Schritt: Startseite aufrufen (zwei Screenshots: 100%-Ansicht und ganze Seite).
2. Schritt: Suchanfrage „Päckch“<sup>47</sup> eintippen (ein Screenshot).
3. Schritt (bei Bedarf): Auswahl der Zielseite aus einer Ergebnisauflistung (ein Screenshot)
4. Schritt: Mit „Päckchen“ gefundene Seite öffnen (zwei Screenshots: 100%-Ansicht und ganze Seite)

## ANALYSE

Die einzelnen Seiten wurden schließlich nach den folgenden Kriterien qualitativ analysiert:

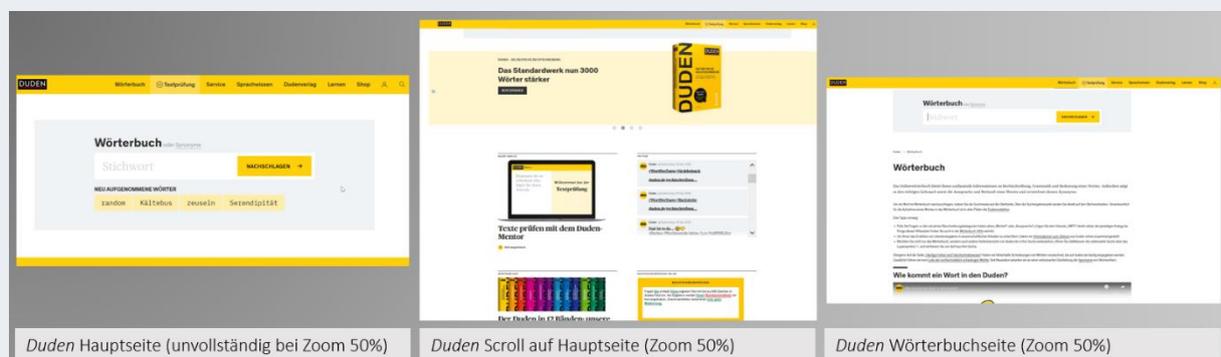
- Design der Startseite
- Struktur des Suchfeldes
- Präsentation der Suchergebnisse
- Informationsfülle und Navigation auf einer Seite
- Struktur und Darstellung der Informationen

Der Sucherfolg und die Wertigkeit des Suchergebnisses spielten in der Analyse keine Rolle, nur graphische, strukturelle und funktionale Aspekte wurden berücksichtigt.

## ERKENNTNISSE ZU LEXIKOGRAPHISCHEN SEITEN

### *Design der Startseite*

Die analysierten digitalen Sprachressourcen zeichnen sich durch eine dezente Gestaltung aus, meist aus schwarzem Text und einer Akzentfarbe. Beim Angebot des *Duden-Verlags* unterscheiden sich die Seiten *duden.de* als moderne Seite zum Scrollen mit animierter Werbung und für Verlagsprodukte und *duden.de/woerterbuch*, die ebenfalls ebenfalls dezent gehalten ist. Bei allen lexikographischen Angeboten steht der einfache Zugang zu einem Suchfeld oder zu einer Übersicht der Informationsangebote im Vordergrund, teilweise wird mit illustrierenden Bildern gearbeitet (bspw. *ellexiko*).



Die Startseiten *dwds.de* und *wortschatz.uni-leipzig.de* unterscheiden sich von den anderen Seiten, weil sie als Portale für unterschiedliche Angebote fungieren (bspw. Korpuszugang, aktuelle Beiträge oder ein Wortspiel). Die Einstiegsseite in die *DWDS*-Wörterbücher hingegen ist sehr textlastig; sie wurde nicht

<sup>47</sup> Um das Autovervollständigen zu untersuchen, wurden die beiden letzten Buchstaben weggelassen. Im Oxford English Dictionary wurde ersatzweise „parc“ gesucht, bei *Zalando* und *Ricardo* „Pullov“.

als übersichtliche Startseite konzipiert. Die Startseite von *wiktionary.org* fällt durch eine recht eigenständige Gestaltung auf: Hier werden viele Informationsangebote in Kacheln präsentiert (bspw. „Allgemeine Hinweise“, „Wort der Woche“ oder „Aktuelles“). Gänzlich anders gestaltet ist die Startseite des *Oxford English Dictionary*, das nur mit einer Bezahlschranke zugänglich ist. Statt eines Suchfelds findet man auf der Seite, die von der animierten Werbung für eigene Produkte dominiert wird, nur die Buttons „Subscribe“ und „Login“.



### Struktur des Suchfelds

Auffällig ist auf vielen Seiten das meist zentral positionierte Suchfeld. Filtermöglichkeiten oder differenzierte Suchanfragen werden selten offensiv angeboten. Die Anleihe an der *Google*-Optik ist unverkennbar. Bei *lexiko* – als Teil des *OWID*-Portals, das auch ein übergeordnetes Suchfeld anbietet – wird auf der Startseite direkt auf die „erweiterte Suche“ verwiesen, die ihrerseits mit zahlreichen Suchoptionen ausgestattet ist. Tippt man eine Suchanfrage ein, erhält man in der Regel Vorschläge zu vorhandenen Einträgen, die sich direkt durch Anklicken auswählen lassen. Unter dem Aspekt der Sichtbarkeit des Werkes auf der Startseite ist es interessant, was als Hintergrundtext im Suchfeld steht: Die Optionen reichen hier von „Suchen“ über „Suchen im DWDS“ bis zu einer konkreten Vorgabe, was man eingeben soll (bspw. „Stichwort“). Auf *duden.de/woerterbuch* wurde zur Markenbildung der gelbe Button zum Ausführen der Suche mit „Im Duden nachschlagen →“<sup>48</sup> beschriftet, während andere beispielsweise eine Lupe als Symbol wählen.



### Präsentation der Suchergebnisse

In der Regel werden die passenden Suchergebnisse in Listenform angeboten, oftmals mit einer gewissen Schreibweisentoleranz. Die eindeutige Lemmazentrierung wird bei der Ermittlung und Darstellung der Suchergebnisse in vielen Nachschlagewerken sichtbar. Im *DWDS*-Wörterbuch wird die alphabetische Makrostruktur genutzt, um im Falle einer erfolglosen Suche alternative Einträge vorzuschlagen oder um die vorausgehenden und nachfolgenden Artikel zum Suchtext anzugeben. Die *Varietengrammatik*

<sup>48</sup> Diese Formulierung wurde mit einer Überarbeitung der Seite offensichtlich in „Nachschlagen →“ geändert.

bezieht auch den Volltext in die Suche ein. In den Ergebnissen wird dann unterschieden, ob das gesuchte Wort im Inhalt vorkommt oder im Seitentitel (= Lemmaposition) steht.

### Darstellung der Artikelseite(n)

Die Präsentation der Inhalte lässt zwei Muster erkennen. Einerseits finden sich stark strukturierte Seiten mit kleinteiligen Textsegmenten, die an die Strukturen von gedruckten Wörterbüchern erinnern (bspw. *Duden* oder die Ressourcen in *OWID*). Es liegt die Vermutung nahe, dass dies datenbankbasierte Systeme mit einer eigenen Programmierung der Oberfläche sind. Die zweite Darstellungsvariante basiert auf Wiki-Systemen und ist daher in den Strukturierungsmöglichkeiten eingeschränkt. Diese Seiten sind wie ein Text mit Textteilen (bspw. *Varietengrammatik*) oder als Listen (bspw. *Wiktionary*) gegliedert. Tendenziell weisen die kleinteilig organisierten Artikel mehr Verweise und Klappmenüs auf, während längere Wiki-Artikel oftmals eine strukturierende Gliederung voranstellen.



Multimodale Darstellungen sind ebenfalls vorhanden, wie etwa die Aussprache-Beispiele bei *Duden*, die Kartenvisualisierung in der *Varietengrammatik* oder die Visualisierung des diachronen Vorkommens eines Wortes mit einer Verlaufskurve im *DWDS*-Portal. Eine Frage stellt sich bezüglich der Positionierung von visuellen Informationen. Gerade in sehr textlastigen Artikeln stehen visuelle Elemente hervor. So zeigt sich beispielsweise im *OWID*-Portal, dass visuelle Elemente strukturierend sind und Überschriften ersetzen können: Dort werden auf der Bedeutungsebene eines Lexems unterschiedliche Aspekte wie Bedeutungserläuterung, Kollokationen oder Grammatik in Form von Tabs angezeigt.



Das rückt einen weiteren Aspekt in den Fokus: Wie gehen digitale Sprachressourcen mit der Darstellung umfangreicher Inhalte um? Generell finden sich hier – neben der Tab-Darstellungsvariante in *OWID* – drei Ansätze, die auch kombiniert vorkommen können.

- Alle Inhalte sind linear angeordnet auf einer Seite mit einem Inhaltsverzeichnis.
- Alle Inhalte sind linear angeordnet auf einer Seite, wobei umfangreiche Inhalte minimiert werden.
- Die Inhalte sind auf mehrere Seiten verteilt und verlinkt.

Ein leitendes Element dabei ist die Hervorhebung der Artikelstruktur durch das Ausblenden von Inhalten oder die Extraposition eines Inhaltsverzeichnisses, was die Navigation erleichtern dürfte.

## ANALYSE DER NICHT-LEXIKOGRAPHISCHEN WEBSITES

Die meistbesuchten nicht-lexikographischen Angebote im Internet wurden im Projekt vor allem im Hinblick auf die Darstellung von Informationen und hinsichtlich der Unterschiede zu lexikographischen Seiten betrachtet. Es zeigen sich einige typische – und man darf davon ausgehen im Netz etablierte – Darstellungen und Funktionen im Zusammenhang mit der Präsentation von Informationen bzw. Produkten (vgl. dazu die folgenden Kapitel).

## 6.2 Plattform statt Einzelressourcen?

Weiter oben wurde ausgeführt, dass die einzelne, passende Information für die Auffindbarkeit eines Angebots an Bedeutung gewinnt und dass auf diesem Umweg über die Artikelinhalte auch das Werk als solches sichtbar werden kann. Des Weiteren wurde argumentiert, dass die Usability sich stärker an der Information als am lexikographischen Artikel orientieren sollte und dass ein großes Potential in der Vernetzung dieser Informationen besteht. Die unterschiedlichen Sprachressourcen und ihre Daten sind allerdings so heterogen gestaltet, dass eine gegenseitige Vernetzung mit erheblichen Schwierigkeiten verbunden ist. Praktikabel wäre eine großflächige Vernetzung vermutlich nur auf der obersten Strukturebene, dem Artikellemma (vgl. das Angebot *woerterbuch-portal.de* oder *www.woerterbuchnetz.de*). Dort sind die Artikel aus verschiedenen Quellen über das Lemma vernetzt, eine Vernetzung von Informationen bleibt aus. Es sind hier zwar viele Artikel auffindbar, aber die Nutzenden müssen die Bereitschaft mitbringen, aus den teilweise umfangreichen Artikeln die passenden Informationen herauszusuchen.

*dwds.de* zeigt, wie sich mit Informationen aus unterschiedlichen einzelnen Sprachressourcen ein übergeordneter Artikel mit erheblichem Mehrwert generieren lässt. Würde man dieses Konzept weiter ausbauen, bekäme man eine Plattform, über welche die Informationen aus vielen Sprachressourcen zugänglich sind.<sup>49</sup> Die aktuelle Anzahl der Quellen – und damit die Informationsvielfalt – ist allerdings recht gering aufgrund des Aufwandes, den ihre Erschließung bedeutet. Umfangreicher sind die Quellen

beispielsweise bei den Portalen *owid.de* oder dem übergreifenden Wörterbuchportal *woerterbuch-portal.de*, das auf Belege und Stichwörter aus zahlreichen digitalen Sprachressourcen verlinkt. In beiden Portalen sind auch digitalisierte Printwörterbücher zugänglich (bspw. im *Wörterbuch-Portal* das Grimm'sche „Deutsche Wörterbuch“ und bei *owid.de* das „Deutsche Fremdwörterbuch“).



Der Artikel „Päckchen“ im Deutschen Wörterbuch online mit Verweisen auf andere Quellen.

Es stellt sich die Frage, wie sich eine Plattform für lexikographische Informationen von bestehenden Portal-Angeboten unterscheiden müsste. Dazu sei nochmals die These aufgegriffen, dass der Lexikographie ein disruptiver Prozess drohen könnte (vgl. Kap. 2.1). Möglicherweise geht diese Gefahr

<sup>49</sup> Weitere, internationale Beispiele für solche Plattformen finden sich bei Boelhouwer et al. 2018.

nicht von anderen, vielleicht innovativeren, lexikographischen Projekten aus, auch nicht von *Wikipedia*, sondern von ganz anderer Seite: Unberechenbar sind Akteurinnen und Akteure, die selbst keine Informationen erzeugen, aber mit fremden Informationen arbeiten: *Google* beispielsweise bedient das Interesse seiner Nutzenden damit, dass es Informationen aus diversen Wissensressourcen anbietet. *Google* selbst besitzt diese Daten und Informationen nicht und hat auch nicht zu ihrer Erarbeitung beigetragen. Auch in unseren Untersuchungen zeigte sich, dass Sprachressourcen wesentlich von der Suchmaschine abhängig sind und dass die Suchmaschine die Nutzung und den Nutzendenkreis verändert (vgl. Kap. 3.2). Sich diesem Zugang zu verschließen, ist keine Handlungsoption, im Gegenteil: Die Optimierung der Informationsangebote für die Algorithmen der Suchmaschinen scheint unumgänglich.

Wäre es nicht erstrebenswert, die Abhängigkeit von Suchmaschinen bei der lexikographischen Informationsrecherche zu minimieren? Die Untersuchung der meistbesuchten Websites zeigt, dass das Informieren ein wesentlicher Aspekt des Internets ist und es keine zentralen Portale für Sprachinformationen gibt. Der Wert der Lexikographie ist, dass sie Informationen anbietet (vgl. Kap. 2.5) und damit genau eine Stärke des Internets bedient. Es ist wichtig, dass diese Informationen auffindbar, sichtbar und nutzbar sind, und langfristig wird dieses Ziel nur zu erreichen sein, wenn sie auch vernetzt und zentral präsentiert werden. Eine solch zentrale Plattform für Sprachinformationen kann im kleineren Stil genau das Prinzip *Google* oder *Amazon* kopieren: Digitale Sprachressourcen könnten dem Open-Access-Konzept folgend dieser Plattform kostenlos ihre Informationen bereitstellen. Möglicherweise würden die Ressourcenbetreibenden sogar für die Publikation ihrer Daten auf einer erfolgreichen Plattform zahlen. Der Vorteil ist, dass die auf der Plattform verfügbare Menge an Sprachinformationen so groß ist, dass fast jede und jeder Suchende fündig wird. Die Folge ist eine hohe User-Experience: Die Menschen kommen wieder und nehmen das Angebot an Sprachressourcen jenseits der Plattform kaum wahr. All die Daten, die man für eine solche Plattform benötigen würde, sind schon jetzt frei im Internet zugänglich. Die Frage ist eher, wer die Plattform umsetzen wird.<sup>50</sup>

Ein solches Angebot könnte die zentrale Adresse werden, die Personen besuchen, wenn sie eine Frage zum Thema Sprache haben (so wie *Wikipedia* zum zentralen Lexikon wurde). Damit die Personen zufrieden sind, müssen sie eine passende Antwort finden. Dafür braucht es gute Informationen aus vielfältigen Quellen. Je mehr Informationen angeboten werden, desto eindrücklicher ist das Nutzungserlebnis. Aufgrund der Vielfalt der relevanten Informationen zu einer Suchanfrage kann die Plattform gegenüber den einzelnen Sprachressourcen-Websites ihre Attraktivität durch Mehrwerte erhöhen (bspw. ein Ranking oder eine Nutzendenbewertung von Informationen). Wenn die Zahl der Nutzenden auf diese Weise steigt, steigt auch die Attraktivität für weitere Anbieter, ihre Informationen auf dieser Plattform verfügbar zu machen.

Das skizzierte Konzept ist nicht neu. Es ist eine Umformulierung der Idee, mit der *Amazon* begann, das Internet zu erobern. Die Idee lautete: Ein Marktplatz mit einer hohen Vielfalt an eigenen und externen Angeboten steigert die Attraktivität für Kaufwillige und Anbietende gleichermaßen.<sup>51</sup> Vor allem aber ist an dem Marktplatzkonzept wichtig, dass *Amazon* die Infrastruktur für die Angebote zur Verfügung stellt, das Einstellen von Warenangeboten fällt den Händlern und Händlerinnen zu. Damit unterscheidet sich das Konzept von einem Online-Shop, in dem nur die Betreibenden etwas einstellen können. Diese

---

<sup>50</sup> Der *Google*-Konzern *Alphabet* hat mit dem Angebot *Google Books* gezeigt, dass er frei verfügbare Publikationen strategisch zu nutzen weiß. Es ist nicht auszuschließen, dass die in lexikographischen Quellen gut aufbereiteten Daten ein ebensolches Interesse wecken könnten.

<sup>51</sup> <https://t3n.de/news/amazon-geschäftsmodell-574526/>, 05.10.2020.

Unterscheidung lässt sich auch auf die Sprachinformationsplattform übertragen, und sie zeigt eine mögliche Zukunftsperspektive: Nicht mehr die Plattformbetreibenden sind für die Integration der Inhalte verantwortlich, sondern die Urheberinnen bzw. Urheber oder die Verwaltenden der Informationen. Das minimiert den Aufwand der Plattformbetreibenden, wenn neue Informationen integriert werden sollen. Das Informationswachstum ist damit nicht mehr von deren Kapazitäten abhängig. Vielmehr können sich die Sprachressourcen-Betreibenden jederzeit im Portal registrieren und ihre Informationen (automatisiert) einstellen. Ein solches Marktplatzkonzept wurde in Bezug auf (lexikographische) Informationen bisher noch nicht umgesetzt. Unter den meistbesuchten Websites allerdings finden sich neben *Google* und *Amazon* weitere Angebote, die nach dem skizzierten Prinzip „eigenes System, fremde Daten“ funktionieren: *ricardo.ch*, dem nach eigenen Angaben größten Online-Marktplatz der Schweiz, *booking.com* oder *YouTube* beispielsweise. Das zeigt, dass die Erfahrung mit solchen Konzepten vorhanden ist. Und es zeigt auch, wo und wie Internetnutzende in Bezug auf das Informieren und Recherchieren geprägt werden.

## 6.3 Digitale Suche, Filter und Ranking

### 6.3.1 Digitale Suche

In diesem Kapitel wird ausgeführt, welche Features das Recherchieren und Auswählen von Sprachinformationen künftig begleiten könnten:

Wie die Seite *booking.com* zeigt, muss ein Suchfeld keineswegs nur eine Eingabemöglichkeit anbieten,



wenn mehr als ein Parameter als für die Suche relevant und sinnvoll erscheinen. Gleiches findet man auf der Seite von *Amazon*, die zwar primär eine einfeldige Suche bietet, aber in einem zweiten Schritt verschiedene Filtermöglichkeiten präsentiert, um die Suchanfrage zu spezifizieren. Es scheint also keine Notwendigkeit zu sein, die Such- und Filteroptionen an *Google* und *Wikipedia* zu orientieren und sich auf ein einziges Suchfeld oder versteckte Filteroptionen (wie bei *Google*) zu beschränken. Ferienportale wie auch Online-Shops bieten detaillierte Such- und Filtermöglichkeiten, die die Vielfalt der verfügbaren Daten und Informationen auf diese Weise für jede\*n Nutzer\*in bestmöglich reduzieren. Solange man in der Lexikographie vom Lemma ausgehend und in Artikeln als konstitutives Element von Wörterbüchern denkt, erscheinen diese Filtermöglichkeiten überflüssig, weil die Informationen in eine klare Hierarchie (bspw. Lemma > Bedeutungen > Belege) eingebunden sind. Nimmt man aber andere Fragen in den Blick, sieht die Situation anders aus: Beispielsweise könnte man die *Variantengrammatik* danach befragen, welche grammatischen



Varianten nur in der Schweiz vorkommen. Diese Ergebnisse mittels eines Filters weiter zu reduzieren, wäre ein nachvollziehbares Anliegen. Man könnte etwa danach filtern, welche Varianten in bestimmten Regionen häufig gebraucht werden oder welche in den Bereich der Verbflexion fallen. Tatsächlich steht eine Datenvielfalt hinter digitalen Sprachressourcen, die sich mit einem einzigen lemmafixierten Suchfeld nicht erschließen lässt. Das Auffinden dieser Informationen hängt aber unmittelbar von den angebotenen Suchoptionen ab (vgl. Dräger 2011, 237). Für die *Variantengrammatik* wären über die

obigen Beispiele hinaus auch Fragen wie „Welche grammatischen Phänomene sind besonders variantenreich?“ oder „Welche Kategorie weist besonders viele Einzelphänomene auf?“ denkbar. Solche artikelübergreifenden Fragestellungen, die man an ein gedrucktes Wörterbuch nicht stellen kann, könnten die Attraktivität und Nutzbarkeit digitaler Angebote deutlich erhöhen. Auch der Mehrwert für ein forschendes Publikum stiege würde steigen und vielleicht könnte auf diese Weise auch bei Personen, die nur kurz eine Information nachschlagen wollten, ein Forschungsinteresse geweckt werden.

### 6.3.2 Filter und Kategorisierungen

Zur detaillierten Suche und Verfeinerung des Suchergebnisses lassen sich Filter und Kategorisierungen nutzen. Wenn diese so präsentiert werden wie bei *Amazon*, werden gleichzeitig auch die Strukturen der gespeicherten Informationen offengelegt. So zeigt das oben abgebildete Beispiel von *Amazon*, dass der gesuchte Begriff (hier „Kollokationenwörterbuch“) in verschiedenen Unterkategorien der Kategorie „Bücher“ vertreten ist. Solche Strukturen gibt es in digitalen Nachschlagewerken auch. In der *Varietengrammatik* werden diese z.B. als Verlinkungen dargestellt (vgl. den Screenshot).

#### Siehe auch

- Grundlagenartikel: Wortbildung (oder direkt Unterkapitel Substantivderivation)
- Überblicksartikel: Suffixvarianten bei Substantiven – Einzelfälle
- Einzelartikel: Abschreiber / Abschreibung, Ausbilder / Ausbildner, Bauherr / Bauherrschaft, Beihilfe / Gehilfenschaft, Fülle / Füllung, Fundament / Fund(ament)ierung / Fundation, Initiator / Initiator, Kai / Kaje / Quai, Kappe / Kapperl, Kommandant / Kommandeur, Krimineller / Kriminologist, Leibchen / Leiberl, Matur / Matura / Maturität, Patronanz / Patronat, Pensionär / Pensionierte / Pensionist, Pfarre / Pfarrei, Spatzen / Spatzle / Spatzli, Sperre / Sperrung, Täter / Täterschaft, Telefon / Telefonat
- Thematisch verwandte Artikel: Packerl, Pluralbildung mit / ohne Endung -n

In *lexiko* wird eine erweiterte Stichwortsuche angeboten, die eine Filtrierung nach unterschiedlichen „Merkmale“ der einzelnen Artikel erlaubt. Auch hier erkennt man, dass durch die Suchmaske die Artikelstrukturen durchscheinen.

Das Beispiel der als Wortwolke dargestellten Diminutiva (vgl. Kap. 5.1.2.5) zeigt eine weiterte Möglichkeit der Visualisierung von verborgenen Zusammenhängen zwischen Informationen. Im Gegensatz zu Portalen wie *booking.com* oder *Amazon*, bei denen die Detailsuche oder Kategorisierung offensiv angeboten oder gar eingefordert wird, fällt bei Sprachressourcen auf, dass Recherche- und Filteroptionen wesentlich zurückhaltender integriert wurden. Dahinter steckt möglicherweise ein Punkt, der noch genauer betrachtet werden müsste: Zeichnet sich eine hohe Usability dadurch aus, dass man mit wenigen Eingaben zum Ergebnis kommt, oder zeichnet sie sich dadurch aus, dass die Möglichkeiten

der Informationssuche innerhalb eines Angebots gut sichtbar und leicht nutzbar angeboten werden?

### 6.3.3 Ranking der Ergebnisse

In den Abschnitten zur Suchmaschinenoptimierung (vgl. Kap. 5) wurde deutlich, dass lexikographische Werke viele vorteilhafte Eigenschaften für eine gute Positionierung in Suchmaschinen mitbringen. Dennoch lässt sich insbesondere die Darstellung in den Suchmaschinenergebnissen nicht so gestalten und beeinflussen, wie es wünschenswert und vorteilhaft wäre. *Google*, aber auch das Video-Portal

*YouTube* zeigen, dass die algorithmische Sortierung oftmals nicht in zufriedenstellender Weise zu den Interessen der Suchenden passt. Doch bei *YouTube* und auch in Online-Shops gibt es die Möglichkeit, bestimmte Sortierungen vorzugeben (etwa nach Preis, Aktualität oder Beliebtheit). Der Rankingmodus lässt sich also auswählen. Das schränkt den kritischen Punkt beim Ranking in Suchmaschinen ein, bei welchen nicht die relevanteste oder am besten geeignete Informationsseite oben steht, sondern jene, die der Logik der Algorithmen am besten entspricht. Auch in der Lexikographie wären solche kriteriengestützte Rankings denkbar (etwa nach Metadaten wie Rezeptionshäufigkeit, Nützlichkeit, Aktualität oder Expertenbewertung). Diese Sortieroptionen lassen sich auch auf Belege übertragen, die üblicherweise in einer von der Redaktion gerankten festen Listenform präsentiert werden (bspw. könnten jeweils von Nutzenden als besonders aussagekräftig bewertete Belege an erster Stelle präsentiert werden). Selbst die lexikographischen Textsegmente eines Artikels (bspw. grammatischer Kommentar, Bedeutungsparaphrase, Belege, etc.) sind kleine Informationscluster, die gemeinsam einen Artikel bilden und deshalb in einer fixen Reihenfolge präsentiert werden. Wieso sollte diese Reihenfolge nicht umsortiert werden können? Hinter der Artikelstruktur steht eine lexikographische Tradition und Theorie, aber natürlich auch eine immanente Logik (vgl. Wiegand 1989), die geübte Wörterbuchnutzende kennen und wahrscheinlich nicht hinterfragen. Doch aus Sicht der Internet-Usability müsste möglicherweise die Information an erster Stelle stehen, die besonders relevant für die Suchanfrage ist. Mit dem Feature unterschiedlicher – und von der Suchanfrage abhängiger – Anordnungen der Informationscluster eines Artikels könnte sich für die Lexikographie eine Möglichkeit ergeben, diese Artikelinformationen leicht zugänglich zu machen, ohne den Umfang eines informativen wissenschaftlichen Artikels reduzieren zu müssen.

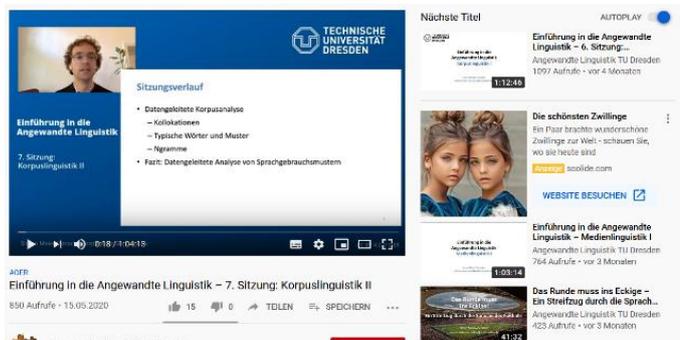
## 6.4 Lexikographische Videos



Die Seite *youtube.ch* zeigt anschaulich, wie Besuchende in den Bann von weiteren Angeboten gezogen werden sollen. Auch hier sind Algorithmen am Werk. Besonders beliebte Videos werden präsentiert und durch die Angabe der Aufrufe (hier 5,7 Mio.) als attraktiv bewertet. Ausschlaggebend für ein solches System sind Nutzendendaten. Wer ein Video betrachtet, bekommt thematisch passende Videos angeboten, die andere auch gesehen haben. Hierfür werden

Informationen aus dem Kontext des betreffenden Videos berücksichtigt (bspw. Schlagworte, Kanalbetreiber, Aufrufe).

Ob das ein Weg für die wissenschaftliche Lexikographie ist, um für ihre Inhalte zu begeistern, scheint fraglich. Allerdings kann man die Idee mitnehmen, dass sich Internetnutzende mit Videos zu einer längeren Rezeption umfangreicher Inhalte bewegen lassen: Der oben aufgeführte Videoclip mit 5,7 Mio. Aufrufen dauert fast 37 Minuten. Liegt darin eine Perspektive für digitale Sprachressourcen, die umfangreiche Informationen bieten und sich ebenfalls längere Rezeptionszeiten wünschen?



Auf [duden.de/woerterbuch](https://www.duden.de/woerterbuch) wird ein Video mit dem Titel „Wie kommt ein Wort in den Duden?“ angeboten. Es wäre interessant, die Nutzung solcher Angebote zu evaluieren und auf diese Weise ein alternatives multimediales Darstellungsformat für lexikographische Artikel zu entwickeln. Das würde u.a. im Hinblick auf die Auffindbarkeit neue Perspektiven eröffnen, weil es neben der *Google-Bildersuche* (vgl. Kap. 5.1.2.5) auch eine Suche für Videos gibt. Das Videoportal *YouTube*, das zum selben Konzern gehört wie *Google*, gilt als zweitgrößte Suchmaschine und ist nach *Google* die zweitmeist besuchte Seite im Netz.<sup>52</sup> Dass sich auf solchen Video-Portalen auch Erklärvideos einer gewissen Beliebtheit erfreuen, kann man als Chance für lexikographische Videoerklärungen sehen. Wie für die Bildersuche beschrieben, könnten beliebte Artikel einer Sprachressource mit einem Video ergänzt werden, das auf diese Weise zu ihrer Auffindbarkeit, Sichtbarkeit und Usability beitragen würde. Für die *Variantengrammatik* läge es aufgrund des schon bestehenden und informativen Bildmaterials nahe, mit animierten Karten und einem auf den lexikographischen Artikeln basierenden mündlichen Kommentar zu experimentieren. In diesem Zusammenhang wären auch Experimente interessant, mit denen untersucht wird, wie sich die Rezeptionsdauer und Informationsaufnahme mit einem Videoformat im Vergleich zu textbasierten Artikeln verändert.

## 6.5 Beurteilung der Wertigkeit

Weiter oben wurde thematisiert, dass Nutzende von digitalen Sprachressourcen Wertigkeit, Vertrauenswürdigkeit und Verlässlichkeit erwarten (vgl. Kap. 3.5). Unklar ist allerdings, wie sie ihr Urteil über diese Qualitäten fällen. Deutlich wurde, dass auf der Artikelseite entsprechende Signale vorhanden sein sollten. Doch es zeigte sich auch, dass man mangels entsprechender Untersuchungen in der Lexikographie auf Vermutungen zurückgreifen muss, welche Signale die entsprechenden Qualitäten betonen. Man sollte keinesfalls dem Irrtum erliegen, dass das, was man als Lexikograph\*in für wertig, verlässlich und vertrauenswürdig erachtet, auch von den Nutzenden so beurteilt wird. Auch hier ist der Blick auf andere Informationsseiten aufschlussreich. Wo beispielsweise finden Kaufinteressierte in Online-Shops und auf Online-Marktplätzen Hinweise über die Qualität eines Produktes? Nach welchen Kriterien beurteilen Reisewillige, welche Unterkunft sie wählen sollen? Einerseits stehen als

Fabelhaft **8,7**  
3.166 Bewertungen  
Lage 9,5  
Preise anzeigen

**Feste Wortverbir**  
von Annelies Häcki Bl.  
★★★★★ ~ 11  
**Gebundenes Buch**  
**29,99€**

Informationsquellen die Titel, die Beschreibungen und mögliche Bilder zur Auswahl, womit man ein Urteil über ein Produkt, eine Unterkunft oder eine Information fällen kann. Andererseits sehen Kaufwillige bei *Amazon* eine mit Sternen visualisierte Kundenbewertung; bei *Booking.com* sind die

Unterkünfte mit Noten und einem entsprechenden Adjektiv (bspw. „Fabelhaft“) bewertet. Das Qualitätsurteil beeinflusst die Wahrnehmung des entsprechenden Angebots. Ist es nun auch in der Lexikographie denkbar, dass die Leserinnen und Leser eines Artikels (oder besser: einer Information) diesen bewerten und anderen damit eine Einschätzung über die Wertigkeit, die Nützlichkeit, die Verständlichkeit oder die Verlässlichkeit geben?

In dem Artikel „Kundenbewertungen in der eTouristik – Segen oder Fluch?“ schreibt Schmeißer (2010, 51): „Empfehlungen und Bewertungen von Kunden sind so alt wie Amazon und eBay. Trotzdem haben sie in den vergangenen Jahren den Reisemarkt grundlegend verändert. Reisende informieren sich anhand von Bewertungen anderer Reisender, sichern ihre Buchungsentscheidung ab und schreiben im

<sup>52</sup> [https://de.wikipedia.org/wiki/Liste\\_der\\_meistaufgerufenen\\_Websites](https://de.wikipedia.org/wiki/Liste_der_meistaufgerufenen_Websites); 04.10.2020.

Internet über ihre eigenen Reiseerfahrungen. Die klassischen Reiseanbieter standen den Nutzerbewertungen lange Zeit kritisch gegenüber, inzwischen sind Bewertungsfunktionen auf fast allen Portalen etabliert. Sowohl Verbreitung als auch Wirkung nutzergenerierter Inhalte werden aber häufig falsch eingeschätzt. Gerade aus Kundensicht sind die Meinungen anderer Nutzer zwar essentiell, unterliegen aber vielfältigen Urteilsheuristiken und erfüllen unterschiedliche Funktionen innerhalb des individuellen Entscheidungsprozesses.“ Er führt aus, dass sich solche Bewertungen langsam etablieren können, dass es Menschen gibt, die gerne Urteile abgeben und dass die Beurteilungen durchaus differenziert und nutzbringend sein können. Ein solches System ist m.E. auch eine Option für digitale Sprachressourcen und könnte durchaus Vorteile mit sich bringen, wenn man davon ausgeht, dass Nutzende wertige Inhalte erkennen und auch entsprechend beurteilen.

Ähnlich verhält es sich mit den Empfehlungen nach dem Prinzip „Wird oft zusammengekauft mit“, wie sie in Kap. 5.3.2.4 thematisiert wurden. Auch hier wird die durch einen Kauf attestierte Entscheidung anderer als Entscheidungs- und implizites Qualitätskriterium genutzt. Der Vorteil im Gegensatz zu einer ebenfalls denkbaren Rubrik wie „Die Redaktion empfiehlt folgenden Artikel“ oder „Artikel des Tages“ (gesehen auf *dwds.de*) ist, dass das Kauf- bzw. Nutzungsverhalten automatisiert ausgewertet werden kann und dass die Interessen und das Entscheidungsverhalten der Nutzenden einfließen.

Helfen Online-Bewertungen also dabei, bestimmte Nachschlagewerke oder Informationen im Sinne einer besseren Benutzbarkeit zu beurteilen? Dies ist in der wissenschaftlichen Community zweifellos umstritten. Doch gerade eine ausformulierte oder in Kategorien differenzierte Kritik (bspw. zur Verständlichkeit, Aktualität oder auch Anwendbarkeit einer Information oder eines Artikels) wäre ein wichtiger Beitrag zu den zentralen Fragestellungen der Benutzendenforschung. Schauen wir hierzu nochmal auf die Darstellung eines Suchergebnisses von *Google*.

Suchen | ekelig [selten] [eklig] - Duden

<https://www.duden.de> > suchen > dudenonline > ekelig [selten] [eklig] > eklig, selten ekelig. Adjektiv - 1a. ekelerregend, widerwärtig, abscheulich; 1b. in seinem Auftreten unangenehm; gemein, ...2. sehr, ganz gehörig, tüchtig.

ekelig / eklig – Variantengrammatik des Standarddeutschen

[mediawiki.ids-mannheim.de](https://mediawiki.ids-mannheim.de) > VarGra > index.php > Ekelig\_ > \_eklig > Die Adjektivvarianten ekelig und eklig werden mit oder ohne -e- in den Bedeutungen 'ekelerregend' und 'unangenehm, merkwürdig' nebeneinander verwendet.

Ekelig oder Eklig? (Schule, deutsch, Rechtschreibung) - Gutefrage

<https://www.gutefrage.net> > frage > ekelig-oder-eklig > 10 Antworten  
16.11.2016 - Beide Varianten sind möglich, allerdings verwendet man eher eklig. Der Ausdruck ekelig ist veraltet.

DWDS – ekelig – Worterklärung, Grammatik, Etymologie u. v. m.

<https://www.dwds.de> > ekelig > Ekel - ekeln - ekel - ekelhaft - eklig. Ekel m. 'Abscheu, heftiger Widerwille', im 16. Jh. vom Md. ausgehend (öfters bei Luther); wohl Substantivbildung zu ekeln (s.

eklig | Rechtschreibung - wie schreibt man eklig | <https://www...>

<https://www.wissen.de> > rechtschreibung > eklig > e k | lig < Adj. > ekelhaft; <auch> ekelig... .. ekelhaft, eklig, Ekel erregend, abscheulich, widerwärtig, unangenehm, schmerzlos, schmerzlos, qualitativ, abstoßend

Hier liefert nicht so sehr der einzelne Eintrag, sondern insbesondere der Vergleich der unterschiedlichen Einträge Entscheidungskriterien zur Auswahl eines Angebots. Denn als besonders schwierig erscheint es, die Wertigkeit und Verlässlichkeit einer Information beziehungsweise einer Quelle in dieser Auflistung zu beurteilen. Im Vergleich dazu bietet die Auflistung von Unterkünften bei Reiseportalen oder von Produkten bei Onlineshops leicht erfassbare zusätzliche quantitative Informationen (Zahl an Bewertungen) und qualitative Informationen (Sternbewertung). Letztlich geht es Unterkunftsanbietenden wie Sprachressourcen-Betreibenden: Sie möchten, dass der Wert und die Qualität ihres Angebots sichtbar und von den Betrachtenden berücksichtigt wird. Dazu passend schreibt Kemmer (2010, 3) über die Schwierigkeit der Beurteilung von Fach-Wörterbüchern und Laien-Wörterbüchern: „In Bezug auf OWB [Onlinewörterbücher] und deren Beurteilung heißt das nun, dass inzwischen ebenfalls lexikografische Laien am Werk sind. Dies kann, muss jedoch nicht, Aussagen über die Qualität eines OWB machen, zumal ohnehin erschwerend hinzukommt, dass anhand von Namen nicht zwangsläufig, d. h. nicht ohne Wissen über die sich dahinter verbergenden Personen, erkennbar ist, ob es sich um Laien oder Experten handelt.“

Wenn es ein Anliegen der wissenschaftlichen Lexikographie ist, die Wertigkeit der empirischen Methodik und des strukturierten Redaktionsprozesses zu einem Werk unter Beweis zu stellen, dann geht das in der Weite des Internets nicht in derselben Weise wie in der analogen Welt. Ein Grund hierfür ist, dass die Signale fehlen, die jenseits des Internets die Wertigkeit eines Nachschlagewerks suggerieren (bspw. hoher Preis, dicker Einband, renommierte Reihe). Kurz vorbeikommende Besuchende im Internet können nur sehr schwer die inhaltlichen Vorzüge, die analytisch-beschreibende Präzision oder auch die Richtigkeit der Inhalte beurteilen. Ist das Erkenntnisinteresse gestillt, werden die meisten ihre Recherche beenden. Ein Vergleich mit anderen Quellen bleibt oft aus, und damit wird die gefundene Information unabhängig von ihrer Qualität und Verlässlichkeit akzeptiert, solange sie nicht offensichtlich falsch oder widersprüchlich erscheint. Um die Informationen zu beurteilen, um mögliche Irrtümer, Ungenauigkeiten oder Widersprüche zu erkennen, müssten mehrere Informationsangebote zu einem sprachlichen Phänomen gegenübergestellt werden. Hierfür würde eine Plattform, die Informationen und nicht nur Zugänge zu Werken präsentiert, hervorragende Möglichkeiten bieten.

## 6.6 Finanzierung lexikographischer Angebote im Internet

Im Zusammenhang mit einer solchen Plattform, die idealerweise sehr viele Quellen zusammenführt, stellt sich die Frage nach der Finanzierung. In der Online-Umfrage zum DL-Projekt konnten sich 64 Prozent der Teilnehmenden Werbung in Online-Nachschlagewerken vorstellen. Die *Duden*-Website zeigt, dass dies akzeptiert wird. Das *Oxford English Dictionary* hat gar eine Bezahlschranke. Allerdings zeichnen sich beide Online-Ressourcen dadurch aus, dass sie auf traditionellen Referenzwerken aufbauen. Es ist interessant zu überprüfen, wie andere Angebote im Netz sich jenseits von Bannerwerbung finanzieren. Hier einige Beobachtungen: Eine gängige Praxis sind kostenpflichtige Zusatzangebote: Auf *Amazon* kann man ohne Zugangsbeschränkung einkaufen, aber mit dem *prime*-Angebot bekommt man exklusive Konditionen (bspw. bei der Lieferzeit). Tageszeitungen bieten nur noch einen Teil an tagesaktuellen Artikeln kostenlos an, andere aufwändiger recherchierte Artikel gibt es nur mit einem Abo. Bei *YouTube* kann man mit einem Premium-Abo die Werbung abstellen und einige Zusatzfunktionen nutzen. Auch *Duden* bietet eine kostenpflichtige Version ohne Werbung und einen kostenpflichtigen „Mentor“ an, der beim Schreiben korrekter Texte helfen soll. Auf *ricardo.ch*, einem schweizerischen Auktionsportal ähnlich wie *eBay*, oder beim Marktplatz *Amazon* zahlen die Verkaufenden eine Gebühr bei erfolgreichem Verkauf ihrer Ware. *Google* verkauft unter dem Namen *Google Ads* die obersten Positionen der Suchergebnisanzeigen an die Meistbietenden. All das funktioniert, weil diese Seiten stark frequentiert sind. Die Besuchenden dieser Seiten kommen nicht, weil *Google* ein hervorragendes Produkt hat, weil *YouTube* die besten Filme produziert oder weil *Amazon* die besten Produkte anbietet, sondern weil man auf den Portalen beziehungsweise Suchmaschinen nahezu alles findet. Die Anbieter ihrerseits kommen, weil sie mit ihrem Angebot von den vielen Besuchenden gesehen werden wollen. Ein interessanter Aspekt dabei ist, dass die Betreibenden der Seiten *Google* und *YouTube* oder *Ricardo* (und mit Einschränkungen auch *Amazon*) die Informationen, Videos oder Produkte nicht selbst anbieten. Sie werden entweder aus dem Netz aufgegriffen (bei *Google*) oder von anderen auf den Portalen eingestellt. Die Kundinnen und Kunden der Plattformen zahlen oftmals sogar dafür, um in diesen Webangeboten vertreten zu sein.

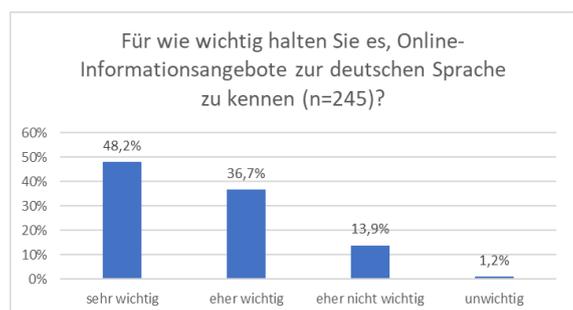
Nun stellt sich die Frage: Würden die Förderer oder Verantwortlichen von wissenschaftlich erstellten digitalen Sprachressourcen auch eine Gebühr dafür zahlen, dass sie auf einem Portal publizieren können, wo eine hohe Nachfrage nach Sprachinformationen besteht und eine gute Auffindbarkeit

gewährleistet ist? Für Händlerinnen und Händler fallen solche Ausgaben unter Werbekosten, weil sie damit die Interessentinnen und Interessenten für ihre Ware erreichen. Die Kaufenden bezahlen nur für das eigentliche Produkt und keine über den Verkaufspreis hinausgehende Gebühr (allenfalls etwas für Zusatzleistungen, bspw. einen schnelleren Versand). Auch Printwörterbücher haben Werbekosten (vgl. Kap. 2.3), die allerdings in den sogenannten Druckkosten versteckt sind, weil die Werbung Aufgabe der Verlage ist. Diese Druckkosten – und damit die Werbekosten – fallen bei Open-Access-Publikationen weg, was zu dem oben geschilderten Dilemma führt, weil online nur der Druck, aber nicht die Werbenotwendigkeit entfällt. Dennoch investiert die Wissenschaft weiterhin Geld für Publikationen in renommierten Verlagsreihen – und dies, obwohl es deutlich günstigere Publikationsmöglichkeiten mit freiem Zugang (Open-Access) für alle gibt. Aber dort ist die Rezeption der Angebote mutmaßlich schlechter, sie werden nicht wahrgenommen. Daher muss es auch ein Gedanke zur Zukunft der Lexikographie werden, dass die Publikation der erstellten Produkte zwar keine Druckkosten verursacht, durchaus aber beträchtliche Werbekosten notwendig sein werden. Eine Plattform, auf der die Informationen aus digitalen Sprachressourcen qualitativ und gut nutzbar publiziert werden, könnte dieses Umdenken begünstigen.

## 6.7 Vision: Ein Marktplatz für Sprachinformationen

In digitalen und gedruckten Wörterbüchern stecken redaktionell und empirisch erarbeitete Informationen über die Sprache und den Sprachgebrauch. Diese Informationen sind prinzipiell interessant für ein Fachpublikum, aber auch für viele andere Personen mit einem entsprechenden Frageinteresse. Die Log-File-Analysen des Projekts „Digitale Sprachressourcen“ deuten darauf hin, dass die Artikel tatsächlich auch gelesen werden, wenn sie denn in den Weiten des Internets gefunden wurden. Auch die Tests mit *Googles* Keyword-Planer zeigen: Es gibt ein großes Potential an Interessierten, das durch eine Verbesserung der Auffindbarkeit der lexikographischen Informationen erschlossen werden kann. Der große Vorteil des Internets, dass prinzipiell alles auffindbar ist, was dort im Open-Access publiziert wird, ist gleichzeitig auch eine große Herausforderung: Die Auffindbarkeit muss gewährleistet werden, sonst bleiben die lexikographischen Informationen ungenutzt. Zwar werden digitale Nachschlagewerke möglicherweise häufiger genutzt als Printwörterbücher in Bibliotheken, aber die Situation ist keinesfalls vergleichbar. Denn die digitalen Informationen werden nicht zwangsläufig bei der Internetrecherche gefunden, sie sind aber prinzipiell frei zugänglich. Das wiederum könnten Anbietende ausnutzen, die die lexikographischen Daten mit einer kommerziellen Absicht abrufen und im Rahmen eigener Angebote nutzen. Dabei blieben die Informationen als solche gratis, würden aber vorrangig zu dem Zwecke zusammengestellt, dass möglichst viele Menschen die Website besuchen – und sich somit beispielsweise Bannerwerbung rechnet.

In der DL-Umfrage zur Wörterbuchnutzung (Adarve 2020) gaben fast 85 % der Befragten an, dass sie es für „sehr wichtig“ oder „eher wichtig“ hielten, „Online-Informationsangebote zur deutschen Sprache zu kennen“. Man darf aus dieser Umfrage nicht schließen, dass diese Personen an fünf oder gar zehn Informationsangebote dachten, die es vielleicht bräuchte, um das Spektrum der fachlichen und sprachlichen Vielfalt abzudecken. Faktisch dürfte es darauf hinauslaufen, dass die Menschen ein oder zwei Angebote kennen und diese regelmäßig



nutzen. Für alle anderen digitalen Sprachressourcen würde das ein Nischendasein mit einer geringen Nutzungshäufigkeit bedeuten. Neben den in Kapitel 6.2 genannten Gründen spricht auch dies dafür, einen zentralen Zugang zu Informationen aus möglichst vielen digitalen Sprachressourcen zu schaffen. Eine solche Plattform wäre das Publikationsorgan für digitale Sprachressourcen und könnte die Anstrengungen zur besseren Auffindbarkeit, Sichtbarkeit und Usability bündeln. Zudem wäre es damit leichter, gegenüber den Förderinstituten zu argumentieren, dass mit einer Website noch keine ausreichende Open-Access-Publikation von lexikographischen Daten gewährleistet ist (vgl. Kap. 2.3). Es würde klar, dass für Auffindbarkeit im Netz auch bei scheinbar günstigen digitalen Publikation Werbekosten anfallen. Das könnte und müsste ein Umdenken in der Förderung lexikographischer Projekte in Gang setzen und verdeutlichen, dass Publikationskosten nicht nur Druckkosten sind, die im Digitalen augenscheinlich wegfallen.

Man kann abschließend festhalten: Über die Analyse verschiedener Internetangebote vom Videoportal, der Tageszeitung bis zum Online-Marktplatz konnten einige Features identifiziert werden, die auch für digitale Sprachressourcen gewinnbringend erscheinen. Auf dieser Basis wurde die These in den Raum gestellt, dass eine Plattform für Sprachinformationen nach dem Konzept von *Amazons* Online-Marktplatz konstruiert sein könnte. Das kehrt die Idee, dass Informationen im Netz oder auf einem Portal gesucht werden, um: Informationen werden von den verschiedenen Sprachressourcen innerhalb vordefinierter Kategorien angeboten. Sowohl für eine Suchmaschinenoptimierung als auch für die Umsetzung einer Plattform für Sprachressourcen sind die nötigen Daten und Informationen wie auch adaptierbare IT-Lösungen vorhanden. Grundsätzlich verfolgt das hier skizzierte keine lineare Fortführung lexikographischer Traditionen und Konzepte. Es wurde vielmehr versucht aufzuzeigen, dass die Digitalisierung neue und noch ungenutzte Potentiale für die Lexikographie mit sich bringt (vgl. Kap. 6.3.). In dieser Hinsicht sind die vielseitigen Inhalte in den zahlreichen digitalen Sprachressourcen eine gute Voraussetzung für eine Erneuerung lexikographischer Online-Publikationskonzepte. Entscheidend dabei ist, dass relevante Informationen gefunden werden und dass auch der Wert der dahinterstehenden Werke erkannt und ihre Benutzungsfreundlichkeit geschätzt wird.

Zum jetzigen Zeitpunkt kann man nur spekulieren, wie das Nachschlagen in der Zukunft aussieht, welche Rolle Künstliche Intelligenz dabei spielen könnte und wer mit den verfügbaren Informationen Geld verdient. Es käme aber einer disruptiven Verkehrung der Ziele der wissenschaftlichen Lexikographie gleich, wenn sie trotz der prinzipiellen Chancen der Digitalisierung nur noch für die Bereitstellung von sprachbezogenen Daten und Informationen zuständig wäre und die Präsentation und Weitergabe dieser Erkenntnisse anderen überlassen würde.

## 7 Literatur

- Adarve, G. (2020): Deutschschweizerische Umfrage zur Nutzung von (digitalen) Sprachressourcen – eine Ergebnisübersicht, Zürich, elektronisch publiziert:  
<https://www.ds.uzh.ch/de/projekte/digitalesprachressourcen.html>.
- Almind, R. (2005): Designing Internet Dictionaries, in: *Hermes, Journal of Linguistics*, Vol. 34, 37 – 54
- Alpar, A./Koczy, M./Metzen, M. (2015): SEO - Strategie, Taktik und Technik. Online-Marketing mittels effektiver Suchmaschinenoptimierung. Wiesbaden.
- Beißwenger, M./Storrer A. (2011): Digitale Sprachressourcen in Lehramtsstudiengängen: Kompetenzen – Erfahrungen – Desiderate, in: *Journal for Language Technology and Computational Linguistics* 1, 119 – 139.
- Bergenholtz, H. (2011): Access to and Presentation of Needs-Adapted Data in Monofunctional Internet Dictionaries, in: Bergenholtz, H./Fuertes-Olivera, P.A. (Hrsg.): *e-Lexicography: The Internet, Digital Initiatives and Lexicography*, London, 30 – 53.
- Boelhouwer, B./Dykstra A./Sijens, H. (2018): Dictionary portals, in: Fuertes-Olivera, P. A. (Hrsg.): *The Routledge handbook of lexicography*, Abingdon, 754 – 766.
- Bowker, L. (2012): Meeting the Needs of Translators in the age of e-lexicography: Exploring the possibilities, in: Granger, S./Paquot, M. (Hrsg.): *Electronic Lexicography*, Oxford, 379 – 398.
- Burke, M. et al. (2005): High-cost banner blindness. Ads increase perceived workload, hinder visual search, and are forgotten, in: *ACM Transactions on Computer-Human Interaction*, 12/4, 423 – 445.
- de Schryver, G.-M. (2003): Lexicographers' Dreams in the Electronic-Dictionary Age, in: *International Journal of Lexicography* (16.2), 143 – 199.
- Dräger, M. (2009): Auf der Suche nach historischen Phrasemen – oder: Wörterbücher als Korpora, in: *Linguistik online* 39/2009, 33 – 64.
- Dräger, M. (2010): Phraseologische Nachschlagewerke im Fokus, in: J. Korhonen (Hrsg.): *Europhras 2008. Beiträge zur internationalen Phraseologiekonferenz vom 13. - 16.8.2008 in Helsinki*, Helsinki, 411 – 421.
- Dräger, M. (2011): *Der phraseologische Wandel und seine lexikographische Erfassung. Konzept des „Online-Lexikons zur diachronen Phraseologie (OLDPhras)“*, Dissertation, Freiburg, elektronisch publiziert: [www.freidok.uni-freiburg.de/volltexte/8528](http://www.freidok.uni-freiburg.de/volltexte/8528).
- Dräger, M. et al. (2014): Kollokationen in der Praxis. Zur Auswahl von Kollokationen für ein Wörterbuch, in: Jesenšek, V.: *Phraseologie und Kultur*, Maribor, 75 – 94
- Dürscheid, Ch. /Elspaß, S./Ziegler, A. (2019): "Variantengrammatik des Standarddeutschen" – das neue Online-Nachschlagewerk zur arealen Variation in der Grammatik des Deutschen, in: Eichinger, L./Plewnia, A. (Hrsg.): *Neues vom heutigen Deutsch. Empirisch – methodisch – theoretisch*, Berlin/Boston, 331 – 334.
- Dürscheid, Ch./Sutter P. (2020): Forschendes Lernen mit digitalen Sprachressourcen. Am Beispiel des Unterrichtsthemas „Sprachvarietäten und ihre gesellschaftliche Bedeutung“, in: *Der Deutschunterricht* 6/2020, 11 – 22.
- Dziemianko, A. (2012): On the use(fulness) of paper and electronic dictionaries, in: Granger, S./Paquot, M. (Hrsg.): *Electronic Lexicography*, Oxford, 319 – 342.

- Dziemianko, A. (2018): The role of online dictionary advertisements in language reception, production, and retention, in: *ReCALL* 31/1, 5 – 22.
- Gouws, R. H./Tarp, S. (2017): Information overload and data overload in lexicography. *International Journal of Lexicography*, 30(4), 389 – 415.
- Häcki Buhofer, A. et al. (Hrsg.) (2014): *Feste Wortverbindungen des Deutschen. Kollokationenwörterbuch für den Alltag*, Tübingen.
- Haß-Zumkehr, U. (2001): *Deutsche Wörterbücher – Brennpunkt von Sprach- und Kulturgeschichte*, Berlin/New York.
- Henrich, A. (2008): *Information-Retrieval 1. Grundlagen, Modelle und Anwendungen*, Vorlesungsskript online:  
[https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/wiai\\_lehrstuehle/medieninformatik/Dateien/Publikationen/2008/henrich-ir1-1.2.pdf](https://www.uni-bamberg.de/fileadmin/uni/fakultaeten/wiai_lehrstuehle/medieninformatik/Dateien/Publikationen/2008/henrich-ir1-1.2.pdf), 28.09.2020.
- Hildenbrandt, V./Klosa, A. (Hrsg.) (2016): *Lexikographische Prozesse bei Internetwörterbüchern*, Institut für Deutsche Sprache, Mannheim.
- Klosa, A./Koplenig, A./Töpel, A. (2014): Benutzerwünsche und -meinungen zu dem monolingualen deutschen Onlinewörterbuch *elexiko*, in: Müller-Spitzer, C. (Hrsg.): *Using Online Dictionaries*, Berlin/Boston, 281 – 381.
- Klosa A./Müller-Spitzer, C. (Hrsg.) (2016): *Internetlexikografie. Ein Kompendium*, Berlin/Boston.
- Koplenig, A./Meyer, P./Müller-Spitzer, C. (2014): Dictionary users do look up frequent words. A log file analysis, in: Müller-Spitzer, C. (Hrsg.): *Using Online Dictionaries*, Berlin/Boston, 229 – 250.
- Koplenig, A./Meyer, P./Müller-Spitzer, C. (2014): Questions of Design, in: Müller-Spitzer, C. (Hrsg.): *Using Online Dictionaries*, Berlin/Boston, 189 – 205.
- Kosem, I./Lew, R./Müller-Spitzer, C./Ribeiro Silveira, M./Wolfer, S. (2018): The image of the monolingual dictionary across Europe. Results of the European survey of dictionary use and culture, in: *International Journal of Lexicography*, 1 – 23.
- Lew, R. (2015): Opportunities and limitations of user studies, in: Tiberius, C./Müller-Spitzer, C. (Hrsg.): *Research into dictionary use/Wörterbuchbenutzungsforschung*, 5. Arbeitsbericht des wissenschaftlichen Netzwerks „Internetlexikografie“, Mannheim, 6 – 16.
- Lorentzen, H./Theilgaard, L. (2012): Online dictionaries – how do users find them and what do they do once they have?, in: Fjeld, R. V./ Torjusen J. M.(Hrsg.): *Proceedings of the 15th EURALEX International Congress*, Oslo, 654 – 660.
- McIntosh, C. et al. (2009): *The Oxford Collocations Dictionary for Students of English*, Oxford.
- Müller-Spitzer, C./Koplenig, A. (2014): Online dictionaries: expectations and demands, in: Müller-Spitzer, C. (Hrsg.): *Using online dictionaries*, Berlin/Boston, 143 – 188.
- Müller-Spitzer, C. (2016a): *Wörterbuchbenutzungsforschung*, in: Klosa, A./Müller-Spitzer, C. (Hrsg.): *Internetlexikografie. Ein Kompendium*, Berlin/Boston, 291 – 342.
- Müller-Spitzer, C. (2016b): Aufgaben und Relevanz der Wörterbuchbenutzungsforschung Mitte der 2010er Jahre, in: Schierholz, S. J. et al. (Hrsg.): *Wörterbuchforschung und Lexikographie*, Berlin/Boston, 275 – 294.
- Müller-Spitzer, C./Ribeiro Silveira, M./Wolfer, S./Kosem, I./Lew, R. (2018): Eine europaweite Umfrage zu Wörterbuchbenutzung und -kultur. Ergebnisse der deutschen Teilnehmenden, in *Sprachreport* 2/2018, 26–35.

- Pariser E. (2011): *The Filter Bubble: What the Internet Is Hiding from You*, New York.
- Quasthoff, U. (2010): Wörterbuch der Kollokationen im Deutschen, Berlin/New York.
- Schmeißer, D.R. (2010): Kundenbewertungen in der eTouristik – Segen oder Fluch? Psychologie der Reiseentscheidung im Social Web, in: Amersdorffer D. et al. (Hrsg.): *Social Web im Tourismus*, Berlin/Heidelberg, 51 – 56.
- Stark, Birgit/Dörr, Dieter/Aufenanger, Stefan (Hrsg.) (2014): *Die Googleisierung der Informationssuche. Suchmaschinen zwischen Nutzung und Regulierung*, Berlin/New York.
- Storrer, A. (2001): Digitale Wörterbücher als Hypertexte. Zur Nutzung des Hypertextkonzepts in der Lexikographie, in: Lemberg, I./Schröder, B./Storrer, A. (Hrsg.): *Chancen und Perspektiven computergestützter Lexikographie. Hypertext, Internet und SGML/XML für die Produktion und Publikation digitaler Wörterbücher*, Tübingen, 53 – 69.
- Tarp, S. (2015): Detecting user needs for new online dictionary projects, in: Tiberius, C./Müller-Spitzer, C. (Hrsg.): *Research into dictionary use/Wörterbuchbenutzungsforschung. 5. Arbeitsbericht des wissenschaftlichen Netzwerks „Internetlexikografie“*, Mannheim, 17 – 27.
- Weidenmann, B. (1995): Multicodierung und Multimodalität im Lernprozeß, in: Issing, L. J./Klimsa, P. (Hrsg.): *Information und Lernen mit Multimedia*, Weinheim, 65 – 84.
- Wiegand, H. E. (1989): Arten von Mikrostrukturen im allgemeinen einsprachigen Wörterbuch. In: Hausmann, F. J. et al. (Hrsg.): *Wörterbücher: Ein internationales Handbuch zur Lexikographie*, Berlin/New York. S. 462 – 501.
- Wiegand, H. E. (Hrsg.) (1998): *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. Teilband 1.*, Berlin/New York.