

## Neologismenlexikographie und das Internet

1	Einleitung	5.4	Rechtschreibung
2	Was sind Neologismen?	5.5	Angaben zur Grammatik
3	Typen von Neologismen	5.6	Angaben zu Bedeutung und Gebrauch
4	Warum Neologismenlexikographie?	5.7	Frequenzangaben
5	Neologismenlexikographie und das Internet	5.8	Publikation
5.1	Grundlagen und Infrastruktur	5.9	Vernetzung mehrerer Ressourcen
5.2	Lemmaselektion	6	Fazit
5.3	Bearbeitung von Artikeln	7	Literatur

Projekt „Digitales Wörterbuch der Deutschen Sprache“, Berlin-Brandenburgische Akademie der Wissenschaften

### *Abstract:*

In this article we will argue that the contemporary lexicography which deals with the collection and description of new words can profit substantially from the use of the internet both as a source of data and as a publication platform. We will plead for a stronger linking of the individual neologism projects in Germany on various levels. This linking of projects by itself can bring an added value to the user.

## 1 Einleitung

Mitte der neunziger Jahre gab es genügend Gründe, um pessimistisch auf einen Zweig der Lexikographie zu blicken, der in Deutschland ohne große Tradition war: die Neologismenlexikographie<sup>1</sup>. Ein in der DDR begonnenes und mit dem Ende der DDR abgebrochenes Projekt war zu der Zeit alles, auf das man mit einigem Stolz verweisen konnte. Dabei ist der Sprachwandel, der sich auf der Ebene der Lexik am deutlichsten manifestiert, schon immer ein Thema, das den sprachlichen Laien interessiert.

Zehn Jahre später bereits stellte sich die Situation deutlich anders dar. Damals wie heute können mindestens drei Projekte für sich beanspruchen, Neologismenlexikographie mit greifbarem Ergebnis zu betreiben: das mittlerweile digitalisierte Neologismenwörterbuch von Herberg et al. ([www.owid.de](http://www.owid.de)), die „Wortwarte“ ([www.wortwarte.de](http://www.wortwarte.de)) und das Neologismenwörterbuch von Uwe Quasthoff (2007). Die Existenz von großen Korpora des gegen-

---

<sup>1</sup> Vgl. Kinne 1996.

wärtigen Sprachgebrauchs ist sicher ein Grund dafür. Ein weiterer, wichtigerer, ist die Verkürzung der Publikationszyklen und damit die elektronische Verfügbarkeit von Texten über das Internet – der Abstand zwischen der Erstellung und der Verfügbarkeit eines Textes für die weitere Analyse ist heute eine Sache von Stunden. Dass die Zeit hier schnelllebiger geworden ist, das gereicht zumindest der Neologismenlexikographie zum Vorteil. Wir wollen im Folgenden darstellen, wie die gegenwärtige Neologismenlexikographie mit dieser Verbesserung der Datenbasis umgeht und welche Rolle das Internet – in der Form des World Wide Web – für die Publikation und die Nutzung der Arbeitsergebnisse spielt.

## 2 Was sind Neologismen?

Zunächst wollen wir den Gegenstand der Neologismenlexikographie näher bestimmen. Nach einer weithin akzeptierten Definition ist ein Neologismus

eine lexikalische Einheit bzw. eine Bedeutung, die in einem bestimmten Abschnitt der Sprachentwicklung in einer Kommunikationsgemeinschaft aufkommt, sich ausbreitet, als sprachliche Norm allgemein akzeptiert und in diesem Entwicklungsabschnitt von der Mehrheit der Sprachbenutzer über eine gewisse Zeit hin als neu empfunden wird (Herberg et al. 2004, S. XII).

Der mit dem Begriff ‚Neologismen‘ bezeichnete Wortschatzausschnitt wird also primär durch die zeitliche Dimension charakterisiert. Dies gilt in dreifacher Hinsicht:

1. Wenn man den Standpunkt des gegenwärtigen Benutzers einer Sprache einnimmt, dann sind Neologismen diejenigen sprachlichen Zeichen, die noch nicht lange in Gebrauch sind. Was *lange* bedeutet, das hängt vom Standpunkt des Benutzers ab. Für die Redaktion eines allgemeinsprachlichen Wörterbuchs ist dieser Horizont wahrscheinlich der Zeitraum seit Fertigstellung der letzten Auflage<sup>2</sup>. Man könnte also Neologismen in Anlehnung an eine ehemalige Rubrik in der Zeitschrift *Lebende Sprachen* als „Wörter, die noch nicht im Wörterbuch stehen“ bezeichnen.
2. Wenn man bei der lexikographischen Arbeit den betrachteten Zeitraum zurückversetzt, wie dies etwa die Projektgruppe Neologismenlexikographie am Institut für Deutsche Sprache macht, dann kann man als Neologismen diejenigen Wörter betrachten, die in dem betrachteten Zeitraum neu waren bzw. als neu empfunden wurden. Der Gegenstand der Betrachtung sind hier „Wörter, die mittlerweile im Wörterbuch stehen (aber damals noch nicht)“<sup>3</sup>.
3. Wenn man Wörterbücher konsultiert, die schon ein gewisses Alter haben, dann wird man evtl. auf Stichwörter stoßen, die z.B. als „Neuwort“ oder „Neulexem“ markiert sind. Dies ist z.B. im *Wörterbuch der deutschen Gegenwartssprache*<sup>4</sup> der Fall. Diese Angaben sind

<sup>2</sup> S. hierzu Scholze-Stubenrecht 2002.

<sup>3</sup> Es gibt möglicherweise auch Wörter, die zu einem bestimmten Zeitraum als neu empfunden wurden, aber nicht in einem aktuellen Wörterbuch verzeichnet sind. Entweder sind diese Wörter schon wieder außer Gebrauch, oder sie haben es nie in ein allgemeinsprachliches Wörterbuch geschafft. Wir belassen es bei diesem Hinweis und gehen nicht weiter auf diesen Sonderfall ein.

<sup>4</sup> Siehe unter dem Stichwort *Service* in der elektronischen Version dieses Wörterbuchs, [www.dwds.de](http://www.dwds.de).

aus heutiger Sicht obsolet. Sie bieten lediglich dem Sprachwandelforscher und dem Sprachhistoriker nützliche Informationen, nicht jedoch dem Benutzer in actu.

Wir wollen die oben erwähnte, von Herberg et al. eingeführte Definition als *Definition des Gegenstandes im engeren Sinn* bezeichnen und ihr eine *Definition im weiteren Sinne* gegenüberstellen; danach ist ein Neologismus

eine lexikalische Einheit bzw. eine Bedeutung, die in einem bestimmten Abschnitt der Sprachentwicklung in einer Kommunikationsgemeinschaft aufkommt und von der Mehrheit der Sprachbenutzer über eine gewisse Zeit hin als neu empfunden wird.

Der Unterschied zwischen diesen beiden Definitionen liegt darin, dass nach der zweiten Definition das Anerkenntnis eines sprachlichen Zeichens als Teil einer sprachlichen Norm und die damit verbundene Lexikalisierung dieses sprachlichen Zeichens nicht als konstitutives Merkmal eines Neologismus festgelegt wird.

Anhand der ersten, engeren Definition lässt sich eine retrospektive Neologismenlexikographie, wie sie das Institut für Deutsche Sprache praktiziert, begründen. Diese hat einen sprachdokumentarischen Wert, besonders wenn daraus so gut gemachte Produkte entstehen wie das Wörterbuch *Neologismen der 90er Jahre im Deutschen* (Herberg et al. 2004). Anhand der zweiten Definition konnte sich eine aktuelle Neologismenlexikographie entwickeln, insbesondere da heutzutage die Ergebnisse in einem Medium präsentiert werden können, das eine zeitnahe Aktualisierung ermöglicht. Ziel der aktuellen Neologismenlexikographie ist es, Wörter in statu nascendi, also noch vor ihrer Lexikalisierung, zu registrieren und zu beschreiben.

### 3 Typen von Neologismen

Auf der obersten Ebene können wir Neulexeme von Neubedeutungen unterscheiden:

- Ein Neulexem ist eine Verbindung einer bisher noch nicht registrierten lexikalischen Form mit einer oder mehreren noch nicht registrierten Bedeutungen (Bsp. *simsen*, das eine neue Art der Telekommunikation bezeichnet);
- Eine Neubedeutung entsteht aus der Verbindung einer bereits registrierten lexikalischen Form mit einer noch nicht registrierten Bedeutung (Beispiel *Lichterkette* in der neuen Bedeutung ‚Form des Protestes, bei der die Demonstranten sich in Form einer Kette anordnen und Lichter in den Händen halten‘<sup>5</sup>).

Die (neue oder bereits existierende) lexikalische Form kann dabei einfach oder komplex sein:

---

de. Die Kennzeichnung von Teilen des beschriebenen Wortschatzes als „Neulexem“ oder „Neubedeutung“ ist eine der Besonderheiten dieses Wörterbuchs, wenn man es mit anderen allgemesprachlichen Wörterbüchern des Deutschen vergleicht. Vgl. WDG, Bd. 1, S. 014.

<sup>5</sup> Vgl. Steffens 2002.

- Hin und wieder kommen neue Wortbildungselemente in Gebrauch. Als Beispiel sei hier das entlehnte Präfix *Cyber-/cyber-* genannt, das seit den neunziger Jahren für mehrere Hundert Neubildungen verantwortlich war. Das Aufkommen neuer nativer Wortbildungselemente ist eher selten. Häufiger ist der Fall, dass lexikalisierte Wörter zu Affixen „herabsinken“ (z.B. *X-mäßig*, *Y-technisch*<sup>6</sup>). Zu den Wortbildungselementen sind auch Lehnaffixe zu zählen, denen bereits in der Gebersprache kein echter Affixstatus zukommt (z.B. *-tainment* wie in *Desastertainment*, *-gate* wie in *Nipplegate*). Diese Elemente sind aber reihenbildend und sollten deshalb in einer Beschreibung des Wortschatzes der deutschen Sprache nicht ignoriert werden.
- Sehr selten sind Neubildungen als echte Simplexformen. Aus den *nuller Jahren* könnte man hier das Wort *simsen* anführen. Man kann diese Neuform aber auch als deverbale Bildung aus dem Nomen *SMS* (mit der Bedeutung ‚SMS verschicken‘) interpretieren. Ähnliches gilt für die ebenfalls recht neuen Kommunikationsverben *voipen*, *skypen* und *twittern*, die allesamt aus Produkt- oder Protokollnamen von Kommunikationsprogrammen gebildet wurden.
- Die bei weitem häufigsten Bildungsmuster für neue Wörter sind Derivation und Komposition. Viele dieser Bildungen sind allerdings spontan gebildete Okkasionalismen (=Gelegenheitsbildungen), die semantisch transparent sind und deshalb nicht einen Prozess der Lexikalisierung durchlaufen. ‚Transparenz‘ ist allerdings ein komparatives Prädikat. Die Neubildung *Abwrackprämie* (das Wort des Jahres 2009, <http://www.gfds.de/aktionen/wort-des-jahres/>) ist transparenter als das Wort *Herdprämie* (Unwort des Jahres 2007, vgl. <http://www.unwortdesjahres.org/>), wenn auch nicht völlig transparent, da der Gegenstand des Abwrackens nicht bezeichnet wird.
- Ebenfalls sehr selten ist die Bildung neuer Mehrwortausdrücke. Meist handelt es sich dabei um die kreative Adaption und Variierung existierender Mehrwortausdrücke (vgl. Wotjak 1992). Doris Steffens (2002) führt als ein Beispiel *den Ball flachhalten* an.

Schließlich kann man noch die lexikalisierten Neuwörter, also die Neologismen im engeren Sinn, von den Okkasionalismen bzw. Gelegenheitsbildungen abgrenzen. Okkasionalismen werden auch als ‚ad-hoc-Wörter‘ oder ‚Gelegenheitsbildungen‘ bezeichnet. Michael Kinne (1996) vergleicht diese Wörter sehr anschaulich mit Bällen, die von einem Spieler in die Luft geworfen, aber von keinem anderen Spieler aufgenommen werden. Betrachtete man die Okkasionalismen in den 90er Jahren noch als durchweg uninteressante Zufallsprodukte, so haben mittlerweile mehrere Arbeiten, u.a. von Corinna Peschel (2002) und Tomášiková (2008), ihren Ruf verbessert und gezeigt, dass diese Elemente zumindest für einige Bereiche der linguistischen Forschung (z.B. Textlinguistik, Wortbildung) und angewandten Linguistik (Übersetzung) höchst relevant sind.

Es sollte auch deutlich geworden sein, dass, wenn man die Definition von ‚Neologismus‘ im weiteren Sinn zugrundelegt, man nicht zwischen Okkasionalismen und lexikalisierten Neubildungen unterscheiden kann. Im Augenblick ihres Entstehens kann man einem Neuwort die Tendenz zur Lexikalisierung noch nicht ansehen.

Den Prozess der Lexikalisierung eines (Neu-)Wortes kann man gar nicht unmittelbar beobachten. Auch die Formulierung in der oben angeführten Definition von Herberg et al.

---

<sup>6</sup> Vgl. Ruge 2004.

ist letztlich schwer zu operationalisieren. Was bedeutet es, dass ein Wort ‚als sprachliche Norm allgemein akzeptiert‘ wird?

Es gibt aber einige indirekte Hinweise, aus denen geschlossen werden kann, wie weit der Prozess der Lexikalisierung vorangeschritten ist:

- Ist ein Wort noch nicht allgemein akzeptiert, dann wird es häufig – jedenfalls häufiger als bei anerkannten Wörtern – mit Distanzmarkern gebraucht. Der Sprecher oder Schreiber signalisiert, dass er dieses Wort zwar verwendet, es aber als Teil der deutschen Sprache noch nicht wirklich akzeptiert hat. Beispiele hierfür sind: Anführungszeichen (*es gibt jetzt schon „eKurse“*), Heckenwörter (*der sog. Schattenhaushalt ...*) und weitere distanzierende Floskeln (*sie hatten sich zum Smirting vor der Tür verabredet, wie man heute in Raucherkreisen sagt*) (pragmatisches Kriterium).
- Vor allem für Wörter, die aus dem Englischen entlehnt werden, muss sich ein stabiles Genus etablieren (*die Engine*), und es müssen vor allem für Verben geeignete Konjugationsformen gebildet werden (z.B. *downloaded* vs. *gedownloaded*). Je dauerhafter sich eine Form auf Kosten anderer Formen durchsetzt, umso weiter ist die Integration eines Wortes in den Wortschatz des Deutschen vorangeschritten (grammatisches Kriterium).
- Je öfter ein Wort als Wortteil in Komposita verwendet wird, umso stärker ist dieses sprachliche Zeichen integriert. Ein Beispiel ist die Verwendung von *Heuschrecke* im Sinne von ‚Finanzkapitalist, der Unternehmen ausschließlich zum Zwecke der späteren gewinnbringenden Kapitalisierung aufkauft und ggf. zerschlägt‘ – eine Bedeutung, die im vergangenen Jahrzehnt aufkam – in Komposita wie: *heuschreckenkompatibel*, *Heuschreckenliste*, *Heuschreckenverbot*<sup>7</sup>. Ein besonderes Kriterium ist für Lehnwörter, ob und wie oft sie mit nativen Wortbestandteilen kombiniert werden (Beispiel *Musikdownload*) (semantisches Kriterium).

#### 4 Warum Neologismenlexikographie?

Worin liegt nun die Relevanz der Neologismenlexikographie? Was kann sie leisten, was über die regelmäßige Aktualisierung der Lemmalisten allgemeinsprachlicher Wörterbücher hinausgeht? Auch bei der Diskussion dieser Frage wollen wir zwischen retrospektiver und aktueller Neologismenlexikographie unterscheiden. Die genuine Leistung der retrospektiven Lexikographie liegt in der Beschreibung von Schlüsselwörtern eines Diskurses, der im betrachteten Zeitraum geführt wurde. Damit dies gelingen kann, sind eine gute Auswahl der beschriebenen Stichwörter und eine Beschreibung des Gebrauchs dieser Wörter mittels gut ausgewählter Belege notwendig. Damit sind die Grenzen dieser Art der Neologismenlexikographie erreicht.

---

<sup>7</sup> Beispiele aus der Wortwarte, [www.wortwarte.de](http://www.wortwarte.de), dort finden sich auch Belege zu den genannten Wörtern.

Die aktuelle Neologismenlexikographie kommt im günstigsten Fall einem weit verbreiteten Informationsbedürfnis entgegen. Beim Gebrauch neuer Wörter sind die Normunsicherheiten der Sprecher größer als in anderen Bereichen des Wortschatzes, wie die folgenden Beispiele illustrieren mögen:

- Schreibt man *Carsharing*, *Car-Sharing* oder *Car-sharing*?
- Wie wird *E-Zine* ausgesprochen?
- Wie wurde das Wort *Bundesbanker* gebildet: *Bundes-banker* oder *Bundesbank-er*? Vom angenommenen Bildungsmuster hängt auch die Aussprache des Wortes ab.
- Heißt es *die Engine*, *der Engine* oder *das Engine*?
- Welche Genitivform ist korrekt: *des Piercings* oder *des Piercing*?
- Wie wird der Plural von *Flyer* gebildet: *die Flyer* oder *die Flyers*?

Dies sind nur einige der möglichen Normunsicherheiten. Normalerweise können solche Fragen durch die Konsultation eines allgemeinsprachlichen Wörterbuchs oder eines Spezialwörterbuchs für den entsprechenden Informationstyp (z.B. Rechtschreibwörterbuch bei orthografischer Normunsicherheit) geklärt werden. Neue Wörter sind aber in diesen Wörterbüchern noch nicht beschrieben. Eine gute aktuelle Neologismenlexikographie könnte diese Lücke schließen.

## 5 Neologismenlexikographie und das Internet

Im Folgenden wollen wir darstellen, welchen Mehrwert die Verwendung des World Wide Web für den lexikographischen Prozess hat, der zur Erstellung von Neologismensammlungen bzw. Neologismenwörterbüchern führt. Bei der Darstellung des prototypischen lexikographischen Prozesses orientieren wir uns an Engelberg/Lemnitzer (2009) und der dort erwähnten Literatur. Danach gliedert sich der lexikographische Prozess in die folgenden Phasen:

- Planung und Konzeption eines Wörterbuchs
- **Schaffung der Grundlagen und Erstellung der notwendigen Infrastruktur (Personal, technische Ausstattung, Ausgangsmaterial)**
- **Lemmaselektion und Erstellen der lexikographischen Beschreibung**
- Überarbeitung der Arbeitsergebnisse
- Wörterbuchherstellung
- **Publikation eines marktfähigen Produktes oder auch mehrerer Produkte**
- Pflege und Weiterentwicklung der lexikalischen Substanz (nach Engelberg/Lemnitzer (2009), S. 228f.)

Wir werden im Folgenden nur auf die hervorgehobenen Prozessschritte eingehen. Wir denken, dass die Nutzung des Internets bei diesen Schritten den größten Mehrwert bietet. Zusätzlich wollen wir auf den bei Engelberg/Lemnitzer an anderer Stelle (Abschnitt 2.6) erwähnten Aspekt der Vernetzung und Integration mehrerer lexikalischer Ressourcen eingehen.

## 5.1 Grundlagen und Infrastruktur

Ein Neologismenwörterbuch kann von einer Person (z.B. die Wortwarte) oder von einer Gruppe von Lexikographen (z.B. das Neologismenwörterbuch am IDS) betrieben werden. Entsprechend klein oder groß sind die Möglichkeiten der Datenaufbereitung und Artikelerstellung.

Eine wichtige Frage ist die der Datenbasis. Bei einem Projekt retrospektiver Neologismenlexikographie kann man, wie beim Neologismenwörterbuch des IDS, auf vorgefertigte Stichwort- und Belegsammlungen in Form einer Kartei und auf ein Referenzkorpus des entsprechenden Dokumentationszeitraums zurückgreifen<sup>8</sup>. Bei der tagesaktuellen Neologismenerfassung ist man auf die tägliche Sammlung von Daten angewiesen. Mittels Webcrawlern kann man z.B. bei den Online-Präsenzen von Tages- und Wochenzeitungen das am jeweiligen Tag bereitgestellte Angebot auf die eigene Festplatte herunterladen und dort weiterbearbeiten. Die Qualität der Daten auf den Websites seriöser Zeitungen ist zumeist ausreichend gut. Ein besonderes Problem stellen die Artikeldubletten dar – man lädt den gleichen Artikel unter Umständen an mehreren Tagen herunter. Ein Programm sollte aber in der Lage sein, textidentische Dateien zu identifizieren und auf ein einziges Exemplar zu reduzieren. Ein weiteres Problem sind die Werbe- und Hinweistexte, die den eigentlichen Artikeltext umrahmen und meist sehr schematisch sind. Aber auch für das Herausfiltern dieser Rahmentexte gibt es mittlerweile Programme<sup>9</sup>. Ein weniger leicht zu lösendes Problem ist das Auffinden und Löschen von Überblicks- oder Teaserseiten, auf denen mehrere Artikel nur angerissen werden – drei Punkte oder ein Wort wie *mehr* fordern dazu auf, die Fortsetzung eines Artikels auf einer weiteren Seite zu lesen. Das Fazit ist, dass ein Monitorkorpus auf Basis einer regelmäßig von verlässlichen Websites heruntergeladenen Datenmenge eine gute Ausgangsbasis für das Auffinden und Beschreiben von neuen Wörtern ist. Es bedarf aber einiger automatischer Schritte der Datenaufbereitung, um diese Daten in eine für den lexikographischen Zweck nutzbare Form zu bringen. Man sollte bei einem auf langfristige Arbeit und öffentliche Sichtbarkeit angelegten Projekt frühzeitig die rechtlichen Bedingungen der weiteren Nutzung der Daten mit den Betreibern der Dienste und Websites, deren man sich bedient, abklären.

Einige eher konzeptuelle Fragen sind damit aber noch nicht gelöst. Die Auswahl der Texte in einem Monitorkorpus ist mehr oder weniger zufällig. Man ist immer weit davon entfernt, eine Sprache in ihrer Gesamtheit zu erfassen. Webkorpora sind (fast) immer schriftsprachlich, in dem o.g. Szenario obendrein auf journalistische Sprachproduktion beschränkt. Die wichtigste Einschränkung, die man auf Grund einer solchen Datenbasis machen muss, ist der Verzicht auf den Erstbeleg eines neuen Wortes. Ein Wort kann an anderer Stelle bereits in Umlauf sein, ehe es in einer der im Monitorkorpus erfassten Publikationen erstmals verwendet wurde. Die Angaben zum „Aufkommen“ eines Wortes sind also immer nur ungefähre Angaben. Wenn man retrospektive Neologismenlexikographie betreibt, ist die Suche nach dem Ursprung eines neuen Wortes nicht immer einfach, aber in vielen Fällen möglich. So lässt sich der Ursprung der o.g. Neubedeutung des Wortes

---

<sup>8</sup> Vgl. Herberg et al. 2004, S. XV-XVI.

<sup>9</sup> Vgl. Evert 2008.

*Heuschrecke* ziemlich sicher auf ein Interview mit Franz Müntefering zurückführen (April 2005, vgl. <http://de.wikipedia.org/wiki/Heuschreckendebatte>).

Möchte man Neologismenlexikographie auf der Grundlage eines Monitorkorpus betreiben, dann empfiehlt es sich, außer den o.g. Bereinigungswerkzeugen die folgenden sprachtechnologischen Werkzeuge zu verwenden:

- Einen Tokenizer, der die erfassten Texte in einzelne Wörter zerlegt,
- einen Lemmatisierer, der die Textwörter soweit möglich auf ihre Grundformen reduziert und
- einen Spracherkenner, um Texte in einer anderen Sprache, die zum Repertoire vieler Tages- und Wochenzeitungen gehören, herauszufiltern und zu löschen.

Darüber hinaus ist der Abgleich der Liste neuer Wörter bzw. Lemmata, die das Ergebnis der sprachtechnologischen Vorverarbeitung ist, mit den Wort- bzw. Lemmalisten großer Referenzkorpora zu empfehlen. Letztere repräsentieren sozusagen die bekannte Sprache, die das Komplement der neuen Wörter ist<sup>10</sup>. Die sog. Referenzwortlisten sollten ebenfalls tagesaktuell sein, also von den Wortlisten des Monitorkorpus gespeist werden.

## 5.2 Lemmaselektion

Der erste kreative Schritt im lexikographischen Prozess ist die Auswahl der Stichwörter, die beschrieben werden sollen. Bezogen auf die Neologismenlexikographie steht ein Projekt vor den folgenden Entscheidungen:

- Sollen nur Neulexeme erfasst werden oder auch Neubedeutungen?
- Sollen nur lexikalisierte Neologismen erfasst werden oder auch Okkasionalismen?
- Sollen nur Wörter oder auch Wortbildungselemente und mehrwortige Einheiten erfasst werden?

Die erste Entscheidung hängt stark von der technischen, insbesondere der sprachtechnologischen Infrastruktur des Projektes ab. Ein Neulexem ist leicht zu ermitteln, da es eine lexikalische Form hat, die bisher noch nicht beobachtet wurde. Eine Neubedeutung einer existierenden Form ist mit den heutigen sprachtechnologischen Mitteln nicht zu identifizieren. Die Ermittlung von Wortbedeutungen auf Grund der Distribution dieses Wortes ist Gegenstand der Forschung in der Computerlinguistik. Da es sich aber bei einer Neubedeutung zu einem bestimmten Wort zunächst immer um eine kleine Zahl von Belegen handelt – die Neubedeutung beginnt sich ja gerade erst zu etablieren –, dürfte die Distributionsmethode zur Erkennung einer Neubedeutung wenig erfolgreich sein. Auch hier ist die retrospektive Neologismenlexikographie im Vorteil, da sie mehr Belege eines Lexems in einer neuen Bedeutung auswerten kann. Komposita, in die diese Neubedeutung mittlerweile eingegangen ist, gehören ebenfalls zu den relevanten Distributionsdaten.

Die zweite Entscheidung ist einfacher, da sie sich aus der Datenlage ergibt. Betreibt man tagesaktuelle Lexikographie, dann ist die Erfassung von Okkasionalismen unvermeidlich – man kann sie natürlich nachträglich entfernen. Betreibt man retrospektive Neologis-

---

<sup>10</sup> Vgl. Lemnitzer/Zinsmeister 2006.



menlexikographie, dann wird man bemüht sein, sich auf die mittlerweile lexikalisierten Neologismen zu beschränken.

Die dritte Frage ist eher hypothetischer Natur. Keines der bekannten Projekte erfasst neue Wortbildungsbestandteile, auch Mehrwortlexeme sind eher die Ausnahme. Im Rahmen der retrospektiven Neologismenlexikographie ist es aber möglich, sowohl Wortbildungselemente als auch etablierte Mehrwortlexeme zu erfassen<sup>11</sup>.

### 5.3 Bearbeitung von Artikeln

Wie bereits oben gesagt, ist die Normunsicherheit der Sprecher bei der Verwendung neuer Wörter größer als bei der Verwendung etablierter Wörter. Das hängt auch damit zusammen, dass sich bei diesen Wörtern im Zuge ihrer Lexikalisierung ihre Schreibweise und Aussprache, ihr Genus, Flexionsparadigma etc. erst festigen. Es kann z.B. zeitweise mehrere konkurrierende Genera oder Flexionsparadigmen geben. Besonders die Erzeugnisse der aktuellen Neologismenlexikographie sind deshalb gefordert, die sich abzeichnende Norm zu kodieren bzw. auf noch konkurrierende Formen und Paradigmen hinzuweisen.

Prinzipiell können Neologismenwörterbücher das gesamte für allgemeinsprachliche Wörterbücher übliche Informationsprogramm haben. Wir wollen hier aber nur auf einige Angabentypen eingehen, die für Neologismenwörterbücher und für die oben angesprochenen Benutzungssituationen charakteristisch sind und bei denen der Bezug auf das World Wide Web von Nutzen ist:

- Angaben zur Orthographie, inkl. gebräuchlicher Schreibungsvarianten
- Angaben zur Grammatik (Wortart, Flexion etc.)
- Angaben zu Bedeutung und Gebrauch
- Frequenzangaben.

### 5.4 Rechtschreibung

Die reformierte Rechtschreibung gibt den Sprachbenutzern mehr Möglichkeiten zu alternativer Schreibung. Dies betrifft vor allem, aber nicht nur, den Bindestrich. Alternative, von den Konventionen abweichende Schreibung sollte registriert werden, vor allem wenn sie weiter verbreitet ist als die konventionelle Schreibung (z.B. *Zwei-Grad-Ziel* vs. *Zweigradziel*, *eLearning* vs. *E-Learning*). Die tagesaktuelle Neologismenlexikographie kann dies natürlich nur durch entsprechende regelmäßige Aktualisierung der Artikel leisten. In statu nascendi erscheint das Wort in einer bestimmten Schreibweise, die dann die Grundlage für den Lemmaansatz bildet. Die ständige Beobachtung und Auszählung von Schreibungsalternativen, die weitgehend automatisch auf den Daten eines aus Webdaten erstellten Monitorkorpus stattfinden kann, kann die Grundlage für diese Beschreibung bil-

---

<sup>11</sup> In Lemnitzer (2007) werden auch einige Wortbildungselemente beschrieben. Das Buch lässt sich als ein Produkt der retrospektiven Neologismenlexikographie einordnen, da in ihm ausschließlich lexikalische Einheiten beschrieben werden, die bereits früher in der *Wortwarte* erfasst wurden. In der *Wortwarte* werden aber nur Einzelwörter erfasst.

den. Im Neologismenwörterbuch von Quasthoff (2007) und in der digitalen Version des Neologismenwörterbuchs am IDS werden Schreibungsvarianten angegeben.

### 5.5 Angaben zur Grammatik

Wie bereits gesagt, ist besonders bei Lehnwörtern aus dem Englischen das Genus und zum Teil, vor allem bei den Verben, die Bildung der Flexionsformen öfters zunächst unsicher. Dies sollte bei der Erstellung tagesaktueller Artikel zunächst vermerkt werden. Auch hier hilft es, aus den Daten des Monitorkorpus entsprechende Angaben zu extrahieren und zu interpolieren und bei der Überarbeitung zu verlässlicheren Angaben zu kommen

### 5.6 Angaben zu Bedeutung und Gebrauch

Bei der Dokumentation des Gebrauchs eines Neuworts liegt eine der Hauptstärken der internetbasierten Neologismuslexikographie. Mit dem Wort erhält man ja auch mindestens einen Beleg. Die Belege können, wenn sie sorgfältig ausgewählt werden, wichtige Informationen zum Gebrauch und sogar zur Bedeutung eines Neuwortes liefern. Oftmals haben frühe Belege definitorischen Charakter, da das Neuwort ja erst einmal bei einer größeren Leserschaft eingeführt werden muss. Diese sind als Belege zu bevorzugen. Aber auch echte, also authentische Belege ohne definitorischen Charakter sind wertvoll. Dagegen tritt die Bedeutungsparaphrase als beschreibende Angabe etwas in den Hintergrund. Es ist zudem nicht leicht, ein gerade neu eingeführtes Wort mit der notwendigen Exaktheit zu definieren. Alle drei Neologismenressourcen verwenden Belege.

### 5.7 Frequenzangaben

Angaben zur Frequenz eines Wortes sind das Ergebnis einer länger währenden Beobachtung und können insofern auch erst nachträglich angebracht werden. Die Frequenz ist ein wichtiges Kriterium zu Unterscheidung von

- Okkasionalismen. Diese Wörter werden ein- bis dreimal und nur in einem sehr engen zeitlichen Rahmen vorkommen.
- Lexikalisierten Neologismen. Diese Wörter werden über einen längeren Zeitraum häufiger und, was die Quelle angeht, breiter gestreut vorkommen.
- Trend- und Modewörtern. Diese Wörter werden über einen gewissen Zeitraum hinweg häufig verwendet, die Frequenz nimmt aber nach einer gewissen Zeit deutlich ab, das Wort verschwindet schließlich ganz aus dem Sprachgebrauch.

Dies festzustellen bedarf es aber wie gesagt einer länger andauernden Beobachtung. Das Leipziger Neologismenwörterbuch (Quasthoff 2007) sowie die Online-Version des Neologismenwörterbuchs des IDS verzeichnen solche Daten, wobei sie Kalenderjahre als Fenster für die Darstellung der Häufigkeiten wählen.

## 5.8 Publikation

Die Vorteile der Publikation von Neologismen-Sammlungen im World Wide Web liegen so deutlich auf der Hand, dass man heute deren Publikation in Buchform als veraltet bezeichnen muss. Gut gemachte Neologismenwörterbücher haben aber einen dokumentarischen Wert und gelegentlich auch einen Unterhaltungswert und sind sicher zum Teil auch kommerziell erfolgreich. Die Benutzung einer Printpublikation dürfte dabei eher der eines Romans oder Erzählungsbandes entsprechen als der eines Nachschlagewerks.

Die Vorteile der Online-Publikation kann vor allem die tagesaktuelle Neologismenlexikographie nutzen:

- Wenn ein gutes Publikationssystem vorhanden ist, können neue Daten zeitnah in das Wörterbuch integriert werden. Das Wörterbuch ist somit immer auf dem aktuellen Stand, was natürlich vor allem bei Neuwörtern wichtig ist.
- Die Datenbestände können auf Grund neuer Erkenntnisse (z.B. zu schwankendem Genus) jederzeit aktualisiert werden, wobei allerdings der Aufwand einer solchen Aufgabe nicht unterschätzt werden sollte.
- Es können verschiedene Sichten auf die Daten ermöglicht werden, die neben eine alphabetische Auflistung treten, z.B. eine Filterung nach Sachgebieten, nach bestimmten Wortbildungstypen und Wortbildungselementen, nach dem Datum der Aufnahme in das Wörterbuch oder nach der Häufigkeit. Die Informationen, auf die diese Filter zugreifen, müssen natürlich vorhanden sein.
- Die Frequenzzählungen können regelmäßig aktualisiert werden.
- Schließlich können Verknüpfungen mit Suchmaschinen, z.B. Google, auf Articlebene gesetzt werden. Durch Anklicken des Links können sofort weitere Belege für das Stichwort angezeigt werden. Auf die Vernetzung mit anderen Ressourcen werden wir im nächsten Abschnitt noch näher eingehen.

## 5.9 Vernetzung mehrerer Ressourcen

Mittlerweile sind mehrere Neuwortsammlungen als Ressourcen verfügbar: das Neologismenwörterbuch des Instituts für Deutsche Sprache unter der Plattform OWID ([www.owid.de](http://www.owid.de)), die Wortwarte an der Berlin-Brandenburgischen Akademie der Wissenschaften ([www.wortwarte.de](http://www.wortwarte.de)) und die Wörter des Tages beim Projekt deutscher Wortschatz an der Universität Leipzig (<http://wortschatz.uni-leipzig.de/wort-des-tages/>). Eine Vernetzung dieser und ähnlicher, in Zukunft entstehender Ressourcen bietet sich an. Einige Vernetzungen dieser Ressourcen mit anderen Ressourcen wurden bereits realisiert. Hierzu gehören:

- Vernetzung einer Ressource mit einer anderen lexikalischen Ressource, z.B. Wikipedia (bei der Wortwarte).
- Vernetzung einer lexikalischen Ressource mit einem oder mehreren Korpora, über die händische Auswahl von Belegen hinaus (Wortwarte, über den Link zur Google-Suche; Wörter des Tages).

- Äußere Integration der Stichwortbestände mehrerer Neologismensammlungen in einer Wörterbuchsammlung<sup>12</sup> (das Neologismenwörterbuch am Institut für Deutsche Sprache ist in das Online-Wortschatz-Informationssystem integriert, zusammen mit zurzeit drei weiteren Wörterbüchern).

Darüber hinaus wären die folgenden Vernetzungsoptionen realisierbar und wünschenswert:

- Äußere Integration der Ressourcen über eine gemeinsame Lemmaliste und Verweis auf die Einträge bei völliger Selbständigkeit der beteiligten Ressourcen;
- Innere Integration der Neologismensammlungen mit anderen Wörterbüchern in ein Wörterbuchnetz über ein gemeinsames Datenmodell für die Mikrostruktur der Artikel aller Wörterbücher. Dies würde den Ansatz des IDS verallgemeinern und Ressourcen mehrerer Standorte einbeziehen.

Eine minimale Voraussetzung für diese Arten der Vernetzung ist es, dass die Artikel aller Ressourcen eindeutig und dauerhaft referenzierbar sind. Dies ist zurzeit noch komplizierter als es klingen mag.

## 6 Fazit

Die Vorteile der Nutzung des Internets sind im Falle der Neologismenlexikographie besonders evident. Dabei profitiert die aktuelle Neologismenlexikographie deutlich mehr von der Nutzung dieses Mediums als die retrospektive Neologismenlexikographie.

Das Internet verbessert die Situation der Neologismenlexikographie in nahezu allen Stadien des lexikographischen Prozesses: Durch das Herunterladen tagesaktueller Daten kann ein Monitorkorpus entstehen und ausgewertet werden; daraus ergeben sich die zu beschreibenden Stichwörter. Die Beschreibung kann auf Belege aus dem Monitorkorpus aufbauen. Durch Langzeitbeobachtung können Schwankungen in der Gebrauchsnorm neuer Wörter dokumentiert werden. Schließlich lässt sich über Frequenzprofile Verlässliches über die Ausbreitung eines neuen Wortes sagen.

Aber nicht nur als Datenquelle ist das Internet wichtig, sondern auch als Publikationsmedium. Mit einem geeigneten Publikationssystem lässt sich der lexikalische Wandel zeitnah darstellen. Das Printmedium verliert hier an Bedeutung, ohne jedoch ganz zu verschwinden.

Es ist allerdings sicher noch Potenzial für die weitere Nutzung von Webdaten vorhanden. Der wichtigste Punkt ist die Fähigkeit, neben Neulexemen auch Neubedeutungen zumindest halbautomatisch zu erkennen. Hier darf man auf die zukünftigen Errungenschaften von Sprachtechnologie und Computerlinguistik gespannt sein.

Auch was die Vernetzung von Ressourcen angeht, könnte man sich noch mehr wünschen: die stärkere Vernetzung der verschiedenen Neologismensammlungen untereinander ist hier ein wichtiges Desiderat. Hierbei sind allerdings neben den technischen auch organisatorische Hürden zu überwinden.

---

<sup>12</sup> Terminologie nach Engelberg/Lemnitzer (2009), Kap. 2.6, S. 73ff.

## 7 Literatur

## Wörterbücher

- Algeo 1941/1991 = Algeo, John: Fifty Years Among the New Words: A Dictionary of Neologisms, 1941–1991 (eine regelmäßige Kolumne in 'American Speech').
- The Barnhart Dictionary of New English Since 1963. New York: Barnhart, Harper & Row, 1973 (Fortführung 1980).
- Bathen, Dirk / Sporer, Josefine /Deinert, Eva / Heiss, Martin: Duden – Das neue Wörterbuch der Szenensprachen. Mannheim: Verlag Bibliographisches Institut, 2009 (Website: <http://szenesprachenwiki.de>).
- Burchfield 1972/1986 = Burchfield, Robert W.: A Supplement to the Oxford English Dictionary. 4 vol., Oxford: Clarendon Press, 1972–1986.
- Harlass 1974 = Harlass, Gertrude / Vater, Heinz: Zum aktuellen deutschen Wortschatz. Tübingen: GNV, 1974.
- Heberth 1977 = Heberth, Alfred: Neue Wörter. Neologismen in der deutschen Sprache seit 1945. Wien: Verl. d. Wiss., 1977.
- Herberg et al. 2004 = Herberg, Dieter / Kinne, Michael / Steffens, Doris: Neuer Wortschatz. Neologismen der 90er Jahre im Deutschen. Berlin, New York: de Gruyter, 2004 (online im lexikalischen Portal des IDS (OWID) verfügbar: [www.owid.de](http://www.owid.de)).
- Hellwig 1972/1983 = Hellwig, Gerhard : Kennen Sie die neuesten Wörter? München: Humboldt-Taschenbuchverlag, 1972/1983.
- Horx 1994 = Horx, Matthias: Trendwörter. Von Acid bis Zippies. Düsseldorf et al.: ECON-Verl., 1994.
- Lemnitzer 2000–2010 = Lemnitzer, Lothar: Die Wortwarte. Wörter für heute und morgen. Tübingen/Berlin, 2000–2010. ([www.wortwarte.de](http://www.wortwarte.de)).
- Lemnitzer 2007 = Lemnitzer, Lothar: Von Aldianer bis Zauselquote. Tübingen: GNV, 2007.
- Lemnitzer 2008 = Lemnitzer, Lothar: Hirndiebstahl im Sparadies. Was so (noch) nicht im Duden steht. Mannheim: Verlag Bibliographisches Institut, 2008.
- Quasthoff 2007 = Quasthoff, Uwe: Deutsches Neologismenwörterbuch. Neue Wörter und Wortbedeutungen in der Gegenwartssprache. Berlin, New York: de Gruyter, 2007.
- sz-Magazin (Hg.) 2006 = sz-Magazin (Hg.): Sprechen Sie Gegenwart? Lexikon des frühen 21. Jahrhunderts. München: Goldmann Verlag, 2006.

## Sonstige Literatur

- Barnhart/Barnhart 1989 = Barnhart, Robert / Barnhart, Clarence : The Dictionary of Neologisms. In: Hausmann / Reichmann / Wiegand / Zgusta. Wörterbücher. Dictionnaires. HSK Band 5, 1. Teilband. Berlin et al. 1989, S. 1159–1166.
- Engelberg/Lemnitzer 2009 = Engelberg, Stefan / Lemnitzer Lothar : Lexikographie und Wörterbuchbenutzung. 4. überarbeitete Auflage. Tübingen: Stauffenburg, 2009.
- Evert 2008 = Evert, Stefan : A lightweight tool for cleaning web pages. In: Proc. LREC 2008, Marrakech.
- Hohenhaus 1996 = Hohenhaus, Peter : Ad-hoc-Wortbildung. Terminologie, Typologie und Theorie kreativer Wortbildung im Englischen. Frankfurt: Peter Lang, 1996.
- Kinne 1996 = Kinne, Michael: Neologismus und Neologismenlexikographie im Deutschen. Zur Forschungsgeschichte und zur Terminologie, über Vorbilder und Aufgaben. In: Deutsche Sprache 24, H. 4, 1996, S. 327–358.
- Lemnitzer 2006 = Lemnitzer, Lothar / Zinsmeister, Heike : Korpuslinguistik. Tübingen: GNV, 2006.
- Peschel 2002 = Peschel, Corinna: Zum Zusammenhang von Wortneubildung und Textkonstitution. Tübingen: Niemeyer (= Reihe Germanistische Linguistik 237), 2002.
- Ruge 2004 = Ruge, Nikolaus: Das Suffixoid ‚-technisch‘ in der Wortbildung der deutschen Gegenwartssprache. In: Muttersprache 114, H. 1, 2004, 29–41.
- Scholze-Stubenrecht 2002 = Scholze-Stubenrecht, Werner: „Die Auswahl der Einträge ist äußerst beliebig“. Warum Jagdherr und Pokémon nicht im Duden stehen. In: Sprachwissenschaft 27, 2002, 225–248.

- Steffen 2002 = Steffens, Doris : Nicht nur Anglizismen ... Neue Wörter und Wendungen in unserem Wortschatz. In: Sprachreport. Informationen und Meinungen zur deutschen Sprache. 19, H. 4, 2002, 2–9.
- Teubert 1998 = Teubert, Wolfgang: Neologie und Korpus. Tübingen: GNV, 1998.
- Tomášiková 2008 = Tomášiková, Slavomira: Okkasionalismen in den deutschen Medien. In: Bočák, Michal / Rusnák, Juraj (Hg.): Média a text II. Prešove. 2008, 246–256.
- Wiegand 1990 = Wiegand, Herbert Ernst: Die deutsche Lexikographie der Gegenwart. In: Hausmann / Reichmann / Wiegand / Zgusta. Wörterbücher. Dictionaries. Dictionnaires. HSK Band 5, 2. Teilband. Berlin et al., 1990, 2100–2246. Zu Neologismenwörterbüchern: 2185–2187.
- Wotjak 1992 = Wotjak, Barbara: Phraseolexeme in System und Text. Tübingen: Niemeyer, 1992.