

## Online publizierte Arbeiten zur Linguistik

Im Auftrag des Instituts für Deutsche Sprache  
herausgegeben von Hardarik Blühdorn, Mechthild Elstermann und Annette Klosa

---

Andrea Abel/Annette Klosa (Hg.)

Der Nutzerbeitrag im Wörterbuchprozess.  
3. Arbeitsbericht des wissenschaftlichen  
Netzwerks „Internetlexikografie“



Institut für Deutsche Sprache  
Postfach 10 16 21  
68016 Mannheim  
[opal@ids-mannheim.de](mailto:opal@ids-mannheim.de)

Mitglied der Leibniz-Gemeinschaft



© 2014 IDS Mannheim – Alle Rechte vorbehalten

Das Werk einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechts ist ohne Zustimmung der Copyright-Inhaber unzulässig und strafbar. Das zulässige Zitieren kleinerer Teile in einem eigenen selbstständigen Werk (§ 51 UrhG) erfordert stets die Angabe der Quelle (§ 63 UrhG) in einer geeigneten Form (§ 13 UrhG). Eine Verletzung des Urheberrechts kann Rechtsfolgen nach sich ziehen (§ 97 UrhG). Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Die zugänglichen Daten dürfen von den Nutzern also nur zu rein wissenschaftlichen Zwecken genutzt werden. Eine darüber hinausgehende Nutzung, gleich welcher Art, oder die Verarbeitung und Bearbeitung dieser Daten mit dem Zweck, sie anschließend selbst oder durch Dritte kommerziell zu nutzen, bedarf einer besonderen Genehmigung des IDS (Lizenz). Es ist nicht gestattet, Kopien der Textdateien auf externen Webservern zur Verfügung zu stellen oder Dritten auf sonstigem Wege zugänglich zu machen. Bei der Veröffentlichung von Forschungsergebnissen, in denen OPAL-Publikationen zitiert werden, bitten die Autoren und Herausgeber um eine entsprechende kollegiale Information an [opal@ids-mannheim.de](mailto:opal@ids-mannheim.de).

## Inhalt

<i>Andrea Abel/Annette Klosa</i>	
Einleitung: „Ihr Beitrag bitte! – Der Nutzerbeitrag im Wörterbuchprozess“ .....	3
<i>Robert Lew</i>	
User-generated content (UGC) in English online dictionaries .....	8
<i>Luca Melchior</i>	
Ansätze zu einer halbkollaborativen Lexikographie.....	27
<i>Karin Rautmann</i>	
Duden online und seine Nutzer .....	49
<i>Katrin Thier</i>	
Das Oxford English Dictionary und seine Nutzer .....	63

## Einleitung: „Ihr Beitrag bitte! – Der Nutzerbeitrag im Wörterbuchprozess“

*Andrea Abel, EURAC Bozen/Bolzano, und Annette Klosa, Institut für Deutsche Sprache, Mannheim*

Unter dem Titel „Ihr Beitrag bitte! – Der Nutzerbeitrag im Wörterbuchprozess“ fand vom 18. bis 21. September 2012 ein Symposium im Rahmen des GAL-Kongresses 2012 „Wissen – Wörter – Wörterbücher“ in Erlangen statt, das sich in Vorträgen und einer Abschlussdiskussion mit der Frage beschäftigte, welchen Beitrag Wörterbuchbenutzer und -benutzerinnen für Internettörterbücher leisten, leisten können bzw. leisten sollten.<sup>1</sup> Mit dem Medium Internet eröffnet sich nämlich für Wörterbücher ein breites Spektrum neuer Möglichkeiten der Nutzereinbindung:

- Das Wörterbuch wird gemeinsam von Benutzern erarbeitet (Stichworte „Bottom-up-Lexikographie“, „nutzergenerierter Inhalt“).
- Die Wörterbuchbenutzer werden aufgefordert, Fehler zu melden und/oder Vorschläge für bestehende Einträge wie Neueinträge zu machen.
- Die Wörterbuchbenutzer werden über Newsletter, Blogs, Twittermeldungen o.Ä. über Entwicklungen am Wörterbuch informiert und hierdurch ans Wörterbuch gebunden.
- Funktionen wie „Häufig gesuchte Wörter“, „Mein Wort“ oder „Meine Suche“, „personalisieren“ das Wörterbuch und tragen ebenfalls dazu bei, den Benutzer stärker zu binden.
- Die Redaktion beantwortet Fragen der Benutzer zum Wörterbuchgegenstand.
- Die Wörterbuchbenutzer beraten sich gegenseitig oder bewerten die Inhalte anderer Beiträge.

Beim Symposium wurden die genannten Möglichkeiten anhand praktischer Beispiele vorgestellt, aber auch kritisch hinterfragt. Die Beispiele umfassten eine breite Palette lexikographischer Angebote (z.B. *OED*, *Duden*, *Wiktionary*) und reichten von der Verlags- und akademischen bis hin zur kollaborativen Lexikographie. In den Beiträgen wurde – kurz zusammengefasst – diskutiert, wie sich nutzergenerierte Wörterbücher zu redaktionell betreuten lexikographischen Produkten verhalten und welche Maßnahmen der Qualitätssicherung in nutzergenerierten Wörterbuchprojekten getroffen werden sollten. Auswirkungen auf den lexikographischen Prozess, die Darstellung und die Versionierung der Wörterbuchinhalte wurden angesprochen. Es wurde außerdem überlegt, wie man die Menge an Nutzerrückmeldungen an ein Onlinewörterbuch vergrößern oder wie man sie effektiver auswerten könnte. Schließlich wurde darüber reflektiert, ob die Fortentwicklung der Wörterbuchprojekte in praktischer wie theoretischer Sicht (soweit sie diesen Anspruch haben) durch die Rückmeldungen der Benutzer befördert wird.

Die vorliegende Onlinepublikation enthält nun vier Beiträge des Erlanger Symposiums (von Karin Rautmann, Katrin Thier, Luca Melchior und Robert Lew), die einen guten Überblick über das Thema geben, indem sie ein breites Spektrum an Möglichkeiten für Nutzer, sich am

<sup>1</sup> Die Idee zu diesem Symposium entstand im Rahmen der Arbeiten des von der DFG geförderten wissenschaftlichen Netzwerks „Internetlexikografie“ (vgl. [www.internetlexikografie.de](http://www.internetlexikografie.de)), das sich generell den speziellen Bedingungen der Erarbeitung von Internettörterbüchern widmet.

Auf- und Ausbau lexikografischer Angebote im Internet zu beteiligen, aufzeigen. In der Einleitung erfolgt außerdem eine Zusammenfassung der wichtigsten in den Beiträgen, Vorträgen und Diskussionen angesprochenen Punkte.

In den Aufsätzen von **Karin Rautmann** („*Duden online* und seine Nutzer“) und **Katrin Thier** („Das *Oxford English Dictionary* und seine Nutzer“), mit Beispielen aus der Verlagslexikographie, steht im Vordergrund, in wie vielfältiger Weise die Webversionen traditionsreicher Printwörterbücher die Mitwirkung der Wörterbuchbenutzer einfordern oder die Wörterbuchbenutzer durch Informationen zum Wörterbuch bzw. dem Wörterbuchgegenstandsbereich an dieses binden. Aufrufe an die Wörterbuchbenutzer zur Mitwirkung haben beim OED eine lange Tradition, die weit in die vorelektronische Zeit zurückreicht, und führen auch heute noch, inzwischen im Wesentlichen durch Nutzung der neuen Medien geprägt, zu überzeugenden Ergebnissen, wie Katrin Thier anhand von Beispielen in ihrem Beitrag vor Augen führt. Karin Rautmann legt in ihrem Beitrag einen Schwerpunkt auf die mithilfe von Marktforschung und einer Onlinebefragung gestützte Entwicklung von *Duden online* und illustriert, wie Nutzerfeedback zur kontinuierlichen Verbesserung des Angebotes eingesetzt wird.

Sowohl beim OED wie auch bei *Duden online* generieren die Wörterbuchbenutzer allerdings nicht selbst lexikographische Inhalte, sondern sie liefern Vorschläge, die alle einen redaktionellen Prozess durchlaufen, und nur die Wörterbuchredaktion kann Inhalte ergänzen oder verändern. Dies stellt sich bei einem sogenannten halbkollaborativen Wörterbuch wie LEO etwas anders dar, wie **Luca Melchior** in seinem Aufsatz über „Ansätze zu einer halbkollaborativen Lexikographie“ ausführt, mit dem der Bogen weg von der Verlags- und hin zu Formen der kollaborativen Lexikographie gespannt wird, bei denen schließlich der gesamte lexikographische Prozess in den Händen der Nutzer liegt. Am Beispiel der bilingualen Wörterbücher von LEO zeigt er, wie die Wörterbuchbenutzer in die Erarbeitung der Substanzen involviert werden und zugleich die lexikographische Qualität garantiert werden kann. Dabei wird deutlich, dass die Nutzer im Wesentlichen Wörterbuchaußentexte wie das Forum ohne redaktionelle Beteiligung gestalten können, während die Er- und Bearbeitung der Makro- und Mikrostrukturen allein der Wörterbuchredaktion gestattet ist, die ihrerseits beträchtlich von den Nutzerbeiträgen profitieren kann. Im Unterschied dazu erstellen in kollaborativen Wörterbuchprojekten wie Wiktionary die Wörterbuchbenutzer selbst den gesamten Wörterbuchbestand.

**Doris Leibold** (Sauerlach) berichtete in ihrem Vortrag ebenfalls über LEO. Sie stellte „Die LEO-Wörterbücher und ihre Benutzer“ vor und legte dabei u.a. Wert auf die Feststellung, dass am LEO-Angebot 10 Techniker und etwa 15 Redakteure (darunter allerdings einige in Teilzeit) arbeiten; Technik und Inhalt würden also Hand in Hand gehen. Sie nannte als Nutzerbeitrag bei LEO vor allem die Schenkungen z.B. von umfangreichen Wortlisten mit Übersetzungen. Daneben erwähnte sie, dass Wörterbuchbenutzer auch Meldungen zu gewünschten Nachträgen oder Ergänzungen machen, was natürlich den redaktionellen Aufwand erhöhe. So betonte Doris Leibold sowohl die Vor-, aber auch die Nachteile des halbkollaborativen Ausbaus des LEO-Wörterbuchangebotes.

Vorteile eines halbkollaborativen Ansatzes können auch für andere Internettörterbücher genutzt werden. Dies wird deutlich, wenn man sich die Beispiele für von Nutzern geschaffene Inhalte in **Robert Lews** Aufsatz zu „User-generated content (UGC) in online English dictionaries“ ansieht. Lew zeigt auf, dass auf die Expertise von Lexikographen nicht verzichtet werden kann, wenn qualitätvolle Internettörterbücher entstehen sollen. So gebe es einen großen Qualitätsunterschied zwischen redaktionell erarbeiteten Wörterbüchern und Beispielen aus

der wirklichen Bottom-up-Lexikographie wie Wiktionary, dessen Inhalte nicht selten durch Copy-and-Paste-Verfahren aus älteren und daher copyrightfreien Quellen übernommen seien. Dies liege unter anderem daran, dass eben nicht jeder selbsternannte Sprachexperte auch tatsächlich über lexikographische Fähigkeiten verfüge (außer möglicherweise in einem streng fachsprachlichen Kontext).

Unter dem Titel „Kollaborative Lexikographie und Nutzerbeteiligung in Wiktionary“ berichteten **Iryna Gurevych** und **Christian M. Meyer** (beide Technische Universität Darmstadt) darüber, wie ein sprachbezogenes Nachschlagewerk im Internet ohne lexikographische Expertise und Beteiligung entsteht. Hier werden alle Einträge von Nutzern verfasst, überarbeitet und diskutiert, wobei es unterschiedliche Typen von Beitragenden gebe. Vom Administrator bis hin zum passiven Leser seien alle vertreten, doch nicht alle Nutzergruppen könnten – so die Vortragenden – vom Wiktionary-Angebot in gleicher Weise profitieren. Da beispielsweise nur gut die Hälfte der Wörterbucheinträge Referenzen verzeichnen und nur etwa ein Viertel Quellenangaben enthält, ist die Nutzbarkeit etwa im wissenschaftlichen Kontext anzuzweifeln.

**Andrea Rapp** (Technische Universität Darmstadt), **Luise Borek** (Technische Universität Darmstadt) und **Vera Hildenbrandt** (Trierer Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften / Trier Center for Digital Humanities) hatten unter dem Titel „Abgeschlossen und doch offen – Nutzerbeteiligung bei retrodigitalisierten Printwörterbüchern“ darüber nachgedacht, wie Wörterbuchbenutzer in den lexikographischen Prozess, der mit der Digitalisierung von Printwörterbüchern, d.h. ihrer Erschließung und Präsentation für das Internet, einsetzt, in sinnvoller Weise einbezogen werden können. Damit wurde ein Bereich der Lexikographie, nämlich der der historischen Lexikographie, charakterisiert durch den Typus des Abschlusswörterbuchs, in den Symposiumskontext einbezogen, den man zunächst kaum mit Möglichkeiten der Nutzerbeteiligung in Zusammenhang bringen würde. So könnten die Nutzer von retrodigitalisierten Internetwörterbüchern beispielsweise Hinweise auf Satz- und Tippfehler geben, aber auch zur Benutzeroberfläche Rückmeldung geben. Wesentlich wünschenswerter wäre, wenn man etwa an die im Trierer Wörterbuchnetz ([www.woerterbuchnetz.de](http://www.woerterbuchnetz.de)) verbundenen Substanzen denkt, dass nutzergenerierte Verweise zwischen Stichwörtern in den einzelnen Wörterbüchern ermöglicht werden. Die Wörterbuchbenutzer könnten sich auch daran beteiligen, die bestehenden Verweise zwischen Wortartikeln zu verifizieren und ggf. zu revidieren, allerdings fehlten hierzu noch die technischen Möglichkeiten, die dies erlauben würden.

**Jörg Schröder** (Berlin) stellte mit einem weiteren Beitrag aus dem Bereich der historischen Lexikographie, ausgehend von den Erfahrungen bei der Neubearbeitung des „Deutschen Wörterbuchs“ von Jacob und Wilhelm Grimm in der Arbeitsstelle an der Berlin-Brandenburgischen Akademie der Wissenschaften, in seinem Vortrag die Frage: „Ihr Beitrag bitte – aber was, wenn er dann wirklich kommt?!“ Er verwies darauf, dass Belege oder Beispiele für den Wortgebrauch, die von Wörterbuchbenutzern eingesandt werden, redaktionell gesichtet und verifiziert werden müssten, was unter Umständen einen erheblichen zeitlichen Aufwand bedeute. So sei abzuwägen zwischen dem Gewinn, den das Wörterbuchprojekt durch solche Beiträge machen könne, und dem Einsatz von Arbeitskraft und Zeit, die an anderer Stelle (z.B. dem Erarbeiten neuer Wortartikel) fehlen könnten.

**Magali Paquot** und **Marie-Aude Lefer** (beide Louvain-la-Neuve) berichteten in ihrem Vortrag aus dem Gebiet der akademischen Lexikographie zu „Fuzzy feedback in the compilation of the UCL-K.U.Leuven University Terminology Database“ davon, welche Erfahrungen sie mit ihrem Onlineangebot zur universitären Terminologie gemacht haben, das

sich mit der gleichen Benutzeroberfläche sowohl an die Nachschlagenden wie die Bearbeiter wendet. Nach Veröffentlichung der terminologischen Datenbank wurde ein Jahr lang das Nachschlageverhalten systematisch ausgewertet (implizit durch Logfileanalysen, explizit durch Auswertung der direkten Nutzerrückmeldungen), im Anschluss wurde die Webseite systematisch ergänzt. Dabei stellte sich heraus, dass die direkten Nutzermeldungen für die Verbesserung des Angebotes besonders hilfreich waren, in Kombination mit der Auswertung der Logfiles wurde darüber hinaus z.B. deutlich, dass als Einträge vor allem Mehrwortausdrücke ergänzt werden mussten. Als Fazit zu diesem Beitrag ist klar festzuhalten, dass das Angebot ohne das Feedback durch die Wörterbuchbenutzer nicht die Qualität hätte erreichen können, die es heute hat.

**Stephan Bopp** (Basel) lieferte mit seinem Vortrag „Der Sprach-Blog ‘Fragen Sie Dr. Bopp!’ in canoo.net“ einen Erfahrungsbericht dazu ab, wie man über Sprache bloggen und so sehr direkt auf Nutzeranfragen reagieren kann. Nutzerfragen zur deutschen Sprache werden in canoo.net in einer Art modernem „Briefkastenonkel-Dienst“ beantwortet, im Sprach-Blog werden darüber hinaus die Antworten, die für einen größeren Adressatenkreis interessant sein könnten, publiziert. Viele Nutzeranfragen entstünden dabei aus der Situation heraus, dass die Nachschlagenden die gesuchte Information in canoo.net nicht finden können oder dass diese Information tatsächlich fehlt. Es ist nicht klar, ob und inwiefern die Reaktionen in Echtzeit im Rahmen des Blogs auf die Inhalte von canoo.net zurückwirken. So verließ das Symposium in diesem Beitrag schließlich den Bereich, der zum eigentlichen lexikographischen Prozess gerechnet wird. Die Antworten und Einträge im Sprach-Blog trügen laut Vortragendem aber auf jeden Fall dazu bei, die Benutzbarkeit des Nachschlagewerkes insgesamt zu erhöhen. Darüber hinaus ist ein Einfluss auf die Nutzerbindung allemal wahrscheinlich.

In der von **Angelika Storrer** (Technische Universität Dortmund) moderierten Abschlussdiskussion diskutierten **Carolin Müller-Spitzer** (Institut für Deutsche Sprache, Mannheim), **Stephan Bopp** (canoo.net, Basel), **Michael Mann** (Universität Erlangen), **Christian M. Meyer** (Technische Universität Darmstadt) und **Eveline Wandl-Vogt** (Wörterbuch des Bairischen in Österreich, Österreichische Akademie der Wissenschaften Wien) über zahlreiche weitere Aspekte des Tagungsthemas, wobei sich die Diskussion rasch ins Publikum, in dem etliche weitere Wörterbuchprojekte vertreten waren (z.B. Deutsches Rechtswörterbuch, Oxford English Dictionary, *Duden online*, LEO-Wörterbücher), öffnete. Im Folgenden werden einige Thesen der Diskussion aufgegriffen, woraus sich auch Forderungen an die Wörterbuchbenutzungsforschung sowie die metalexikographische Beschäftigung mit dem Nutzerbeitrag bei Onlinewörterbüchern ableiten lassen.

Generell sollte im Kontext des Einbezugs des Wörterbuchbenutzers in Ausbau und Verbesserung des Wörterbuchs weniger von Kosten bzw. Nutzen als vielmehr von Aufwand bzw. Bereicherung gesprochen werden. Durch den Nutzerbeitrag haben Wörterbuchredaktionen nämlich die große Chance, vollständigeres Material zu gewinnen. Insbesondere Dialektale oder Kontexte in bestimmten Soziolakten können hierdurch, besonders aber auch in der kollaborativen und halbkollaborativen Lexikographie, besser abgedeckt werden. Die Rückmeldungen können auch genutzt werden, um beispielsweise die Recherchemöglichkeiten zu optimieren.

Lexikographen nehmen bei Internettörterbüchern (insbesondere in halbkollaborativer Lexikographie) verstärkt die Rolle von Organisatoren ein. Das Gesamtkonzept des Wörterbuchs sowie die Festlegung der Lemmaform bleiben hier Aufgabe des Lexikographen. Was die inhaltliche Beschreibung der Lexeme anbelangt, so waren Lexikographen insbesondere im fachsprachlichen Bereich schon immer auf die Mithilfe von Fachleuten angewiesen. Hieran wird sich auch zukünftig nichts ändern. Im Bereich der Quellenbeschaffung und -auswertung

können Lexikographen nun aber verstärkt Aufgaben an die Wörterbuchbenutzer abgeben. Im Kontakt mit dem Wörterbuchbenutzer bzw. Laienbeiträger bleibt der Lexikograph in der Rolle des Moderators, der den Überblick über das ganze Wörterbuchprojekt behält und für Vollständigkeit und Kontinuität sorgt. Wie wichtig die Rolle des Lexikographen bleiben wird, zeigt sich auch darin, dass die Beiträger zum Wiktionary viel eher auf die lexikographischen Traditionen zurückgreifen (z.B. was den Lemmaansatz, das Angabeinventar oder die Form der Bedeutungsangaben anbelangt) und konservative Ansätze überliefern helfen, als neue zu entwickeln (z.B. auf der Grundlage von Wörterbuchbenutzungsforschung).

Die Relevanz des Lexikographen wurde im Zusammenhang mit kollaborativen Wörterbüchern wie Wiktionary auf unterschiedlichen Ebenen angesprochen: Einerseits führten Copy-Paste-Verfahren aus copyrightfreien Wörterbuchversionen mit teilweise veralteten Daten mitunter zu qualitativ minderwertigen Wiktionary-Einträgen. Andererseits zeige die Tendenz des Kopierens, das im Übrigen nicht immer aus veralteten Quellen geschehen muss, wie beispielsweise Quellenangaben in der deutschen Wiktionary-Version nahelegen (z.B. DWDS), dass selbst kollaborative Projekte ohne Rückgriff auf lexikographisch-redaktionell erarbeitete Produkte nicht möglich zu sein scheinen.

Großen Einfluss auf die Menge und Verwertbarkeit von Nutzerbeiträgen haben im Übrigen Platzierung und Gestaltung von Feedbackformularen auf den Wörterbuchseiten. Vor Freischaltung einer solchen Möglichkeit der Nutzerbeteiligung sollte diese deshalb sorgfältig geplant werden, wozu auch die Festlegung des Umgangs mit den Nutzerrückmeldungen gehört. Erfahrungen in Wörterbuchredaktionen zeigen, dass es sehr wichtig ist, den Wörterbuchbenutzern direkt Rückmeldung zu ihren Beiträgen zu geben und den Vorgang, wie damit redaktionell verfahren wird, transparent zu machen. Ansonsten wird die Bereitschaft von Laien, auf Lemmalücken hinzuweisen, Beleglücken schließen zu helfen usw. schnell zurückgehen.

Als Fazit des Symposiums kann festgehalten werden, dass Beiträge von Nutzern zu Onlinewörterbüchern in vielfältiger Form erbracht werden und häufig zur inhaltlichen Qualität der Nachschlagewerke beitragen. Voraussetzung hierfür ist allerdings, dass die inhaltliche Verantwortung bei den lexikographischen Erarbeitern bleibt, die zunehmend auch organisatorische Aufgaben im Kontakt mit den Nutzern übernehmen, sodass der lexikographische Prozess den neuen Gegebenheiten angepasst werden muss. Die Nutzung der vielfältigen Möglichkeiten, die das Medium Internet zur Verfügung stellt (Twitter, Feedbackformulare usw.), kann die Nutzer außerdem näher ans Wörterbuch binden. Dass akademische Lexikographie und Verlagslexikographie gegenüber Bottom-up-Lexikographie Bestand haben werden, ist vor allem dann wahrscheinlich, wenn sich wissenschaftliche Wörterbuchprojekte und Wörterbuchverlage verstärkt mit den Nachschlagenden auseinandersetzen und deren Expertise zur Verbesserung von Wörterbuchinhalten und -präsentation nutzen. Ob und inwiefern sich die Qualität von kollaborativ erarbeiteten Produkten am Rückgriff auf professionell erarbeitete Quellen messen lässt, werden künftige Studien zeigen müssen.

# User-generated content (UGC) in English online dictionaries

Robert Lew, Adam Mickiewicz University in Poznań

## 1. Introduction

User-generated content rests on two pillars: the philosophy of the role of the collective of ordinary people, and the technical capability to realize this philosophy in practice.

The philosophy behind UGC is actually anything but new. It is based on the belief in the *wisdom of crowds* (Surowiecki 2005), which in essence puts the collective opinion of a group of people above that of a single expert. A practical and important social application of this principle is *trial by jury*, which functions in numerous societies and has many advocates.

The technological underpinnings of the concept of user-generated content have to be sought in the rise of *Web 2.0*. This playful name (which draws on the convention of referring incrementally to an improved version of a software application) stands for the advanced interactive stage in internet experience, and it is the modern affordances that come with it which make user-generated content happen. Users of Web 2.0 are no longer passive recipients of packaged content. Increasingly, they actively contribute to the creation and provision of self-made content. This double capacity of newly empowered users can be aptly captured in the neologism *prosumer*, which is a blend of *producer* and *consumer*. In this way, web users' social roles become blurred. In the context of online lexicography, the new model puts dictionary users in the shoes of lexicographers.

Another concept relevant to UGC is *new media*. In opposition to traditional media such as newspapers or motion pictures, new media involve a two-way exchange of information: it is *interactive* or *conversational*. Another of its characteristics is the immediacy with which new media can be accessed, being available on demand, essentially at any time, and on a range of digital devices (PC's, netbooks, tablets, mobile phones, ...). It is generated in real time and largely unregulated. Unlike in traditional media, the creation, publishing, distribution, and consumption of new media involve a level of democratization.

## 2. What is User-Generated Content?

It is instructive to consider what UGC actually signifies. A workable characterization of the concept (albeit under a slightly different name of User-Created Content) is provided in the relevant OECD report (Vickery/Wunsch-Vincent 2007); this publication cites three broad characteristics that have to obtain in conjunction (although the authors are open to a degree of fuzziness in the third criterion):

**Publication requirement:** A principle [sic!] characteristic is that the work is *published* in some context, for example on a publicly accessible website or on a page on a social networking site only accessible to a select group of people (e.g. fellow university students), even though UCC could be made by a user and never published online or elsewhere. This characteristic excludes e-mail, two-way instant messages and the like.

**Creative effort:** A *certain amount of creative effort has to be put into creating the work or adapting existing works to construct a new one; i.e. users must add their own value to the work.* UCC could include user uploads of original photographs, thoughts expressed in a blog or a new music video. The creative effort behind UCC may also be collaborative, for example on websites that users edit collaboratively. Merely copying a portion of a television show and posting it on an online video website (a frequent activity on UCC sites) would not be considered UCC. Nevertheless the minimum amount of creative effort is hard to define and depends on the context.

**Creation outside of professional routines and practices:** User-created content is usually created *outside of professional routines and practices.* It often does not have an institutional or commercial market context and UCC may be produced by non-professionals without expectation of remuneration or profit. Motivating factors include: connecting with peers, achieving fame, notoriety or prestige, and expressing oneself. (Vickery/Wunsch-Vincent 2007, p. 18)

The *creative effort* stipulation becomes especially problematic in the context of lexicographic activity. This is so because dictionary-making has traditionally been an activity of an accretive, and often duplicative nature. For centuries, numerous dictionaries were written by copying and collating content from dictionaries already in existence, and a considerable degree of acceptance of this fact exists in the general public to this day, even among people directly involved in lexicography (Landau 2001, p. 403).

### 3. Motivation behind UGC

It is interesting to examine the motivation behind user-generated content. The gains of site hosts are rather obvious: being able to obtain potentially relevant content that is essentially cost-free, keeping the site alive with activity, and attracting users. But what exactly is it that drives web users to donate their time and (sometimes) expertise in order to create online content? It seems that at least three types of factors are at play as far as motivation for creating UGC is concerned: psychological, social, and economic.

#### 3.1 Psychological motivation

Many individuals find contributing content *psychologically satisfying.* There is the do-gooder feeling of benefiting others which comes under the general rubric of *altruism.* Another psychological aspect important for some is that contributing content can fulfil the individual's need for self-expression. Finally, being able to provide advice on language, as often happens in lexicographic resources, makes people feel like teachers, and there are individuals who derive enjoyment out of mentoring others.

#### 3.2 Social motivation

The second category of reasons which drive users to create UGC has to do with social aspects. One reason is the urge to be part of an online community, connecting and interacting with others: it is much the same motivation that is responsible for the current success of the phenomenon of social networking. Of course, in order for this aspect to be effective, the

service needs to have a reasonable-sized active user base. In other words, this type of social incentive requires *scale* in order to be effective. This is also the prerequisite for another reason in the social category. Quite a few individuals engage in content creation in hopes of becoming famous (or notorious, as the case may be). Top contributors of UGC often succeed in securing certain prestige in the community. Another reason for engaging in user-supported lexicographic projects is that many individuals are honestly and publicly enthusiastic about words, their meanings and uses, and matters of language usage. This *word enthusiast* factor provides an important social incentive for lexicographic UGC.

### 3.3 Economic motivation

Apart from psychological and social motivation, there may also be material rewards for supplying UGC. This may take the form of prizes, coupons for physical or virtual goods, access to premium features, etc.

## 4. UGC on the Web

Unlike in its early years, a significant section of today's World Wide Web relies on user-generated content. Web 2.0 is an interactive experience, with distinctions between consumers and producers of content becoming increasingly blurred.

With respect to specific websites, UGC can be exploited to varying degrees. This, of course, is a continuum, but for practical reasons it may be convenient to distinguish between at least three configurations with regard to the proportion of content originating from users:

1. content is entirely or almost entirely user-generated;
2. UGC is combined with professional (institutional) content; this option is sometimes termed half-collaborative (Storrer 2010 has *Internet-Wörterbücher mit Nutzerpartizipation*; Melchior 2012 has *halbkollaborative Wörterbücher*);
3. professional/institutional content dominates, but is supplemented or enriched with user input; the latter may take the form of reviews, comments, etc.

As we shall see below, these three models are all represented in online dictionaries. Before we examine such dictionaries, let us take a quick look at a selection of pre-lexicographic examples, so as to get some view of the models underlying UGC in online lexicography.

Web 2.0 has brought in an opportunity for people to share their (amateur) artistic output, and rate and comment on the work of other users. [Quizilla](#) (Figure 1) is a relatively early (2002) website with user-generated content, which started as a platform for exchanging quizzes, but was soon extended to allow for sharing amateur prose and poetry, thus providing an example of how a community-based site is shaped according to genuine needs of its members. One of the key features that is now so much part and parcel of Web 2.0 is the user-rating of content.

An urge to produce and share creative content is one that is present in many humans and it is websites like [Quizilla](#) that allow such content to be shared and discussed with netizens of like interests. One of the most popular services of this type, but dealing with visual rather than verbal art, is the photograph-sharing service [Flickr](#) (Figure 2).

The screenshot shows the Quizilla Poems Hub page. At the top, there's a navigation bar with links for QUizzes, STORIES, POLLS, POEMS, LYRICS, GAMES, MYSTUFF, and HELP. Below the navigation is a search bar with a 'SEARCH' button and a 'GO' button. On the left, there are three tabs: TOP RATED, MOST POPULAR, and NEWEST. A red banner at the top of the main content area says 'Freedom (From My Ready Rope)'. Below it, a message says 'Created by Angeltruth233 on 06/23/2012'. There's a 'READ MORE!' button with an arrow pointing right. To the right of the main content, there's a 'LOG IN' form with fields for 'username' and 'password', and buttons for 'Forgot Password?' and 'GO'. Below the log-in is a 'Remember Me' checkbox. Further down, there's a 'GOT AN IDEA? GET STARTED!' section with a dropdown menu set to 'Poem' and a 'title' input field with a 'GO' button. At the bottom right, there's a 'NEW TO QUIZILLA?' section with a link to 'Take a Personality Quiz'.

Figure 1: [Quizilla](#), an early UGC website for the sharing of prose and poems (originally quizzes)

The screenshot shows the Flickr homepage under the 'Explore / Most recent photos and videos' section. It displays a grid of 20 thumbnail images, each with a caption indicating the user who uploaded it. The users include Circles of..., kevjin, p11118, El Bingle, lsAurora, Eidenmark, Pattys-photos, edlaw01, kimilife, Actua, sieffert, Justin Porterfield, gwazulu, SVNews, tiredbees, and randylin27. To the right of the grid, there are two links: '» Popular' and '» Creative'.

Figure 2: The photo-sharing website [Flickr](#)

Of course, the use of UGC-based or UGC-enhanced websites does not stop at aesthetic appreciation. Just as many [Flickr](#) users value the art appreciation side of the service, for many others it is mainly about the practical utility of being able to conveniently share family photos from holidays, trips, or family events. One very practical application of the UGC-model is the air travel website [SeatGuru](#). The idea of this service is to assist potential travellers with flight seat selection. Its database accumulates data on plane seating quality for most of the world's airlines, listing specific models of aircraft in their fleets, including flight number information. Much of the actual detailed data come from the users themselves. Via feedback forms (Figure 3), travellers supply their comments on the quality of the seats they have occupied on their recent flights. Thanks to all this data, prospective passengers are in a position to make an informed choice of seating, particularly when checking in online. They can be warned of potential inconveniences, such as legroom restrictions due to media boxes, backrests failing to recline in front of bulkheads, or proximity to a lavatory. What this service has in common with lexicographic tools is that it serves specific consultation needs, in this particular case a type of needs which has sometimes been called *cognitive* (Tarp 2008; Ptaszyński 2009).

The screenshot shows a web page for the Lufthansa Airbus A319-100 NEK. On the left, there is a sidebar with a list of aircraft models. At the top right, there are icons for food and infant services. Below these are two sections: one for general amenity icons and another for seat-specific comments. The main area contains a feedback form with fields for flight number, comments, and seat preferences. It also includes a 'Seatmap Key' and a search bar for finding cheap flights.

General Information
Airbus A319-100 IIEK
Airbus A320-200 IIEK
Airbus A321 NEK
Airbus A330-300 (333)
Airbus A340-300 Vers. 1 (34V)
Airbus A340-300 Vers. 2 (34P)
Airbus A340-300 Vers. 3 (343)
Airbus A340-600 (346)
Airbus A380-800 (388)
Avro RJ85 Avroliner (AR8)
Boeing 737-300
Boeing 737-530 NEK
Boeing 747-400 (744) Version 1
Boeing 747-400 (744) Version 2
Boeing 747-400 (744) Version 3
Boeing 747-8i Intercontinental
Bombardier Q-400
Canadair CRJ-700 (CR7)
Canadair CRJ-900 (CR9)
Embraer ERJ-195 (E95)
Malaysia Airlines
MDLR
Mesa Airlines
Monarch Airlines
Oman Air
Openskies
Philippine Airlines
Porter Airlines
Airlines: O-R-S-T
Airlines: U-V-W

**Food** **Infant**

The above amenity icons represent the services that are available on this aircraft. Click on the individual icons for more detailed information. If an icon is not pictured, that service is not offered on this aircraft.

For seat specific comments on this aircraft, mouse over the plane illustration below. View our Seat Map Key below or visit the [FAQ page](#) for further assistance and explanations.

**SeatGuru.com Seat Feedback Form**

**SEATGURU** by **tripadvisor**

**Feedback for Lufthansa Airbus A319-100 NEK**

My flight number was:

My comments are for seat:

For the existing seat comments:  
 I agree  I disagree  I have something to add

Your Comments:

Please Enter your Email Address:

For other correspondence, see the FAQ link on the homepage.

**do you know this plane?**

**submit comments**

**Seatmap Key**

- Good Seat
- Some Drawbacks
- Poor Seat
- Standard Seat
- Blocked Seat
- Crew Seat
- Power Port
- Emergency Exit
- Overhead TV
- Galley
- Lavatory
- Closet

**Find Cheap Flights!**

**NEW!**

Flights Hotel Vacation Car Rental

From: Warsaw, Poland - Frederik  To:

Depart: Aug 4 Return: Aug 11 Travelers: 1

More options > one-way, multi-city, etc.

Find cheap flights at TripAdvisor.

14   
15   
16

Figure 3: Seat feedback form on [SeatGuru.com](#)

One important area of online content where internet users provide meaningful contributions is citizen journalism, in which non-professionals get involved in reporting news on websites such as Korea's [OhmyNews](#) (Figure 4). But many mainstream news services also try to involve their regular readers in collaborative journalism. For instance, [CNN](#) has *iReporters*: registered individuals who contribute reports, including video recordings. Many other professional services now include similar kinds of user involvement, and it is in fact quite rare today to have news services which would have no provision for reader comments at least.

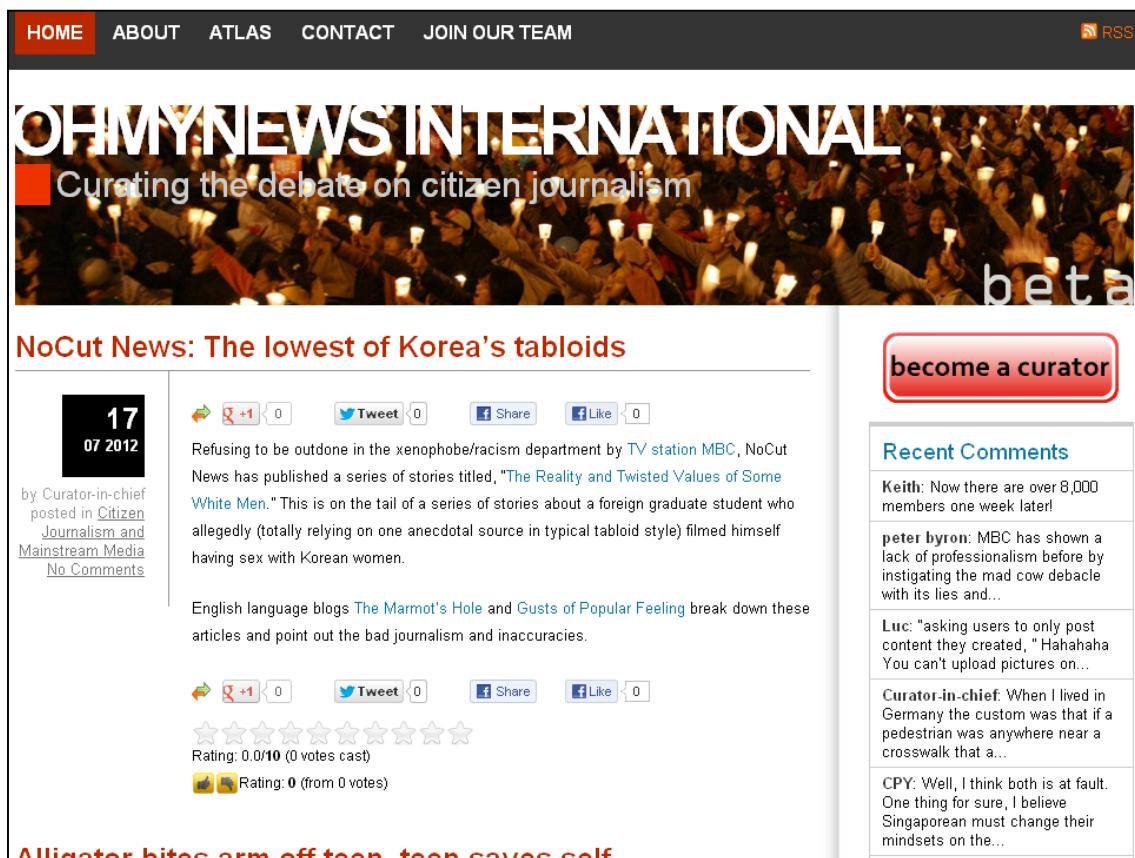


Figure 4: [OhmyNews](#), an example of citizen journalism

Film lovers have for some years enjoyed the extensive collaborative database [IMDb](#) (short for Internet Movie Database), where, subject to subsequent IMDb staff approval, registered users can contribute most of the pivotal data (film titles, appearances by actors, etc.), as well as provide supplementary content such as summaries and reviews, user ratings, or detailed content advisory for concerned parents.

## 5. UGC in English lexicography

### 5.1 [Wiktionary](#) as an example of a collaborative dictionary project

As we have seen above, user-generated content has spread to many domains of the Web, thus it should not be surprising that it has entered online lexicography as well. What may be thought of as an immediate precursor to crowdsourced dictionaries is a reference work with an encyclopaedic rather than lexical focus: the collaborative encyclopaedia [Wikipedia](#) (Figure 5). This extensive resource is a large-scale implementation of the idea of the *wisdom of crowds* (Surowiecki 2005). It relies on the assumption that in the online world, there will be experts on a specific topic willing to share their time and expertise pro bono with the world at large. There is little doubt that this model has turned out to be highly successful, at least in some areas. A peer-review based comparative assessment of science articles in [Wikipedia](#) and the highly respectable *Encyclopaedia Britannica* (Giles 2005) found an average of four errors in [Wikipedia](#) as compared to three in the latter, suggesting that the factual accuracy of the two resources was at a comparable level. Presumably, [Wikipedia](#) got better rather than worse since that time.



The screenshot shows the Wikipedia article page for "Lexicography". The page title is "Lexicography" and it is indicated as being from "Wikipedia, the free encyclopedia". Below the title, there is a note: "For the term in mathematics, see [Lexicographical order](#)". A yellow warning box states: "This article includes a [list of references](#), but its sources remain unclear because it has insufficient citations. ([May 2009](#))". The main content discusses the division of lexicography into practical and theoretical branches, and provides a definition of a lexicographer. It also notes the focus on design, compilation, use, and evaluation of general and specialized dictionaries. A sidebar contains a "Contents" section with links to etymology, aspects, recommended reading, and external links. At the bottom of the page, there is a link to the "Etymology" section.

Figure 5: An entry for *lexicography* in *Wikipedia*, the collaborative encyclopaedia

Initiated as a sister project of *Wikipedia*, *Wiktionary* came to life in 2002, initially as an English-language edition only, to be joined less than two years later by separate language editions of the then 143 active languages (Meyer/Gurevych 2012, p. 262). At the time of this writing (12 January 2013), the English edition of *Wiktionary* reports nearly three million entries, and three further editions stand at more than a million articles each. However, a non-trivial proportion of items are automatically generated inflectional forms; another reason behind the high numbers is the wholesale incorporation of entries from older out-of-copyright dictionaries. Even more importantly, only a fraction of the entries in the English edition are actual English words: a given language edition A uses language A as the primary language of explanation for its entries, but the headwords themselves are not restricted to language A. And so, for example, the English *Wiktionary* covers the English lexicon (which it explains with definitions in English), but also words from numerous other languages (which it explains with English translations and/or English definitions). Meyer and Gurevych (2012) stress that this aspect is poorly understood, even by some experts. As shown in Figure 6, the English Wiktionary entry for HANDLE covers the English word (at the top), but also the Danish, German, and Norwegian words so spelled in those languages.

*Wiktionary* is compiled through a collaborative process by a large community of Web users, in this context called *Wiktionarians*. Wiktionarians pattern into a pyramid-like structure with the top represented by (democratically elected) 98 *administrators* (Meyer/Gurevych 2012, p. 271), who hold the most extensive permission rights. *Registered users* are users holding personal accounts, and it is they who do the greatest proportion of the work. However, as in many collaborative enterprises, the distribution of the workload tends to be Zipfian, with only a very small minority of users taking care of most of the work. For example, the English *Wiktionary* has about 400,000 registered users, but only some 4,000 (a mere one per cent)

have at least ten edits to their names (Meyer/Gurevych 2012, p. 272). Unregistered users can also edit entries, but this happens rarely, with only about five per cent of all article edits being by unregistered users (Meyer/Gurevych 2012, p. 272).

The screenshot shows the English Wiktionary entry for the word "HANDLE". The page includes a sidebar with navigation links like Main Page, Preferences, and Toolbox, and a footer with links to other projects like Wikipedia and Wikipedia. The main content area features a table of contents for the word "HANDLE" in English, Danish, German, Norwegian, and Verb forms. Below the table of contents, there are sections for Pronunciation, Etymology, and Definition. A note indicates that the word has articles on Handle (grip) and Handle (mathematics). The page also includes standard wiki functionality like Edit and History buttons.

Figure 6: Top of the English *Wiktionary* entry for HANDLE

The collaborative aspect is supported in *Wiktionary* with the usual wiki-type functionality. Each article has a *talk page* for the exchange of comments. Differences of opinion are resolved through discussion, and, where consensus has not been reached, through voting. There are also pages in *Wiktionary* discussing more general matters of policy, such as criteria for inclusion of terms and senses in the dictionary. As far as I am aware, a systematic evaluation of the lexicographic quality of *Wiktionary* articles has not yet been attempted. Those assessments that are available tend to address aspects analyzable by computer, such as coverage of terms and senses, or sense ordering (Krizhanovsky 2011; Meyer/Gurevych 2012). Unfortunately, such automated evaluation, though fast and convenient, cannot at present even begin to address the essence of the quality of lexicographic description. Isolated comments about the latter have been made, however, and overall have been critical of *Wiktionary's* lexicographic quality (Fuentes-Olivera 2009; Hanks 2012). Hanks (2012, pp. 77-78) observes that “[i]n the English Wiktionary, the etymologies are taken from or based on those in older dictionaries, as are the definitions, which are extremely old-fashioned and derivative, taking no account of recent research in either cognitive linguistics or corpus linguistics”. This observation is certainly largely true, and a close comparison of a Wiktionary entry with one from a quality professional dictionary for the same headword can be a sobering experience to enthusiasts of crowd-sourced lexicographic description.

<p><b>handle</b> (<i>third-person singular simple present handles, present participle handling, simple past and past participle handled</i>)</p> <ol style="list-style-type: none"> <li>1. (<i>intransitive</i>) To use the hands.           <ul style="list-style-type: none"> <li>• They have hands, but they <b>handle</b> not - <i>Psalm 115:7</i></li> </ul> </li> <li>2. To touch; to feel with the hand.           <ul style="list-style-type: none"> <li>• <b>Handle</b> me, and see; for a spirit hath not flesh - <i>Luke 24:39</i></li> </ul> </li> <li>3. To use or hold with the hand.           <ul style="list-style-type: none"> <li>• About his altar, <b>handling</b> holy things - <i>John Milton</i></li> </ul> </li> <li>4. To manage in using, as a spade or a musket; to wield; often, to manage skillfully.           <ul style="list-style-type: none"> <li>• That fellow <b>handles</b> his bow like a crowkeeper - <i>Shakespeare, King Lear, I/iv</i></li> </ul> </li> <li>5. To accustom to the hand; to work upon, or take care of, with the hands.           <ul style="list-style-type: none"> <li>• The hardness of the winters forces the breeders to house and <b>handle</b> their colts six months every year - <i>Sir W. Temple</i></li> </ul> </li> <li>6. To receive and transfer; to have pass through one's hands; hence, to buy and sell; as, a merchant <b>handles</b> a variety of goods, or a large stock</li> <li>7. To deal with; to make a business of.           <ul style="list-style-type: none"> <li>• They that <b>handle</b> the law knew me not - <i>Jeremiah, 2:8</i></li> <li>• 2011 December 16, Denis Campbell, "Hospital staff 'lack skills to cope with dementia patients'", <i>Guardian</i>: The findings emerged from questionnaires filled in by 2,211 staff in 145 wards of 55 hospitals in England and Wales and 105 observations of care of dementia patients. Two-thirds of staff said they had not had enough training to provide proper care, 50% said they had not been trained how to communicate properly with such patients and 54% had not been told how to <b>handle</b> challenging or aggressive behaviour.</li> </ul> </li> <li>8. To treat; to use, well or ill.           <ul style="list-style-type: none"> <li>• How wert thou <b>handled</b> being prisoner - <i>Shakespeare, Henry VI, Part I, I-iv</i></li> </ul> </li> <li>9. To manage; to control; to practice skill upon.           <ul style="list-style-type: none"> <li>• You shall see how I'll <b>handle</b> her - <i>Shakespeare, Measure for Measure, V-i</i></li> </ul> </li> <li>10. To use or manage in writing or speaking; to treat, as a theme, an argument, or an objection.           <ul style="list-style-type: none"> <li>• We will <b>handle</b> what persons are apt to envy others - <i>Francis Bacon</i></li> </ul> </li> <li>11. (<i>soccer</i>) To touch the ball with the <b>hand</b> or <b>arm</b>; to commit handball.           <ul style="list-style-type: none"> <li>• 2011 February 12, Les Roopanarine, "Birmingham 1 - 0 Stoke", BBC: Robert Huth <b>handled</b> a Bentley shot, only for the offence to go unnoticed.</li> </ul> </li> </ol>	<p><b>handle</b></p> <p>1 <b>DO WORK</b> [transitive] to do the things that are necessary to complete a job:</p> <ul style="list-style-type: none"> <li>• I handled most of the paperwork.</li> <li>• The case is being handled by a top lawyer.</li> <li>• The finance department handles all the accounts.</li> <li>• Computers can handle huge amounts of data.</li> </ul> <p>2 <b>DEAL WITH A SITUATION</b> [transitive] to deal with a situation or problem by behaving in a particular way and making particular decisions:</p> <ul style="list-style-type: none"> <li>• The headmaster handled the situation very well.</li> <li>• I knew I had handled the matter badly.</li> <li>• Leave it to me. I can handle it.</li> <li>• Most customers were satisfied with the way their complaints were handled.</li> <li>• Opposition leaders will be watching carefully to see how the Prime Minister handles the crisis.</li> </ul> <p>3 <b>DEAL WITH A PERSON</b> [transitive] to deal with a person or behave towards them in a particular way, especially in order to keep them happy:</p> <ul style="list-style-type: none"> <li>• Some customers are quite difficult to handle.</li> </ul> <p>4 <b>NOT BECOME UPSET</b> [transitive] to not become upset in a difficult situation:</p> <ul style="list-style-type: none"> <li>• She can't handle it when people criticize her.</li> <li>• He doesn't handle stress very well.</li> </ul> <p>5 <b>HOLD</b> [transitive] to touch something or pick it up and hold it in your hands:</p> <ul style="list-style-type: none"> <li>• He had never handled a weapon before.</li> <li>• We teach the children to handle the animals gently.</li> <li>• He was roughly handled by the mob.</li> </ul> <p>6 <b>CONTROL A VEHICLE</b></p> <p>a) [transitive] to control the movement of a vehicle or an animal:</p> <ul style="list-style-type: none"> <li>• I didn't know if I'd be able to handle such a large vehicle.</li> </ul> <p>b) [intransitive] the way a vehicle handles is how easy it is to control</p> <p><b>handles well/badly</b></p> <ul style="list-style-type: none"> <li>• The car handles well, even on wet roads.</li> </ul> <p>7 <b>MOVE GOODS</b> [transitive] to move goods from one place to another:</p> <ul style="list-style-type: none"> <li>• The Post Office handles nearly 2 billion letters and parcels over the Christmas period.</li> </ul> <p>8 <b>BUY/SELL GOODS</b> [transitive] to buy or sell goods:</p> <ul style="list-style-type: none"> <li>• Bennet was charged with <b>handling</b> stolen goods.</li> </ul>
--	--

Figure 7: Entries for the verb HANDLE in the English *Wiktionary* (left) and the *Longman Dictionary of Contemporary English*, free online version (right)

In Figure 7, two entries for the English verb HANDLE are presented side by side: on the left, an entry from the English *Wiktionary*, and on the right, one from the free online version of the *Longman Dictionary of Contemporary English*. The verb HANDLE was chosen at random, only considering relatively common verbs. It is immediately striking that the very first sense covered in *Wiktionary* represents a use that is highly unfamiliar in contemporary English, if ever there indeed was a genuine intransitive pattern, rather than a peripheral absolute use with an elliptical object. Even the *Oxford English Dictionary* does not record such a pattern. In general, senses are not very clearly distinguished, and their ordering appears haphazard. Definitions themselves tend to be made up of lists of rather general words, often used in non-contemporary or rare senses, such as *use* in sense 8, where there is no indication that the object is meant to be human. Examples are mostly citations from the Bible or old literary classics, and are all archaic without indicating this fact. Much of this is traceable to the 1913 edition of *Webster's Revised Unabridged Dictionary*, available free of charge on many dictionary aggregator sites (Lew 2011). Although two modern citations have been added from journalistic texts, the one under sense 7 is much too long: only its very final part is at all relevant here. No attention is given to word combinations.

Compared to this, the entry for the same word from the [\*Longman Dictionary of Contemporary English\*](#) is like a breath of fresh air, with its clearly signposted (see e.g. senses 2 and 3) and perspicuously defined senses. Examples of use are contemporary, authentic, and natural-sounding. They illustrate the combinatory potential of HANDLE, and offer usable models for production. All in all, the comparison of the two entries aptly illustrates the significant progress that lexicography has made over the last decades or so. Ironically, it is the collaborative entry, one produced under an innovative Web 2.0 model, which exemplifies outdated ways of lexicographic description. It seems that the web community, while enthusiastically embracing the novelty of online collaboration, propagates the traditional model of lexicographic description perhaps received from their repeated encounters with respectable volumes sitting proudly on their parents' bookshelves.

While a UGC-driven model, perhaps best known from [\*Wikipedia\*](#), has been a resounding success in terms of encyclopaedic content (factual data), it does not seem to work so well in lexicographic endeavours, such as [\*Wiktionary\*](#). [\*Wikipedia\*](#) revolves on the principle that somewhere out there in the world there are experts on every little bit of knowledge, willing to give of their time to freely share their knowledge with other people. While this model works surprisingly well for the reporting of encyclopaedic facts, it is much less robust when it comes to the job of describing words and expressions of a language: their meaning, pronunciation, morphology, syntax, word combination (collocation and colligation), and usage. To put it simply, while it makes good sense that somewhere out there there is an expert on a piece of specialized knowledge (say, a rare species of nettle) who is willing to share their expertise with the world, there is normally no such expert on the meanings of a particular everyday word (say, the word *field*), who would be capable of teasing apart the senses and providing a nuanced treatment of the many combinations and uses of the word. Lexicographers do not usually specialize in isolated words or narrow subsets of the vocabulary: they are normally equipped to supply lexical descriptions of complete vocabularies. But this is all they do and few can afford to do the job for free. Rather than language professionals, UGC lexicography tends to be in the hands of self-proclaimed language experts with a willingness to share, but their best efforts cannot match the output of professional, trained lexicographers. This is so because the quality of lexicographic description heavily depends on a good grasp of lexicographic principles, procedures, and tools of the trade (such as skilled use of corpus data or structural markup). Such appears to be the situation for general language items; things might be looking somewhat better, though, at the interface of encyclopaedic and linguistic description: specialized vocabulary and terminology, where “[e]ach contributor has a certain field of expertise [...] [and this] fosters the encoding of a vast amount of domain-specific knowledge” (Meyer/Gurevych 2012, p. 259).

## 5.2 Define your world: [\*Urban Dictionary\*](#)

Another area of lexis which might benefit from the bottom-up approach are the elusive domains of slang, cant and jargon. This is the focus of another collaborative dictionary project, [\*Urban Dictionary\*](#) (Figure 8), which invites site visitors to “define your world”. [\*Urban Dictionary\*](#) actually predates [\*Wiktionary\*](#): its origins go back to 1999. Mainstream lexicography has traditionally been cautious about including fresh slang terms and fashionable words before they have proven themselves as likely to stay in active use for some time. The idea of [\*Urban Dictionary\*](#) is to capture such fleeting words and expressions, and it is an idea that has clearly caught on, given that the resource now holds a staggering 6,90 million definitions (as of 14 January 2013). This does not mean that anywhere near as many terms

are defined in [Urban Dictionary](#), as most terms have multiple (sometimes hundreds) alternative definitions from different users. By adding up the numbers of entries as presented for each respective letter page, I estimated the number of different terms at 1,38 million (as of 14 January 2013).

The screenshot shows the Urban Dictionary homepage with the search term 'bootyism' entered in the search bar. The results are listed under three main sections:

- 1. bootyism**: Definition: "Tiger Woods' real faith and religion." Rating: 179 up, 47 down. Examples: "Simon: did you hear that Tiger have cured his nymphomania with Buddhism?" and "John: No, mate! It was Bootyism...". Buttons: buy bootyism mugs & shirts, share this, add a video.
- 2. bootyism**: Definition: "a religion originating in the east(coast) encouraging the endowment of large quantities of (female) booty, getting booty, banging booty, etc." Rating: 435 up, 169 down. Examples: "1)I've recently converted to bootyism. 2)Since the divorce, I've been able to practice bootyism.". Buttons: buy bootyism mugs & shirts, share this, add a video.
- 3. Bootyism**: Definition: "1) Like Alcoholism, bootyism is extreme sex addiction triggered by craving booty (hind quarters) to the point of making life unmanageable. 2) Often a poor attempt at humor, bootyism is an ignorant or scathing reference of the religion of Buddhism." Rating: 65 up, 20 down. Buttons: buy Bootyism mugs & shirts, share this, add a video.

The sidebar on the left lists various related terms under the heading 'bootyism'.

Figure 8: An [Urban Dictionary](#) entry for BOOTYISM (15 September 2012)

As seen in Figure 8, an [Urban Dictionary](#) entry contains user-supplied definitions, examples of use, and tags. Tags may be useful for searching the dictionary and cross-linking related entries (the dictionary also has an alternative *thesaurus* access interface). Other users can rate the different definitions for quality, and the order in which alternative definitions are displayed depends on those collective ratings. As a result, the ‘best’ definitions are not necessarily, and not usually, those that make a genuine attempt at explaining the meaning, but rather those that are seen as clever or otherwise amusing. Thus, for some of the entries at least, [Urban Dictionary](#) has become a playing field for put-down, scorn, displays of in-group identification, and all sorts of ideology. For example, the top-rated definition of BOOTYISM in [Urban Dictionary](#) (Figure 8) is not so much a comment on the (meaning of the) headword, but rather on Tiger Woods's character. And the second most popular ‘definition’ of RELIGION runs as follows:

The biggest lie in human history. It has been responsible for more deaths throughout human history than all other unnatural causes combined. For a thousand years the Church was a tyrannical dictatorship that used religion to control the uneducated masses. Free your minds and come into the 21st century. (<http://www.urbandictionary.com/define.php?term=Religion&defid=6179390>)

Another user's definition of RELIGION is this: "Apparently what most people on Urban Dictionary hate." — clearly a meta-comment on the ideologization of lexicography in [Urban Dictionary](#). Sometimes, however, a *bona fide* definition will make it into first page, an example being number 3 in Figure 8, or the following one for RELIGION: "An organized group of people with a common belief. Most religions strongly stress ethics and morals along with setting guidelines for people to follow in their day-to-day life."

It is doubtful if users of [Urban Dictionary](#) truly come to the site in search of the basic meaning of such common words as *religion*, so here the strategy of providing something else may actually not be incompatible with the needs of the users of the service. In contrast, for actual slang terms such as *bootyism*, meaning may be what users genuinely seek (especially non-native speakers of English who have failed to find an item of slang in a mainstream dictionary). Well aware that their definitions do not help with this, witty users will sometimes compensate for the lexicographic vacuity of their definitions in the other field of substance, namely *examples of usage*. Interestingly, instances of genuine, or even remotely plausible, examples of usage are even rarer than *bona fide* definitions. It is hard to say whether this is more a matter of misunderstanding the design principle behind this particular data field, flouting it, or else living proof of the claim (by corpus linguists and lexicographers such as John Sinclair or Patrick Hanks) that usage cannot be invented.

### 5.3 [Wordnik](#)

[Wordnik](#) (Figure 9) is a relatively new (founded in 2009) collaborative resource which offers a rather successful combination of professional lexicographic content from existing dictionaries, its own corpus-based content, content pulled from [Twitter](#) and [Flickr](#), and user-generated content.

The screenshot shows the Wordnik website interface for the word 'rosemary'. At the top, there is a navigation bar with links for 'Community', 'Word of the day', 'Random word', 'Search' (with a magnifying glass icon), and 'Log in or Sign up'. Below the navigation bar, the word 'rosemary' is displayed in a large, bold, black font. To the right of the word are several buttons: 'Define' (blue), 'Relate' (green), 'List' (pink), 'Discuss' (orange), 'See' (teal), 'Hear' (yellow), and 'Love' (red). Underneath the word, there is a link to 'Log in or sign up to show \'rosemary\' some love.' On the left side, under the heading 'Definitions', there are two entries from the 'American Heritage® Dictionary of the English Language, Fourth Edition': one for the plant and one for the leaves used as a seasoning. On the right side, under the heading 'Examples', there are four examples from Twitter posts, each with a timestamp: 'From Twitter 07-13-2009', 'From Twitter 07-07-2009', 'Today's Garden', and 'He gets it from his dad.'

Figure 9: Top of a [Wordnik](#) entry for ROSEMARY

According to the co-founder and CEO of [Wordnik](#) (McKean 2011), its users (calling themselves *Wordniks*) like to think of the service more as a game than a dictionary. Very few users want to write definitions; most prefer to supply metalinguistic information instead, such as tags, comments, or word lists. [Wordnik](#)'s users also like to invent new words; so much so, that they have even come up with an in-house adjective to describe such words: *madeupical*. The structure of [Wordnik](#) is user-activity-based, as indicated by the coloured-font options in the top right corner in Figure 9, with the sections stacked on top of one another in one lengthy page. In the default view of a word, definitions from existing works of reference, including out-of-copyright dictionaries, [Wiktionary](#) and [WordNet](#), are accompanied by citations, including some from [Twitter](#). The technology behind [Wordnik](#) prioritizes citations with self-defining contexts, especially when no prior dictionary definitions are available for a term.

Below the definition-and-examples section, a related-words section is presented, which includes hypernyms (more general words), words found in similar contexts, words that contain the headword in their definitions, and semantic tags for the headword. Further down, user-created lists (Figure 10) are presented. Interestingly, word lists are amongst the favourite features of the service.

The screenshot shows the Wordnik website interface. At the top, there's a navigation bar with links for 'Community', 'Word of the day', and 'Random word'. A search bar is on the right. Below the header, the main content area has a title 'Lists' and a sub-instruction: 'These user-created lists contain the word "rosemary".'

The lists are categorized and include:

- I Found A Rose**: List of words containing the string "rose" - from aprosexia nasalis to prosectionist. Includes: aprosexia nasalis, prosectionist, erose, roseate, tuberose, morose, grosella, retrosexual, serose, arose, Drosera, haroset and 154 more...
- Herbs: How Fragrant, How Delicious**: Delicious scents in an edible nibble. Includes: zedoary, zahtar, yarrow, wormwood, wasabi, verbena, valerian, thyme, Thai basil, tarragon, sweet basil, sorrel and 59 more...
- If-Christ-Had-Not-Died-For-Thee-Thou....**: Words used as baby names, including virtue names, nature names, place names, etc. Includes: Words that have been used as baby names, including virtue names, nature names, place names, etc.
- food collection**: bread, peel, pot, chorizo, Filet, olive, fill, Phylo, dough, bake, mat, pinot and 988 more...
- Scrabble Names**: Given names acceptable for play the last time I checked the OWL. Includes: kris, ray, barb, morris, kat, mark, maria, erica, marge, mason, hunter, hazel and 164 more...
- the first list**: an immense, grandiloquent list that loads like a thousand years sentence in stone. new words are in the other lists. includes: ridiculous, brummagem, predicament, sanctimonious, vapid, eschew, admonish, auspicious, capitulation, enumerate, lachrymose, tenet and 1648 more...
- #1**: Words I Like

On the right side, there's a sidebar titled 'More lists containing "rosemary"' with a list of items:

List Name	Number of words
the herbalist	69 words
Herbs and Spices	74 words
Wordies for a Foodie	25 words
Food Mmmmm's	72 words
What Joanna Wrote	17 words
Culinary	18 words
Sunday	14 words
Beauty Is Truth, Truth Beauty	50 words
vegetation	62 words
mary, mary, quite contrary	59 words
Words Covered in Faery Dust...	78 words
Whatever you rolled in, it ...	79 words
Grocery List	50 words
foodie list - well seasoned	44 words
in the kitchen	24 words
I don't give a damn	31 words

Figure 10: Word lists in [Wordnik](#) containing the word *rosemary*

Not all of the user-created lists would impress a professional lexicographer, such as a list of words containing the string *rose* ("I Found a Rose"), many of which bear no linguistic relationship to *rose* except the mere sequencing of letters (e.g. *aprosexia*). Word lists are not [Wordnik](#)'s original invention: they were inherited from [Wordie](#) (Figure 11), a former service with which [Wordnik](#) merged, incorporating its user-generated content.

*Wordie*'s tongue-in-cheek slogan was "Like Flickr, but without the photos" (Figure 11). [Wordnik](#), in contrast, does include photos, pulling them from none other than [Flickr](#). Image

selection is based on tags, but this does not guarantee that the pictures so chosen are lexicographically relevant. In the case of ROSEMARY, for instance (Figure 12), how most of the images are related to the headword cannot be appreciated by someone not already familiar with the different uses of the word and its various designates.

**Wordie [wərd̩ . ē] ↗**  
Like Flickr, but without the photos.

Login ↗ or Register ↗ | word search | go

Wordie lets you make lists of words -- practical lists, words you love, words you hate, whatever. You can then see who else has listed the same words, and talk about it. It's more fun than it sounds.

When adding comments and citations wrapping words in [brackets] now links them to their Wordie page. Mad, mad props to [angharad ↗](#) for contributing a Bookmarklet ↗, which lets you select words on any web page and import them directly into Wordie. Works with Safari and Firefox (Mac or PC), does *not* work with IE 6. To install, bookmark the 'Add to Wordie' link below (on Safari and Firefox you can just drag it onto your bookmarks bar).

bookmark this -> [Add to Wordie](#)

Wordie hacking is heartily encouraged: please let me know if I can do anything to facilitate a project. Wordie suggestions and bug reports always appreciated (via [email](#) or my [profile ↗](#)). If you like Wordie, please [dig it ↗](#). Yahoo! ↗ has a short video segment on Wordie. Ridiculous. We're vid #2.

**Recent Words (rss ↗)**

penumbra ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
lilt ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
schadenfreude ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
canter ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
snowth ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
drawl ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
melody ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
splay ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
sprawl ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
damask ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
grin ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
nimbus ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗
rain ↗	... ninjawords ↗   tfd ↗   dictionary.com ↗   M-W ↗   UD ↗   Wikipedia ↗   Google ↗

Total Words: **29,689**  
Unique Words: **18,812**  
Word Lists: **1,142**  
Wordies: **737**

**Wordiest Wordies ↗**

colleen ↗	(1714 words)
Kaichi ↗	(1330 words)
kat ↗	(1281 words)
robotjohnny ↗	(1118 words)
threedotted ↗	(1094 words)
dienekes ↗	(972 words)
lanas ↗	(930 words)

Figure 11: The service *Wordie* (no longer operative) as seen on 5 December 2006

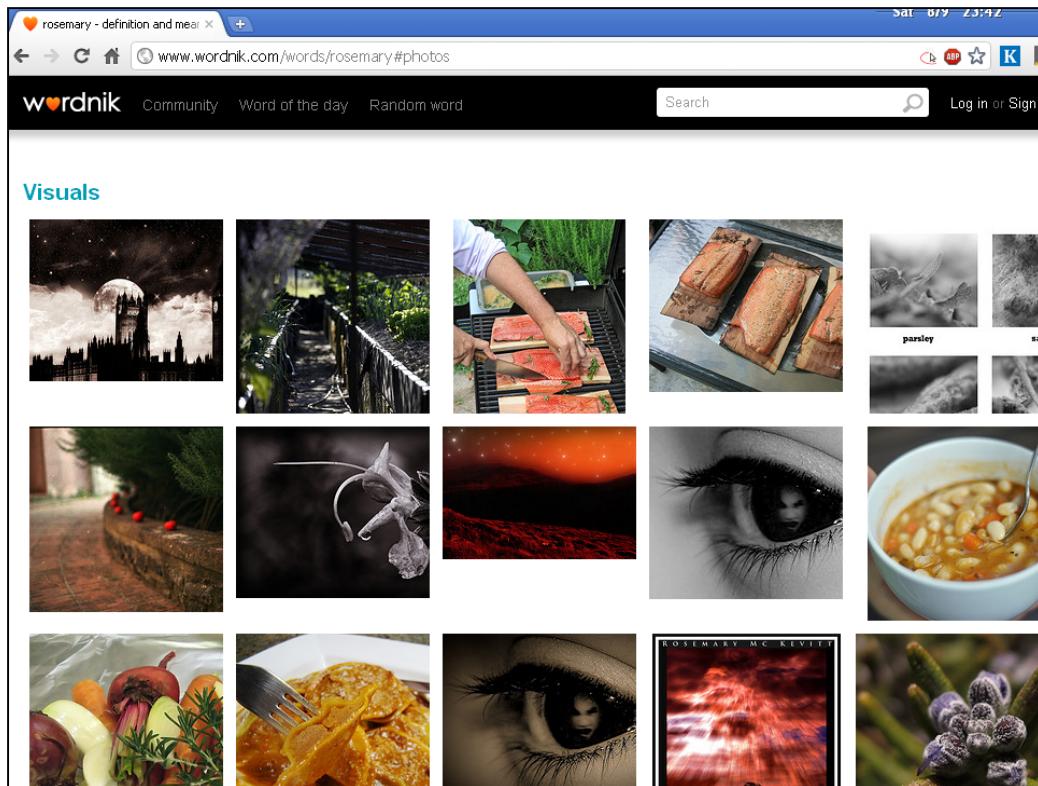


Figure 12: [Flickr](#) images for ROSEMARY in *Wordnik*

## **6. User-generated content in professional dictionaries**

Until recently, most major English-language publishers had rarely engaged users in bottom-up lexicography, one exception being the *Oxford English Dictionary*, which has a fairly long tradition of user participation. Much public attention was drawn to the issue with an announcement made by Collins on 17 July 2012, covered in *The Independent*, *The Guardian*, and *Daily Mail*, and an ensuing discussion in the social media, including the Lex group on [LinkedIn](#). The new feature is called “What’s your word?”. Alex Brown, Head of Digital at Collins was quoted as saying “Most dictionaries are static. By allowing the public to truly participate, we’re ensuring that we stay on top of the evolving English language.” (O’Neill 2012). However, in the same article we learn that “Collins will ultimately decide which words are included based in part on its 4.5 billion-word database of language called the Collins Corpus … based on a number of criteria including frequency of use, number of sources and staying power”. Having read this, someone who knows about dictionary-making cannot but wonder what new words users can possibly offer, and how come such words are not *already* in the dictionary. It could be that the primary motivation behind engaging users is different from that stated: the publisher may be counting on the sense of community borne out of users being given a chance to, on the one hand, contribute their own entries, and, on the other hand, enjoy entries edited by their peers. In this way, casual dictionary users may become more meaningfully (pun intended) involved.

In order to suggest an item for inclusion in the *Collins English Dictionary*, registered users need to fill in a submission form. The form (Figure 13) is very simple and assumes no lexicographic sophistication on the part of the user, with only three text boxes provided for, respectively: (1) the new term, (2) its definition, and (3) additional information.

You are here > Submit a word to the Collins English dictionary

# Submit a word to the Collins English Dictionary

All submissions will be reviewed for inclusion in the next updates of our online and print dictionaries

Your suggested word: \*

Definition: \*

Additional information (e.g. word origin, example sentences):

Post updates to your facebook wall:

[Submit >](#)

Figure 13: [Collins English Dictionary](#) new word submission form

The screenshot shows a user-submitted word entry for 'A/C'. At the top, it says 'You are here > Home > Latest word submissions > A/C'. The main title is 'A/C New Word Suggestion'. Below that, it says 'Submitted By: sarathbabu.s' and 'Approval Status: Pending Investigation'. There are two sections under a bullet point: 'Definition of A/C' which defines it as 'Stands for Air conditioner. It also stands for Account. Also means Assistant Commissioner.' and 'Additional Information' which notes 'Many origins.'

Figure 14: Example entry suggestion to the *Collins English Dictionary*

Submitted entries are subject to acceptance by Collins editors. A review of the recent submissions shows many to be mediocre at best, and users' lack of sophistication shows through in accepted entries. A typical submission is shown in Figure 14. Quite often, users will suggest words which are already included in the dictionary: a recent example of suggested addition to the *Collins English Dictionary* is YOUSE, whereas YOUS or YOUSE is already a headword in the dictionary.

Despite the publicity it had received, Collins was not the first authoritative English-language dictionary publisher to introduce bottom-up lexicography as part of its online offerings. A notable case is *Macmillan English Dictionary*, a dictionary primarily but not exclusively for learners of English. In February 2009, its UGC-model *Open Dictionary* was launched, at the same time as the main *Macmillan English Dictionary Online*. Since then, the dictionary site has invited users to contribute entries to this user-created supplement. In order to submit a new item, a user fills in a special submission form (shown in Figure 15), very similar to the Collins form (though, of course, the Macmillan form was there earlier). Submitted entries are reviewed by Macmillan staff, and a selection are allowed to enter the word list of the *Open Dictionary*.

When researching entries submitted to Macmillan's *Open Dictionary*, in September 2012 I located an entry as in Figure 16 containing an ungrammatical definition ("used to parts..."). At the time of this writing (15 January 2013), the defective definition has been corrected to "used to refer to parts...". This is a positive example of quality control exercised by professional Macmillan editors over user input, which at present amounts to about 1,400 accepted entries. But if this number continues to grow, the challenge of maintaining and updating all the additional entries as well as the main dictionary may become too much of a job. Is it worth the trouble then? Assessing three years of experience with Macmillan's crowd-sourced *Open Dictionary*, Michael Rundell (2012, p. 80) writes that "the most fruitful areas where users can make a contribution are neologisms, regional varieties, and technical terminology" (this is sometimes referred to in the industry as the *long tail* of lexis). This observation dovetails with my (informal) assessment of the quality of general-language entries in user-constructed dictionaries.

**New Word**

What's your word?

What does it mean?

Could you give us an example sentence?  
(optional)

**Your Details**

Name/Nickname

Where in the world are you?

**Submit**

Figure 15: *Macmillan Open Dictionary* new word submission form

**Most recent entries**

**Add a word**

[Browse all entries](#)

**OE**

**NOUN**

original equipment; used to parts made for a car by the car manufacturer itself, or made exclusively for the car manufacturer by a third party

*There's good news for anyone called on to repair a Saab. You will still be able to source OE replacement parts, even though Saab is currently going through bankruptcy proceedings and has brought a halt to manufacturing.*

Submitted from United Kingdom on 12/07/2012 15:10:00

Figure 16: Entry for OE submitted to *Macmillan Open Dictionary* (15 September 2012)

## 7. Conclusion

Serious lexicography requires specialized skills and expertise; for this reason, a large part of crowdsourced content is of questionable lexicographic value, as we have seen in the entry for HANDLE in *Wiktionary* and BOOTYISM in *Urban Dictionary*. Pulling content from other social sites, though an easy way to enrich content, often produces irrelevant material, and sometimes considerable embarrassment, as in the former *Google English Dictionary* entry for KILT (Figure 17), which for several months in 2009 inadvertently appended a picture of a male wearing a kilt with no underwear, because it happened to be the most popular Google image for the word *kilt*.

The screenshot shows the Google English Dictionary interface. At the top, there's a search bar with 'English dictionary' and a dropdown arrow, followed by another search bar containing 'kilt' and a 'Search' button. Below the search bar, the word 'Dictionary' is displayed. The main content area starts with 'Found in dictionary: English > English.' A 'Starred words »' link is shown with a yellow star icon. The word 'kilt' is highlighted with a blue star icon and its pronunciation '/k'ilt/'. To the right of the word, there's a 'Images' link. Below the word, under 'Synonyms:', it says 'verb: pleat'. Under 'Kilts', it says 'plural'. A numbered definition follows: '1. A kilt is a skirt with a lot of vertical folds, traditionally worn by Scottish men.' Below this, it says 'Kilts can also be worn by women and girls.' with a note '(N-COUNT)'. To the right of the text, there's a small image of two men in kilts sitting on a bench outdoors. Below the image, the URL 'web.ripnet.com' is visible.

Figure 17: The *Google English Dictionary* entry for KILT (screenshot taken on 18 December 2009)

However, I have mostly looked at monolingual dictionaries, and constructing effective definitions is known to be notoriously difficult. For non-native speakers of the language, definitions, however skilfully written, are not usually the best way to convey meaning (Lew 2004; Adamska-Sałaciak 2010). There may be more room for collaboration between professional lexicographers and active users on bilingual (and multilingual) tools, and some bilingual dictionary publishers, for instance in Germany (to just mention *Duden* and *PONS*) or Japan (cf. Eijiro at <http://eijiro.jp>) are trying to go down that road. One possibility of engaging online users which, to my knowledge, has not been tried yet would be to get users to suggest interlingual equivalents based on professionally supplied raw data, such as automatically-generated concordances illustrating source-language use, or perhaps parallel concordances. This might provide a way for non-professionals to engage in useful lexicography, making the job closer to more intuitive text translation.

For now, areas of vocabulary where amateur lexicographers can make the most significant contribution are neologisms, slang, regional varieties, and technical terminology. In contrast, general vocabulary is probably best left to professional lexicographers.

## 8. References

- Adamska-Sałaciak, Arleta (2010): Why we need bilingual learners' dictionaries. In: Kernerman, Ilan J./Bogaards, Paul (eds.): English learners' dictionaries at the DSNA 2009. Tel Aviv, pp. 121-137.
- Fuertes-Olivera, Pedro A. (2009): The function theory of lexicography and electronic dictionaries: WIKTIONARY as a prototype of collective free multiple-language internet dictionary. In: Bergenholz, Henning/Nielsen, Sandro/Tarp, Sven (eds.): Lexicography at a crossroads: Dictionaries and encyclopedias today, lexicographical tools tomorrow. Bern, pp. 99-134.
- Giles, Jim (2005): Internet encyclopaedias go head to head. In: Nature 438, pp. 900-901.
- Hanks, Patrick (2012): Corpus evidence and electronic lexicography. In: Granger, Sylviane/Paquot, Magali (eds.): Electronic lexicography. Oxford, pp. 57-82.
- Krizhanovsky, Andrew (2011): A quantitative analysis of the English lexicon in Wiktionaries and WordNet. In: Trudy SPIIRAN 19, pp. 87-101.
- Landau, Sidney I. (2001): Dictionaries: The art and craft of lexicography. 2nd edition. Cambridge.
- Lew, Robert (2004): Which dictionary for whom? Receptive use of bilingual, monolingual and semi-bilingual dictionaries by Polish learners of English. Poznań.

- Lew, Robert (2011): Online dictionaries of English. In: Fuertes-Olivera, Pedro A./Bergenholtz, Henning (eds.): e-Lexicography: The Internet, Digital Initiatives and Lexicography. London/New York, pp. 230-250.
- McKean, Erin (2011): Wordnik: Notes from an online dictionary project. Electronic lexicography in the 21st century: New applications for new users (eLex2011). Bled, Slovenia.
- Melchior, Luca (2012): Halbkollaborativität und Online-Lexikographie. Ansätze und Überlegungen zu Wörterbuchredaktion und Wörterbuchforschung am Beispiel LEO Deutsch-Italienisch. In: Lexicographica 28, pp. 337-372.
- Meyer, Christian M./Gurevych, Iryna (2012): Wiktionary: a new rival for expert lexicons? Exploring the possibilities of collaborative lexicography. In: Granger, Sylviane/Paquot, Magali (eds.): Electronic lexicography. Oxford, pp. 259-291.
- O'Neill, Cordelia. (2012, July 17). Collins dictionary invites word suggestions. The Independent. Retrieved from <http://www.independent.co.uk/arts-entertainment/books/news/collins-dictionary-invites-word-suggestions-7952775.html>.
- Ptaszyński, Marcin Overgaard (2009): Function-related secondary user needs and secondary data? A critical examination of some central concepts in the modern theory of lexicographical functions. In: Lexikos 19, pp. 242-261.
- Rundell, Michael (2012): 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. In: Fjeld, Ruth Vatvedt/Torjusen, Julie Matilde (eds.): Proceedings of the 15th EURALEX International Congress. Oslo, pp. 47-92.
- Storrer, Angelika (2010): Deutsche Internet-Wörterbücher: Ein Überblick. In: Lexicographica 26, pp. 155-164.
- Surowiecki, James (2005): The wisdom of crowds. New York.
- Tarp, Sven (2008): Lexicography in the borderland between knowledge and non-knowledge: General lexicographical theory with particular focus on learner's lexicography. (= Lexicographica Series Maior 134.). Tübingen.
- Vickery, Graham/Wunsch-Vincent, Sacha (2007): Participative web and user-created content: web 2.0, wikis and social networking. Paris.

# Ansätze zu einer halbkollaborativen Lexikographie

Luca Melchior, Universität Graz

## 1. Ziele des Beitrags

Im vorliegenden Beitrag sollen einige Überlegungen zum Potential eines halbkollaborativen Ansatzes für die lexikographische Arbeit – insbesondere im Rahmen onlinelexikographischer Projekte – und für die metalexikographische Reflexion vorgestellt werden. Ausgehend von praktischen Beispielen aus einem konkreten Wörterbuchprojekt – die bilingualen Onlinewörterbücher von LEO ([dict.leo.org](http://dict.leo.org)) – werden Entwicklungsmöglichkeiten gezeigt, die durch eine stärkere Involvierung der Benutzer/Sprecher im Wörterbuchprojekt eröffnet werden und in deren Zentrum die Sprecher, ihre Sprache, ihre Wünsche und Bedürfnisse stehen, ohne dabei auf die Qualitätskontrolle einer lexikographisch geschulten Redaktion zu verzichten.<sup>1</sup> Des Weiteren sollen zusammenfassend die Vorteile, aber auch die Probleme des Ansatzes präsentiert und über das einzelne konkrete Projekt hinausgehende Anwendungsmöglichkeiten desselben in der Lexikographie und anderen sprachwissenschaftlichen Bereichen skizziert und damit ihre Darstellung aus Melchior (2012) ergänzt und fortgeführt werden.

## 2. Das Internet und die Lexikographie

Klassische Referenzwerke der Wörterbuchforschung, die in Zeiten der fast ausschließlich papiernen Wörterbucharbeit erschienen sind, wie Bergenholz/Mugdan (1990), Wiegand (1998), Zgusta (1971), aber auch neuere Einführungen in die Wörterbuchforschung, die in Zeiten elektronisch- und onlinegeprägter Lexikographie entstanden sind, wie die neuere Ausgabe von Engelberg/Lemnitzer (2009, S. 235-238), nennen drei grundsätzliche Informationsquellen, die dem Wörterbuchmacher zur Verfügung stehen:

Die *primären Quellen* sind vor allem (aber nicht nur) Texte, welche aus natürlichen oder quasi-natürlichen Kommunikationssituationen stammen, oder größere zusammenhängende Ausschnitte aus solchen. Meistens handelt es sich um schriftliche Texte. Dies ist aber nicht notwendigerweise der Fall. Auch Tonbänder, Schallplatten, Videoaufnahmen etc. können als primäre Quellen in Frage kommen. Die Menge der primären Quellen, welche Texte oder größere zusammenhängende Ausschnitte sind, bildet das *lexikographische Korpus*. Dieses ist also immer ein *Textkorpus* [...]. – Zu den *sekundären Quellen* gehören alle Wörterbücher, die nach dem Instruktionsbuch entweder obligatorisch oder fakultativ konsultiert werden sollen, und zu den *tertiären Quellen* gehören alle sonstigen Sprachmaterialien, die benutzt werden, wie z.B. linguistische Monographien und Grammatiken [...]. (Wiegand 1998, S. 140)

Obwohl die mediale Revolution, die durch die Verbreitung des Internets erfolgt ist, weitreichende Auswirkungen auf die lexikographische Arbeit sowohl bezüglich der Konzipierung als auch der Gestaltungsmöglichkeiten von Wörterbüchern gehabt hat, hat sie an der grundsätzlichen Art dieser Quellen nichts geändert, sehr viel hingegen im Hinblick auf die Menge, die Verfügbarkeit und Erreichbarkeit sowie (vielleicht) auf die Qualität derselben. Google, YouTube, Blogs, soziale Netzwerke sowie die publizistische Arbeit im Netz stellen den Wörterbuchmachern gewaltige Mengen an Sprachmaterialien zur Verfügung, sowohl in frei

<sup>1</sup> Vgl. dazu auch de Schryver/Prinsloo (2000a, S. 3).

zugänglichen Korpora als auch in Form von unterschiedlichsten sprachlichen Daten;<sup>2</sup> metasprachliche Nutzerdiskussionen liefern wichtige Sprechermeinungen; lexikographische und linguistische Arbeiten aller Art sind aus dem bzw. im Netz beziehbar. Das Internet hat Möglichkeiten für die Wissensschaffung, -tradierung und -aneignung im Allgemeinen und für die lexikographische Arbeit als kulturelle Praxis (vgl. Wiegand 1998, insb. S. 38-42) im Besonderen eröffnet, welche die Wörterbuchproduktion z.T. tief greifend erneuert haben, vor allem bezüglich der Konzipierung des Produkts „Wörterbuch“. Dies ist in zweierlei Hinsicht zu verstehen: Auf der einen Seite ermöglicht das Internet eine ständige Aktualisierung, Erweiterung, Verbesserung und Anpassung – sowohl was die inhaltliche Seite als auch was die Benutzerfreundlichkeit und den Umfang des Angebots angeht. Man hat daher die Notwendigkeit gesehen, für offene Onlinewörterbücher das neue Merkmal der Dynamik einzuführen (vgl. Lemberg 2001, S. 81).<sup>3</sup> Andererseits erlaubt die digitale Onlineform einen raschen und ständigen Austausch mit den Nutzern, so dass es für die Wörterbuchredaktion möglich ist, deren Wünsche, Anforderungen und Kritik zu erfahren und ggf. das Wörterbuch entsprechend zu ändern und anzupassen. Für die Nutzer selbst bietet sich darüber hinaus die Möglichkeit, sich in vielfältiger Weise aktiv an der Entwicklung des Wörterbuchprojekts zu beteiligen. Das Web 2.0 hat Kollaborativität und *bottom-up*-Ansätze<sup>4</sup> zu Grundprinzipien der Wissensbeschaffung, -tradierung und -multiplizierung gemacht. Dies wurde auch von der lexikographischen Praxis erkannt und in unterschiedlichem Maße sowie auf unterschiedliche Art und Weise genutzt.

### 3. Der Wörterbuchnutzer und sein Beitrag in der Onlinelexikographie

Vor fast zehn Jahren beschrieb de Schryver (2003) in seinem Beitrag über „Lexicographer's Dreams“ die Erwartungen, die der mediale Wechsel bei den Wörterbuchmachern erweckt hatte, wie etwa die Demokratisierung des lexikographischen Prozesses. Das Beispiel Wiktionary mit seinem kollaborativen Prinzip scheint eine dieser Erwartungen zu erfüllen, wie Fuertes-Olivera (2009, S. 106, Kapitälchen im Original) schreibt:

Collective free-multiple-language Internet reference works such as WIKIPEDIA, WICTIONARY, and OMEGAWIKI, aim at including data for all the languages of the word by promoting the possibility for any person in any place to add and/or edit entries in any language following a protocol detailing the structure of the entry and the steps to be taken

Ein kollektives Nachschlagewerk mit Wörterbuchfunktion also, zu dem jeder beitragen kann und das nach dem Kollaborativitätsprinzip und dem Schwarmintelligenzgrundsatz unter Einsatz und Kontrolle der Usergemeinde wachsen sollte (vgl. Fuertes-Olivera 2009). Aber auch andere Onlinewörterbücher, die nicht auf dem Prinzip der Kollaborativität beruhen, haben die Vorteile der Einbeziehung ihrer Nutzer entdeckt,<sup>5</sup> wenn auch auf traditionellere Art und Weise und weniger radikal. So bieten die Wörterbücher vom Verlag Garzanti (Italienisch einsprachig sowie zweisprachig Italienisch-Englisch und Italienisch-

<sup>2</sup> Zum Internet als Korpus, vgl. unter anderem Rakotonolina (2011), zum Internet als mögliches Korpus für die Lexikographie und zu den damit bezogenen Fragen und Problemen vgl. Heid (2012), mit besonderer Berücksichtigung der Fachlexikographie vgl. Fuertes-Olivera (2012).

<sup>3</sup> Storrer/Freese (1996, S. 117) sprechen von „Wörterbücher[n] im Aufbau“, Schröder (1997, S. 60) von „Ausbauwörterbuch“, in Melchior/D'Agostini (2011, S. 261) ist von „offene[n] Projekte[n]“ die Rede, Martin (2012, S. 3; S. 6; S. 22) spricht von „lexicographie évolutive“. Für eine ausführliche Diskussion vgl. Klosa (2013). Zu den Aktualisierungsmöglichkeiten von Internettörterbüchern vgl. auch Engelberg/Lemnitzer (2009, S. 152f.), Haß/Schmitz (2010, S. 3).

<sup>4</sup> Zu *bottom-up*-Lexikographie vgl. die frühen Beobachtungen und Überlegungen von Carr (1997, S. 214f.).

<sup>5</sup> Bei usergenerierten Wörterbüchern ist es selbstverständlich, dass die Projekte durch die Mitarbeit der Nutzer wachsen. In den vorgestellten Fällen – anders als bei *Wiktionary* oder bei LEO – handelt es sich aber um Retrodigitalisierungen von Printwörterbüchern, die nicht direkt als Internetprodukte konzipiert worden sind.

Französisch, <http://garzantilinguistica.sapere.it/> [Stand: 13.04.2013]) die Möglichkeit, neue Einträge vorzuschlagen sowie Fehler zu melden; auch bei den zweisprachigen Onlinewörterbüchern von *PONS* (<http://www.pons.eu> [Stand: 13.04.2013]) dürfen die Nutzer neue Einträge verfassen und haben darüber hinaus die Möglichkeit, sich in einem Forum auszutauschen. Diesen Weg hatte der Verlag Zanichelli mit dem Wettbewerb „*Segnala una nuova parola*“ zur Meldung von Neologismen schon vor Jahren für das berühmte italienische Wörterbuch *Zingarelli* eingeschlagen.<sup>6</sup> Die Wörterbuchbenutzer können in den genannten Fällen ihren Beitrag höchstens in der Bearbeitung der *Inhalte* liefern, und zwar in Form von Verbesserung vorhandener Inhalte und von Erweiterung derselben. Sie können sich nicht an der aktiven Gestaltung des Wörterbuchprojektes, an seiner Konzipierung und grundsätzlichen Ausrichtung beteiligen. In den genannten Beispielen werden die Nutzer also lediglich als „Lektoren“ bzw. „Informanten“ für ein Produkt genutzt, das, wenn auch nicht fertig gestellt und unveränderbar, so doch zumindest als abgeschlossenes Werk konzipiert ist.<sup>7</sup> Experten und Laien werden klare und getrennte Rolle zugewiesen.

#### 4. Das Prinzip der Halbkollaborativität

Einen anderen Weg gehen die LEO-Wörterbücher, die nicht auf Kollaborativität wie *Wikipedia* oder *Wiktionary* basieren,<sup>8</sup> sondern auf einem Grundprinzip, das ich als Halbkollaborativität bezeichnet habe (vgl. Melchior 2009, 2012, Melchior/D'Agostini 2011). Das heißt, dass die Wörterbücher nicht benutzergeneriert, sondern benutzerunterstützt sind: Die Nutzer werden ins Wörterbuchprojekt mit einbezogen, ihr Beitrag zur Gestaltung des Wörterbuchs und zu seinem stetigen Ausbau ist vom lexikographischen und technischen Team erwünscht – das Wörterbuch ist nicht abgeschlossen, sondern erweitert sich und entwickelt sich stets weiter. Jede Entscheidung über mögliche Veränderungen – sowohl auf technischer als auch auf inhaltlicher Ebene – liegt jedoch in der alleinigen Verantwortung der Mitarbeiter. Das Wörterbuch ist ein Produkt des Lexikographen und des technischen Teams. Die User<sup>9</sup> können keine Änderungen an den Inhalten bzw. an der Struktur desselben vornehmen, sondern nur Vorschläge und Wünsche formulieren – die im „Wörterbuchaußentext“ *Forum* gepostet werden. Wenn man ein Kontinuum zwischen Experten- (akademischer, aber auch Verlags-) und Laienwörterbuchpraxis annimmt, sind die Wörterbücher von LEO also eher am ersten Pol anzusiedeln.<sup>10</sup> In LEO scheinen allerdings die Möglichkeiten des elektronischen internetbasierten Mediums bezüglich der Interaktion mit den Usern intensiver ausgeschöpft zu werden als dies

<sup>6</sup> Dagegen nutzen andere retrodigitalisierte kostenfrei zugängliche und benutzbare Wörterbücher wie z.B. die ein- und zweisprachige Wörterbücher von Larousse (<http://www.larousse.fr/dictionnaires/bilingues> [Stand: 14.03.2014]), Duden (<http://www.duden.de/woerterbuch> [Stand: 9.12.2012]) oder Treccani (<http://www.treccani.it/vocabolario> [Stand: 13.04.2013]) die Möglichkeit zur Einbindung der Nutzer nicht.

<sup>7</sup> Die User des Wörterbuchs werden dementsprechend nicht „as consumers of products of mass consumption“ (Fuentes-Olivera 2012, S. 50) gesehen.

<sup>8</sup> Abels Charakterisierung als kollaborative Wörterbücher (Abel 2006, S. 44) trifft daher nicht zu, wenn nicht auf die ersten Jahre des Projektes LEO Deutsch-Englisch. Zur Geschichte von LEO vgl. Melchior (2012, S. 338-343). Auch voll kollaborative Wörterbücher wie *Wiktionary* ermöglichen den Usern aber nur einen eingeschränkten Entscheidungsraum in der konzeptionellen Gestaltung und Ausrichtung derselben.

<sup>9</sup> Auch hier wie in Melchior/D'Agostini (2011, S. 263 FN24) sowie in Melchior (2012, S. 352 FN30) unterscheide ich zwischen *Nutzern* und *Usern*. Während die ersten (vereinfacht gesagt) eher passive Rezipienten des Produktes Wörterbuch sind, beteiligen sich die letztgenannten mit Vorschlägen, Anmerkungen und kritischen Beiträgen an der Gestaltung des lexikographischen Werkes.

<sup>10</sup> Auch wenn nicht alle Wörterbücher des Projektes hinsichtlich ihrer Qualität und der Kontrolle gleichzusetzen sind. Zu den unterschiedlichen Ausprägungen des Benutzerbeitrages vgl. Engelberg/Lemnitzer (2009, S. 152f.), Lew (in diesem Band), Gurevych/Meyer (in diesem Band) und, obwohl mittlerweile etwas veraltet, Nesi (2009, S. 474ff.). Vgl. auch die frühen Überlegungen von Storrer/Freese (1996, S. 118f.).

in anderen Wörterbüchern geschieht (vgl. dazu auch Müller-Spitzer 2003, S. 140f.).<sup>11</sup> Dies ermöglicht es, die Vorteile des Austausches mit den Usern und deren Beiträge mit der qualitativen Arbeit und der Kontrolle einer ausgebildeten lexikographischen Redaktion zu vereinen. Auf der einen Seite werden bewährte Standards lexikographischer Praxis garantiert und mithilfe metalexikographischer Reflexion implementiert, auf der anderen Seite stehen die Sprecher – wie in Melchior/D'Agostini 2011 beschrieben – im Vordergrund. Deren Sprache (sowie deren sprachliches Wissen) ist Ausgangspunkt der Beschreibung im Wörterbuch, in einem – der Begriff ist durchaus positiv gemeint – zirkulären Prozess, in dem der Sprecher nicht nur Endbenutzer des „Wörterbuchprodukts“ ist, sondern Ausgangsbasis, Mitarbeiter und Nutznießer zugleich. Die Sprecher selbst sind die Experten, die ihr metasprachliches Wissen der *community*, aber auch den Wörterbuchmachern zur Verfügung stellen. Sprecher, die nicht anonym einen „Schwarm“ bilden, sondern deren (Sprach-)Biographie häufig rekonstruierbar ist, was vor idiosynkratischen Meinungen besser zu schützen hilft.<sup>12</sup> Im Rahmen dieses Ansatzes erfolgen die *Autorisierung* der Sprache und deren Beschreibung nicht durch eine übergeordnete „Autorität“ (Akademie, nationale lexikographische Institution, literarischer Kanon), sondern durch den Sprecher selbst; im Wörterbuch wird sie dann systematisiert.

Die User können auf unterschiedlichen Ebenen zum Wörterbuch beitragen.<sup>13</sup> Im Folgenden werden nur exemplarisch einige Bereiche genannt, in denen ihr Beitrag für das Wörterbuch sinnvoll und nützlich sein kann, wie sich im genannten Projekt der LEO-Wörterbücher gezeigt hat.<sup>14</sup> Das Prinzip der Halbkollaborativität ist m.E. auch jenseits dieses konkreten Beispiels auf andere Vorhaben anwendbar, wie hier verdeutlicht werden soll.

Der Beitrag der Nutzer zu einem Wörterbuchprojekt kann auf verschiedenen Ebenen angesiedelt werden:

- a) Verbesserung der vorhandenen Inhalte;
- b) Erweiterung der Inhalte;
- c) Erweiterung des Wörterbuchangebots;
- d) Weiterentwicklung des Projekts.

#### 4.1 Die Ebenen der Nutzerbeiträge

##### 4.1.1 Verbesserung der vorhandenen Inhalte

Die erste, einfachste Ebene, auf der die Nutzer beitragen können, ist nicht anders als bei den schon vorhin genannten Onlinewörterbüchern, eine inhaltliche: Die Nutzer werden aufgefordert, mögliche Fehler zu melden und damit die Inhalte des Wörterbuchs zu verbessern.<sup>15</sup> Dies

<sup>11</sup> Die Konzipierung für das Internet ist in dieser Hinsicht maßgeblich, vgl. auch Klosa (2013, S. 2): „The original form of publication affects to a great extent the online design, the degree and methods of hypertextualization, the degree of interaction with the user, the integration of multimedia, and the creation of ways of access to the dictionary.“

<sup>12</sup> Dazu vgl. Melchior (2009), Melchior/D'Agostini (2011). Zu möglichen Problemen beim Beitrag anonymer User ohne redaktionelle Überwachung vgl. Haß/Schmitz (2010, S. 4).

<sup>13</sup> Zur Interaktion mit den Usern vgl. auch Klosa (2013, S. 7f.), zur theoretischen und praktischen Stellung des Userbeitrags vgl. Gouws (2011, S. 28f.).

<sup>14</sup> In der Tat weisen die verschiedenen Wörterbücher des LEO-Projektes jedoch einen unterschiedlichen Umsetzungsgrad der Möglichkeiten auf, die das allen zugrunde liegende Halbkollaborativitätsprinzip bietet.

<sup>15</sup> Es handelt sich also um „Fehlerbehebung“ (Storrer/Freese 1996, S. 118).

geschieht auf zwei unterschiedlichen Wegen. Auf der einen Seite hat man die Möglichkeit, einfache formale Fehler, z.B. Tippfehler oder falsche Übersetzungen, die keine inhaltliche Diskussion benötigen, durch ein Fehlerformular (auch anonym) zu melden.

Die zweite Ebene, die strikt von der gerade geschilderten getrennt ist, bildet dagegen die Möglichkeit zur Meldung inhaltlicher Fehler sowohl auf der Äquivalenzebene – so in bilingualen (bzw. multilingualen) Wörterbüchern wie LEO – oder auf der Definitionsebene in einsprachigen Wörterbuchprojekten, als auch auf der Ebene der Informationen im mikrostrukturellen Aufbau der einzelnen Artikel. Diese Option, in Form einer für Diskussionen offenen Rubrik<sup>16</sup> im Forum, wird dann genutzt, wenn die Fehler nicht offensichtlich sind. Dies kann z.B. der Fall sein, wenn die angebotenen Übersetzungsäquivalente bzw. Definitionen von den Sprechern als nicht (ganz) korrekt empfunden werden, wie z.B. wenn die diasystematischen Markierungen des Lemmas als nicht korrekt betrachtet werden oder nicht mit denen des Äquivalenten übereinstimmen, oder aber wenn die mikrostrukturelle Gestaltung eines einzelnen Eintrags als nicht zufriedenstellend wahrgenommen wird. Die User können in der genannten Rubrik einen Diskussionsstrang eröffnen, in dem sie ihre Anmerkungen bezüglich des Eintrags und ihre Änderungs- bzw. Verbesserungsvorschläge, im Idealfall versehen mit Quellenangaben bzw. authentischen Beispielen, posten können. Diese Anmerkungen und Vorschläge werden dann sowohl vom lexikographischen Team als auch von den anderen Usern diskutiert, so dass ein fruchtbarer Austausch der User untereinander und mit den Wörterbuchmachern entsteht, der zu einer Verbesserung der einzelnen Einträge (z.B. geeigneter Äquivalente oder andere mikrostrukturelle Darstellung) führen kann. In den folgenden Beispielen (Abbildungen 1 und 2) aus dem Forum des LEO Deutsch-Französisch-Wörterbuch wird die Adäquatheit der Übersetzungsäquivalente aufgrund der unterschiedlichen diasystematischen Markierungen der zwei Lexeme *vioque* vs. *alt* (vgl. Abbildung 1) bzw. aufgrund der unikalen Kombinierbarkeit des Lexems *acabit* (vgl. Abbildung 2), das eine andere Distribution und Kombinatorik als das im Wörterbuch vorgeschlagene Äquivalent *Charakter* aufweist, diskutiert.

<sup>16</sup> Die Rubrik trägt den Namen „Falscher Eintrag“ (bzw. „Wrong entry“, „Entrée erronée“, „Entrada incorrecta?“ usw. in der jeweiligen Fremdsprache). Sie ist über das Forum erreichbar, wie z.B. im Falle des französischen Forums unter der URL <http://dict.leo.org/forum/previewWrongentry.php?idForum=16&lp=frde&lang=de> [Stand: 13.04.2013]. Zu den LEO-Foren und deren Rubriken, die hier nicht weiter beschrieben werden, vgl. Melchior (2012).

Falscher Eintrag \*vioque m./f. adj. - alt

Korrektur vioque m./f. adj. [fam.] - alt

Beispiele/  
Definitionen mit PONS bietet als Entsprechung "oll", "der Olle/die Olle" an.

Quellen <http://de.pons.eu/franzosisch-deutsch/vioque>

[http://www.larousse.com/de/dictionnaires/francais/vioc\\_vioque/82061](http://www.larousse.com/de/dictionnaires/francais/vioc_vioque/82061)

Kommentar Ich denke hier müsste noch präzisiert werden, dass es sich um einen Argot-Begriff handelt.  
(Auch wäre zu klären, ob er überhaupt noch verwendet wird. Mir ist er noch nie begegnet.)

Außerdem wird er sowohl als Adjektiv, als auch als Substantiv verwendet.

Verfasser Katz 2 (473027) 02 Nov 12, 12:01

Kommentar ""vioque " est toujours en usage. Heureusement, ce terme péjoratif n'est pas trop souvent utilisé.

#1 Verfasser ZucchiniMann (238313) 02 Nov 12, 15:34

Kontext/ Beispiele vioque [n. |adj.]

¶ Vieux, ancien ; vieillard, vieil homme, vieille femme, vieille prostituée ; vieux galantin qui se croit encore jeune ; parents, père

<http://www.languefrancaise.net/bob/detail.php?id=540>

Kommentar Est encore largement utilisé.

#2 Verfasser patapon (677402) 02 Nov 12, 18:13

Kommentar la correction pourrait ressembler à ça, qu'en pensez-vous?

vioque adj. m./f. fam. péj. - alt

le/ la vioque nom m./f. fam. péj.- der/die Olle fam. péj.

P.S. C'est plus souvent utilisé comme nom qu'en tant qu'adjectif.

#3 Verfasser JosephineB (455714) 03 Nov 12, 16:25

Abb. 1: <http://dict.leo.org/forum/viewWrongentry.php?idThread=1178002&idForum=16&lp=frde&lang=de>

(Stand: 13.04.2013)

Falscher Eintrag \*acabit m. - der Charakter

Korrektur de + adj. + acabit / de + déterm. + acabit [pej.] - des gleichen Schlags / von meinem (seinem)/diesem Schlag [pej.]

Beispiele/ Definitionen mit (wie jd.)

Quellen être du même acabit - vom gleichen Kaliber sein

B.- [Dans la constr. de + adj. + acabit] Fam., le plus souvent péj. Qualité, sorte :  
<http://www.cnrtl.fr/definition/acabit>

C.- Loc. De + adj. + acabit; de + déterm. + acabit. [En parlant d'une pers., rarement d'une qualité hum. ou d'une chose]  
<http://www.cnrtl.fr/definition/acabit>

de cet acabit, du même acabit → vom gleichen Schlag  
<http://dictionnaire.reverso.net/francais-allemand/acabit/forced>

de cet ou du même acabit (péjoratif) → von diesem ou gleichem Schlage  
<http://www.larousse.fr/dictionnaires/francais-allemand/acabit>

des gens de cet acabit → Leute Pl dieses Schlags  
de [ou du] même acabit péj vom gleichen Schlag  
il n'est pas du même acabit que les autres → er ist nicht aus dem gleichen Holz geschnitzt wie die anderen  
<http://fr.pons.eu/dict/search/results/?q=acabit&l=defr>

aus demselben/dem gleichen Holz (geschnitzt) sein (wie j.) → être du même acabit/de la même trempe (que qn)  
vom gleichen Kaliber sein wie j. → être du même acabit que qn  
Menschen/Leute/... meines/deines/unseres/... Schläges → des gens/... de (mon)/ton/notre/... acabit/(mon)/ton/notre/... espèce  
<http://212.227.53.204/dict/owDFc.cgi?l=0&dt=&en=acabit>

- L'argumentation sans polémique est du même acabit. → Die unpolemische Argumentation ist vom gleichen Kaliber.  
- De même, le premier devoir d'un État est toujours de protéger ses citoyens innocents contre de nouveaux attentats commis par des groupes terroristes de cet acabit. → Außerdem besteht die erste Pflicht eines jeden Staates darin, seine unschuldigen Bürger vor künftigen Angriffen derartiger terroristischer Gruppen zu schützen.  
<http://fr.bab.la/dictionnaire/francais-allemand/acabit>

Kommentar Siehe Wörterbuch: acabit

Verfasser JosephineB (455714) 07 Jun 12, 11:58

Kommentar Bemerkenswert sind auch die stilistischen Hinweise, die das CNRTL dazu gibt:

Stylistique – Les emplois A et B, vieillis ou arg., ont une vitalité extrêmement faible. D'un bon acabit est un emploi dial. (ex. 8) ou arg. et iron. (ex. 7) peu usité. Le peu de vitalité que possède acabit est donc limité à son emploi dans des constr. figées (C : de + adj. + acabit). Acabit appartient alors au lang. fam. Indiquant, plus que genre ou même sorte, une certaine méfiance ou un certain mépris, il a le plus souvent une coloration dépréci., éventuellement renforcée par le cont. immédiat (cf. l'adj. pire ex. 10), ou par des éléments répartis sur l'ensemble de la phrase (cf. ex. 11, 13, 15, 17, 18, 21, 22). Cependant le mot acabit, à lui seul, suffit le cas échéant à traduire cette nuance (cf. ex. 12). Parmi les 3 loc. figées : du même acabit, d'un autre acabit, de tout acabit, qui toutes trois sont fam., seule la 3e est orientée vers un usage nettement péj., les 2 autres prenant dans leur cont. une valeur styl. moins marquée.

#1 Verfasser Pierrot (236507) 07 Jun 12, 13:24

Kommentar Weis: Charakter, Kaliber (Mensch)  
de bon acabit = (Sache) von guter Qualität  
= (Mensch) tadellos, einwandfrei

In gewisser Weise hat die Übersetzung "Charakter" seine Berechtigung, sie kann aber nicht alleinstehen.

vom gleichen Kaliber oder vom selben Kaliber?  
(Bin mir da nicht ganz schlüssig.)

#2 Verfasser hannabi (554425) 08 Jun 12, 11:31

Abb. 2: <http://dict.leo.org/forum/viewWrongentry.php?idThread=1155008&idForum=16&lp=frde&lang=de>  
(Stand: 13.04.2013)

In beiden Fällen wird nicht der denotative Charakter der Übersetzungsäquivalente in Frage gestellt, sondern konnotative und kombinatorische Merkmale. Dies verlangt von den Sprechern hohes Sprachbewusstsein und metasprachliches Wissen. Die Interaktion der User untereinander und mit den Lexikographen sowie die Diskussionsbeiträge der *community*, die stets mit authentischen Beispielen bzw. Quellen belegt sind, stellen den Wörterbuchmachern begründete metasprachliche Daten zur Verfügung, die, in ihrer Zusammenschau einen überindividuellen Mehrwert erhalten, der idiosynkratische Bewertungen zu vermeiden hilft, und die es ihnen ermöglichen, über eventuelle Änderungen zu entscheiden und diese ggf. in geeigneter Form vorzunehmen, um die Qualität eines lexikographischen Eintrags zu verbessern. Vor allem im Bereich der diasystematischen Markierungen helfen die Diskussionen im Forum dem Lexikographen, eventuelle Verschiebungen im Diasystem der Sprache nachzuvollziehen und dienen als „*tertiäre Quelle*“, da sie metasprachliche Urteile und andere Angaben liefern, die für die adäquate linguistische und lexikographische Aufarbeitung des Sprachmaterials von großer Bedeutung sein können.

#### 4.1.2 Erweiterung der Inhalte

Die User können darüber hinaus auf einer zweiten Ebene aktiv sein und Wörterbuchinhalte erweitern. Dies kann ebenfalls auf unterschiedliche Weise geschehen. Die erste Möglichkeit dafür ist die Schenkung bzw. die Bereitstellung von eigenen Glossaren und Wortlisten – vielfach im fachsprachlichen Bereich, in dem die jeweiligen User Spezialisten sind –, die in die Datenbank integriert werden. Hier wird deutlich, dass der halbkollaborative Ansatz sich von der Kollaborativität unterscheidet: Die Wortlisten und Glossare werden zur Verfügung gestellt, sie müssen aber von der Wörterbuchredaktion an die formalen und inhaltlichen Richtlinien und Anforderungen des Wörterbuchs angepasst werden. Das heißt einerseits, dass die Mikrostruktur der Einträge mit den nötigen und erwünschten Informationen ergänzt wird, andererseits, dass die Angemessenheit und Korrektheit der Übersetzungsäquivalente vom Wörterbuchmacher durch Vergleich mit anderen lexikographischen Werken, mit linguistischen Arbeiten oder durch Introspektion überprüft wird *bevor* die neuen Einträge aufgenommen werden. Die Überprüfung und eventuelle Korrektur erfolgt daher in diesem Fall nicht *a posteriori* durch die *community* oder durch die Redaktion, sondern schon *a priori* durch die Wörterbuchredaktion.

Als zweite Möglichkeit ist die Verwendung der Rubriken im Forum zu werten, die für den Vorschlag fehlender Einträge vorgesehen sind. Dabei kann es sich um allgemein- oder fachsprachliche Einträge handeln, die im Wörterbuch noch nicht vorhanden sind.<sup>17</sup> Dadurch, dass die Vorschläge mit Quellen bzw. anderen Verweisen belegt werden müssen, und dass dies häufig durch die Angabe (als Verweis und/oder als Zitat) weiterer lexikographischen Quellen geschieht, erleichtern die User den Zugang zu sekundären Quellen, die für die Arbeit am Wörterbuch unabdingbar sind (selbst wenn sie ihn damit nicht gänzlich ersetzen können). An einem Beispiel (siehe Abbildungen 3a und 3b) – wiederum aus dem französisch-deutschen Forum – zeigt sich in der Diskussion der User bezüglich der Berechtigung eines Eintrags das Potenzial des halbkollaborativen Ansatzes deutlich: Die Existenz des Eintrags in verschiedenen lexikographischen Quellen wird nicht von allen Usern als ausreichend für die Einbindung in LEO betrachtet, so dass zusätzliche Belege gesucht werden. In der Diskussion werden von den Usern (als Laien!) grundsätzliche Prinzipien der lexikographischen Praxis kritisch beleuchtet.

<sup>17</sup> Es handelt sich nicht nur um „Schließung von Lemmalücken“ (Storrer/Freese 1996, S. 118), sondern auch um die Erweiterung der vorhandenen Artikel, wenn z.B. im Falle von Polysemie bzw. von unterschiedlichen Lesarten eines Lexems in der jeweiligen anderen Sprache entsprechende Äquivalente fehlen.

**Neuer Eintrag \*fumidôme m. [tech.] - die Rauchabzugskuppel**

Beispiele/ Langenscheidts Taschenwörterbuch 1982  
 Definitionen mit fumidôme / Rauchabzugskuppel  
 Quellen

Kommentar pour laisser "envoler" la mousse du fumiste :)

Verfasser stevelefrancais (725002) 25 Jun 12, 20:48

Kommentar Très joli mot, qui n'existe que dans quelques dicos.  
 Pas dans la vraie langue.  
 Non soutenu.

#1 Verfasser Pauvre Lelian (719620) 25 Jun 12, 20:53

Kontext/ Beispiele ftp://usagers.ville-nice.fr/MAPA/916/CCTP.pdf p. 11 CAHIER DES CLAUSES TECHNIQUES, REHABILITATION DE LA SALLE BRANCOLAR  
 5.2.2.3 : LANTERNEAUX DE DESENFUMAGE  
 Fourniture et pose d'un lanterneau de type fumidôme avec coffret à commande manuelle pour l'action d'une manœuvre de désenfumage pneumatique, type CA1 de chez Comtra ou équivalent  
<http://www.klekoon.com/boamp/boamp-appels-offres-extension-aerogare-aeroport-bez...> Détail de l'appel d'offre, Extension de l'aérogare de l'aéroport de Béziers-Agde-vias à Béziers:  
 Lot(s) 01. - gros-œuvre - démolition - maçonnerie - plâtrerie.  
 Extension de la salle d'embarquement de l'aérogare (77 mètres carrés), VRD, espaces verts (90 mètres carrés, toiture (fumidôme)  
<http://positif-sep.heberg-forum.net/topic147-0.html>  
 Bon, cette année on devrait consommer moins, j'ai dans la salle de bain un "fumidôme" (un truc pour évacuer les fumées en cas d'incendie) et ce truc n'était pas hermétique ...  
<http://termiumplus.gc.ca/tvp2alpha/alpha-eng.html?lang=eng&srctxt=plafond&i=1&i...>  
 fumidôme Source 414, record 102, fumidôme  
 CORRECT, MASC  
 DEF – Coupole pratiquée dans un plafond et destinée à l'évacuation (smoke dome)  
[http://termiumplus.gc.ca/tvp2source?lang=fra&srctxt=plafond&i=1&index=frr&src\\_i...](http://termiumplus.gc.ca/tvp2source?lang=fra&srctxt=plafond&i=1&index=frr&src_i...)  
 1. fumidôme  
 Dictionnaire de termes nouveaux des sciences et des techniques / sous la direction de G. Quemada. –  
 Paris : Conseil international de la langue française, (c1983).  
 xix, 605 p.; Comprend des termes en anglais, allemand et espagnol.; Publié conjointement avec l'Agence de coopération culturelle et technique.; ISBN 2853191117.

Kommentar rare mais un terme utilisé en France et au Canada.

#2 Verfasser stevelefrancais (725002) 26 Jun 12, 10:28

Kommentar Steve, ich unterstütze deinen Vorschlag. Wenn ich deine Belege sehe und zudem lese, dass PL dem Wort bescheinigt, in einigen Wörterbüchern zu stehen, dann sehe ich nicht ein, warum es, wie in 1 angeregt, aus LEO herausgehalten werden soll. Was ist schon "la vrai langue"?

#3 Verfasser mars (236327) 26 Jun 12, 12:03

Abb. 3a: <http://dict.leo.org/forum/viewWrongentry.php?idThread=1158126&idForum=17&lp=frde&lang=de>

(Stand: 13.04.2013)

Kontext/ Beispiele <http://www.google.com/search?q=%22Rauchabzugskuppel%22&hl=de#q=%22Rauchabzugskup...>

Kommentar PL hat "la vraie langue" geschrieben.

Auch im Deutschen nur 32 Treffer. Davon etwa die Hälfte von LEO und anderen Online-Dicos wie proz.com, dictionariste, quickdict. etc.

#4 Verfasser Ceesem (719060) 26 Jun 12, 12:33

Kommentar Es ist immergefährlich zu behaupten, dass ein Wort in einer Sprache nicht existiere. Die Tatsache, dass ein Wort in google nicht oft zu finden ist, heisst ja nicht, dass es nicht existiert oder nicht gebraucht wird. Speziell wenn Fachwörter ins Spiel kommen, dann sind Google oder andere Suchmaschinen kein sicherer Hinweis darauf, dass Wörter nicht existieren oder nicht gebraucht werden.

Als Beispiel diene der Meißel (ciseau, burin)  
 Flachmeißel (Leo: ciseau/burin plat) google ok  
 Spitzmeißel (Leo: -) google ok  
 Fugenmeißel (Leo: -) google ok  
 Fliesenmeißel (Leo: -) google ok  
 Maurermeißel (Leo: -) google ok  
 Nutenmeißel (Leo: -) google ok  
 Aushaumeißel (Leo: -) google 411  
 Trennstemmer (Leo: -) google ok  
 Kreuzmeißel (Leo: bédane m. aussi: bec-d'âne/burin en croix [techn.]  
 ) google: Bédane 85'000, burin en croix 711  
 Scharriereisen (Leo: -) google ok  
 Bossiereisen (Leo: -) google ok  
 Schlageisen (Leo: -) google ok  
 Zahneisen (Leo: -) google ok  
 Sprengeisen (Leo: -) google ok

Es handelt sich hier um die gängige Basisausstattung eines Tischlers oder eines Steinhauers (und vieler verwandter Berufe). Solche Meissel - und noch viel mehr - werden weltweit millionenfach produziert und verwendet und natürlich v.a in der Fachsprache täglich unzählige Male gebraucht. Bei Google findet man diese im Bereich von zigtausend Einträgen oder manchmal auch nur eher selten (in Leo nur ausnahmsweise), eventuell auch noch falsch übersetzt. Sind sie deshalb **pas de la vraie langue?**

#5 Verfasser stevelefrancais (725002) 27 Jun 12, 10:34

Abb. 3b: <http://dict.leo.org/forum/viewWrongentry.php?idThread=1158126&idForum=17&lp=frde&lang=de>

(Stand: 13.04.2013)

Ähnlich ist der Fall, wenn fachsprachliche Einträge von Spezialisten vorgeschlagen werden und sich dann im Forum Diskussionen über die Angemessenheit der vorgeschlagenen Lexeme und der entsprechenden Äquivalente entspinnen. Das Fachwissen der Experten im jeweiligen Fachbereich und ihre (häufig stark ausgeprägte) sprachliche Kompetenz in den betreffenden Sprachen helfen dem Lexikographen, das Wörterbuch um Fachlemmata zu erweitern, zu denen im Wörterbuchaußentext „Forum“ zusätzlich einsprachige Definitionen, fachliche Diskussionen,<sup>18</sup> Verlinkungen zu externen Inhalten (encyklopädische Einträge, Abbildungen, Fachtexte) zur Verfügung stehen.<sup>19</sup> Sprachliches Material, dessen Auswertung lexikographisch häufig interessant ist, liefern die User im Forum auch auf indirekte Weise in ihren Beiträgen in den unterschiedlichen Diskussionssträngen. Am Sprachgebrauch der User, die in diesem Fall richtige „Informanten“ sind, kann man unterschiedliche Entwicklungen (Sprachwandelphänomene aller Art) der Sprache beobachten und z.B. evtl. Neologismen, aber auch Bedeutungserweiterungen und -verschiebungen feststellen, die in anderen lexicographischen Werken noch nicht verzeichnet sind, und nach denen man mühsam in Korpora suchen müsste. Es stellen sich in diesem Fall auch nicht die methodologischen und ethischen Fragen, die Marcoccia (2011) bezüglich der Verwendung von Internetforen als „Korpora“ für die linguistische Forschung stellt:<sup>20</sup> Die User wissen, dass die Wörterbuchredaktion ihre Beiträge liest und auswertet bzw. lesen und auswerten kann. Der Kontext, in dem die

<sup>18</sup> Die Diskussionen zwischen Fachexperten verlassen jedoch teilweise die sprachliche Ebene und werden nur auf der Fachebene fortgeführt, so dass sie für die Belange des Wörterbuchs und seiner User weniger interessant, wenn nicht gar schwer zugänglich werden (vgl. dazu Melchior 2012, S. 360ff.).

<sup>19</sup> Zu den Möglichkeiten und Gefahren von Verweisen und von der Verwendung externen Daten vgl. u.a. Heid/Prinsloo/Bothma (2012).

<sup>20</sup> Allgemeiner zu ethischen Fragen im Umgang mit Daten aus dem Internet vgl. Markham/Buchanan (2012).

Beiträge entstanden sind, ist dem Wörterbuchmacher, der die Diskussionen verfolgt, bekannt. Auch das mögliche Eintreten des Beobachter-Paradoxons wird vermieden, da die User, obwohl sie um die Präsenz des Lexikographen wissen (und dieser also nicht „en cachette“ (Marcoccia 2011, S. 56) ist), der sie teilnehmend beobachtet, ihre Sprachverwendung bzw. ihre metasprachlichen Urteile nicht deswegen *ad hoc* anpassen. Die auf diese Weise gewonnenen Materialien können allerdings teilweise unsystematisch sein, so dass ihr Beitrag in gewissem Maße eher dem Zufall überlassen und nicht planbar ist. Dies verringert den Wert der einzelnen Beiträge für die lexikographische Arbeit aber nicht.<sup>21</sup>

#### 4.1.3 Erweiterung des Wörterbuchangebots

Wenn man die Ebene der Erweiterung bzw. Verbesserung des Lemmataumfangs und der Qualität der einzelnen Artikel verlässt, kann der Blick auf eine weitere Ebene gerichtet werden: Die User können auch einen Beitrag zur Erweiterung des Wörterbuchangebots liefern, und zwar auf vielfältige Art und Weise. Einige Beispiele mögen dies exemplarisch illustrieren.

Einen ersten wichtigen Beitrag können die User leisten, indem sie Informationen liefern, welche für andere Nutzer interessant sind, aber keine notwendigen Angaben in einem bilingualen Wörterbuch darstellen, wie z.B. im Bereich der Wortetymologie (und der Wortgeschichte) bzw. zur Herkunft von Redensarten und anderen sprachlichen Elementen. Abbildungen 4a und 4b zeigen ein Beispiel aus dem Forum des LEO-Wörterbuchs Deutsch-Französisch, zu finden in der metasprachlich ausgerichteten Rubrik „Sprachlabor/Labo de langues“ (ähnliche Beispiele finden sich aber auch in Diskussionssträngen, die nicht explizit der Etymologie gewidmet sind).

Sujet	<b>Etymologie von Wörtern, Redensarten, Namen</b>
Commentaire	Auf vielfachen Wunsch wird nun die Serie der fünf bisherigen Etymologiefäden fortgesetzt. Im Fadentitel wollte ich andeuten, dass es nicht nur um die Etymologie einzelner Wörter geht, sondern auch um die von Vor- und Familiennamen sowie um die Herkunft von manchen Redensarten wie dieser: „Das geht auf keine Kuhhaut.“  Nach diesen Etymologien könnt ihr fragen, oder ihr teilt sie einfach als interessanten Fund mit. Das darf ruhig wie <i>Kraut und Rüben</i> durcheinandergehen.  Als besonders kommunikativ hat es sich auch bewährt, passende Etymologien als Rätsel zu verpacken. Für das Procedere schlage ich wieder vor: Wer glaubt, die Antwort gefunden zu haben, nennt sie nicht gleich direkt, sondern deutet sie nur an, um den anderen das Weitersuchen / Weiterspekulieren zu ermöglichen. Und wer die Rätselfrage gestellt hat, erläutert sie abschließend selbst, und zwar zu einem Zeitpunkt, den er für richtig hält, in der Regel nicht früher als 48 Stunden nach dem ersten Impuls.  Gewisse Beiträge in vorangegangenen Fäden (Siehe auch: 5. Etymologiefäden - Siehe auch: 4. Etymologiefäden - Siehe auch: 2. Etymologie-Fäden - Siehe auch: Etymologie - Siehe auch: ETYMOLOGIE ), legen es mir nahe, deutlich Folgendes zum Ausdruck zu bringen: In diesen Etymologie-Fäden sind alle „Beiträge“ unerwünscht, die mit Wortgeschichte nichts zu tun haben. Wer sich aus Etymologie nichts macht, braucht ja in den Faden gar nicht hineinzuschauen; er findet für seine Interessen genug andere Fäden in diesem Forum.
Auteur	<b>mars</b> (236327)
	29 Aug 11, 09:06
Commentaire	Beginnen möchte ich mit der Etymologie eines Familiennamens, nämlich mit dem der angeblich mächtigsten Frau der Welt, Andrea Merkel:  Merkel ist eine mit i-Suffix versehene Koseform von Rufnamen, die mit <i>marcha</i> gebildet sind, wie zum Beispiel von Markwardt aus ahd. <i>marcha wārt</i> , was soviel bedeutete wie „im Grenzland ein Schutz“. Dass Angela Merkel zwar in Hamburg geboren wurde, aber im Landkreis Uckermark aufwuchs und heute auch <i>Schutzwart</i> dieser Gegend betrachtet werden kann, ist so gesehen ein netter Zufall.
#1	Auteur <b>mars</b> (236327)
	29 Aug 11, 09:07

Abb. 4a: <http://dict.leo.org/forum/viewGeneralDiscussion.php?idThread=1098435&idForum=10&lp=frde&lang=fr>  
(Stand: 13.04.2013)

<sup>21</sup> Die Wörterbuchredaktion hat in diesen (nicht moderierten) Foren Zugang zu allen Beiträgen, auch zu denen, die eventuell *a posteriori* entfernt bzw. unsichtbar gemacht werden, so dass die Kontextualisierung der einzelnen Sprachmaterialien immer nachvollziehbar ist.

Commentaire Herzlichen Dank, liebe Mars, für den neuen Etymologiefaden.

Wie sehr ich ihn schätze: Ich habe an ihm einen Narren gefressen :-))

**"einen Narren an jemandem gefressen haben"**

**Bedeutung:** umgangssprachlich: jemanden sehr gern haben, in jemanden geradezu vernarrt sein, jemanden bevorzugen.

**Herkunft:**  
Es existieren unterschiedliche Erklärungen zur Herkunft dieser Wendung:  
Es kann sein, dass sie sich von den Hofnarren früherer Zeiten herleitet. Diese standen weit oben in der Gunst des Herrschers und konnten sich allerlei erlauben. Der Herrscher hatte also, einen Narren an ihm gefressen und sich somit die Narrheit seines Günstlings beinahe selbst einverlebt.  
Möglich ist es aber auch, dass die Redensart sich von alten Vorstellungen von Besessenheit herleitet. Man glaubte zum Beispiel, dass Geisteskranke von Dämonen, den Narren, besessen seien, die sich in ihren Körpern befänden.

**Synonyme:**  
einen Affen an jemandem gefressen haben  
jemandem ans Herz gewachsen sein  
eine Schwäche für jemanden haben

Darf es noch ein ganz berühmter Narr aus der Opernliteratur sein,  
an dem ich auch einen Narren gefressen habe:  
**"Freundlich blick ich ..... auf diesen Faden :-))**  
[http://www.youtube.com/watch?v=NWwg\\_VtrTdg](http://www.youtube.com/watch?v=NWwg_VtrTdg)

Wie würde man diese Redewendung übersetzen können ?

Herzlichst oopsy

#2 Auteur [oopsy](#) (491382) 29 Aug 11, 09:22

Commentaire merci mars pour ce nouveau fil d'étymologie tant attendu !

oopsy - je pense que la traduction de "einen Narren an jemandem gefressen haben" est "avoir le béguein pour quelqu'un"

et ma première question à mars dans ce fil : d'où vient le mot béguein ? (ce n'est pas une devinette !!!)

#3 Auteur [kitine](#) (633017) 29 Aug 11, 10:44

Commentaire Servus Kitine! Vielen herzlichen Dank für Deine schnelle Antwort! Allerherzlichst oopsy

#4 Auteur [oopsy](#) (491382) 29 Aug 11, 11:14

Commentaire Kitine,

ich freue mich über deine Frage, denn es ist in einem Faden wie diesem immer befriedigender, eine Anfrage zu beantworten als Etymologien im Verlautbarungsstil einfach bekanntzugeben.

Wie es von den *Beguinen* zu dem Ausdruck *avoir le béguein pour qn* kam, wird so erklärt:

"Au XIIe siècle, le "béguein" était une coiffure féminine qui fut à l'origine de l'ordre religieux des Béguines. Plus tard, il s'agira d'une coiffe que l'on faisait porter aux enfants sous leur bonnet. C'est d'ailleurs de là que naîtra l'expression "être coiffé de quelqu'un", signifiant "être aveuglé par quelqu'un", comme si le béguein était trop basse sur les yeux, empêchant celui qui le porte d'y voir clair. Par extension, l'expression est devenue "avoir le béguein pour", et dont la traduction serait "être fasciné par une personne, être amoureux". Une autre origine possible est l'expression "avoir le béguein à l'envers", qui signifierait que le bonnet est tellement mal placé sur le crâne qu'on en a la tête à l'envers. Prise dans son sens figuré, cette expression pourrait également se traduire par "être amoureux". (<http://www.linternaute.com/expression/langue-francaise/155/avoir-le-beguin-pour-...>)

Zur Etymologie von *bégueine* selbst kann ich aus dem Wörterbuch meines verehrten Lehrers Gamillscheg zitieren: *bégueine* ist fem. Neubildung zu mittelfrz. *begard* "Bettelmönch", dieses aus mndl. *beggaert* "Bettler" (siehe auch engl. *beggar* bzw. das satirische Singspiel *The Beggar's Opera* von John Gay und die danach benannte Progressive-Rock-Gruppe aus Schottland).

Bégueine ist dann ins Niederländische zurückgewandert, wie z. B. *tunnel*, im Englischen nach *la tonnelle* gebildet, ins Französische 'zurückkehrte'.

#5 Auteur [mars](#) (236327) 29 Aug 11, 11:25

Abb. 4b: <http://dict.leo.org/forum/viewGeneralDiscussion.php?idThread=1098435&idForum=10&lp=frde&lang=fr>

(Stand: 13.04.2013)

Die Etymologie stellt einen Bereich der Sprachwissenschaft dar, der für die Sprecher besonders interessant zu sein scheint, so dass solche Diskussionsstränge teils über Jahre gepflegt und erweitert werden. Sowohl reine etymologische als auch wortgeschichtliche Angaben sind jedoch in vielen Wörterbuchtypen keine notwendigen Angaben. So werden sie in zweisprachigen (Gebrauchs-)Wörterbüchern in der Regel nicht angeführt (sie werden auch von Haussmann/Werner (1991) nicht erwähnt), da sie nicht funktional für die Zwecke derselben sind. Die Beiträge der Nutzer stellen daher eine wichtige und interessante Ergänzung des Wörterbuchsangebots dar, die meist mit Quellenangaben und Belegen gut dokumentiert ist. Für den Linguisten liefern darüber hinaus einige Beiträge in diesen Diskussionen auch nützliche Informationen über mögliche volksetymologische Deutungen.<sup>22</sup>

Einen weiteren Beitrag leisten die User, indem sie z.B. Fragen zu grammatischen Aspekten beantworten: Dadurch werden die Außentexte des Wörterbuchs ergänzt, die dazu dienen, auch Aspekte der Grammatik zu behandeln, die sonst nicht in voller Breite dargestellt werden könnten. Abbildungen 5a und 5b illustrieren ein solches Beispiel, wieder aus der Rubrik „Sprachlabor/Labo de langues“ des Französisch-Deutsch-Wörterbuchs.

Betitell Conjugaison - I. Gruppe, Anzahl der Unregelmäßigen...

Kommentar Hallo wertes Forum,

hab grade meinen Bescherelle aus dem Regal gezogen, um mir die Konjugationsformen wieder ein wenig in Erinnerung zu holen. Dabei ist mir aufgefallen, dass es im Bescherelle zwar eine annähernd vollständige Liste der Verben gibt, die in die dritte Gruppe fallen, jedoch keine Liste für die 2. Gruppe und für die Verben der 1. Gruppe die nicht nach dem Schema von -aimer- konjugiert werden (in meinem Bescherelle sind das genau 11. Untergruppen, konkret die Nummern 7-18).

Für die zweite Gruppe habe ich jetzt beim Wiktionnaire eine Liste gefunden:  
[http://fr.wiktionary.org/wiki/Cat%C3%A9gorie:Verbes\\_du\\_deuxi%C3%A8me\\_groupe\\_en\\_f...](http://fr.wiktionary.org/wiki/Cat%C3%A9gorie:Verbes_du_deuxi%C3%A8me_groupe_en_f...)

Für die Verben der 1. Gruppe, die nicht wie -aimer- konjugiert werden, habe ich jedoch keine Liste gefunden. Kann mir da jemand von euch eventuell weiterhelfen?

Merci bien,

Jan

Verfasser Jan Hase (247478) 18 Aug 10, 00:27

Kommentar Mehr locker nehmen

1.  
Zuerst ein Danke für den Link. Zum ersten Mal in meinem Leben sehe ich alle regelmässigen ir-Verben auf einer Liste. Es gibt also davon 765, ich hätte eher um die 600 getippt. Wenn mir also das Verb „raffolir“ begegnet, dann brauch ich nur hier nachschauen und sehe, dass es ein regelmässiges ir-Verb ist. Also „nous raffolissons“ und nicht „nous raffolons“, was es ja auch gibt.  
Die Verbenbüchlein werden nie so etwas ausführlich angeben, sondern konjugieren das Verb „finir“ durch und schreiben: Alle anderen regelmässigen ir-Verben nach diesem Modell.

2.  
Es gibt schätzungsweise 5000 er-Verben, das sind die allermeisten. Nur ein er-Verb ist vollständig unregelmässig (aller), die anderen werden als regelmässig betrachtet, wenngleich bei einigen wenige Umformungen vorzunehmen sind z.B. bei ger-Verben ein eingeschobenes „e“,  
nous mang e ons, je mang e ais  
bei cer-Verben ein c cédille  
nous commençons, je commençais

Gewöhnlich wird in Verbenbüchlein das Verb „chanter“ durchkonjugiert als Modell und ein ger-, cer-Verb usw. als Musterbeispiel. „Aimer“ ist nicht so 100% regelmässig, denn es fängt schon gut an mit dem elidierten „j'aime“. Ich glaube nicht dass es irgendwo eine vollständige Liste aller er-Verben gibt. Aber man soll die Hoffnung nicht aufgeben.

Es tut mir leid, dir nicht weiter helfen zu können und nur diese paar Allerwelts-Informationen gebracht zu haben.

#1 Verfasser Josef - Joseph (324940) 18 Aug 10, 06:32

Abb. 5a: <http://dict.leo.org/forum/viewGeneralDiscussion.php?idThread=970363&idForum=10&lp=frde&lang=de>  
(Stand: 13.04.2013)

<sup>22</sup> Diese ermöglichen wiederum Rückschlüsse auf die Strukturierung des mentalen Lexikons der Sprecher und über die Zuordnung von Wörtern zu Wortfamilien.

Kommentar Hallo Joseph,

vielen Dank für deine Antwort. Ich fand die Liste auch sehr hilfreich, der Bescherelle verhält sich da ja eher abstrakt. In meinem gibt es zwar im Anhang eine Liste von 12000 Verben und mit der dazugehörigen Zuordnung zu den 82 Konjugationsschemata, die die 3 Gruppen aufweisen - diese Liste ist aber alphabetisch und eben für alle Verben aus den 3 Gruppen. Für die 3. Gruppe gibt es dagegen zusätzlich eine Liste der entsprechenden Verben, mit numerischer Zuordnung zu den entsprechenden Konjugations schemata. Nicht aber für die zweite Gruppe und die Verben aus der 1. Gruppe, die abweichend von -aimer- konjugiert werden.

Das kann einen ja auch vor ein Problem stellen:

Ersten finde ich es mitunter recht schwierig, anhand der Struktur eines beliebigen Verbs aus der 1. Gruppe zu wissen, ob es sich um ein Verb handelt, dass wie -aimer- oder im Sinne einer der abweichenden Verben konjugiert wird. So finde ich es schwierig, mir zu merken, das alle Verben, die auf -ecer, -emer, -eper, -eser etc. etc. enden, wie -peser- konjugiert werden. Eine Liste mit den entsprechenden Verben, aus der ich mir nur die für mich relevanten Verben fische, wäre für mich praktischer.

Zweitens, sind nicht alle Angaben des Bescherelle stringent: So wird angegeben, dass viele Verben auf -tir wie -sentir- konjugiert werden. Wenn ich jetzt ein Verb wie -amortir- vor mir habe, kann ich nicht sofort eindeutig sagen, ob es zur dritten oder zweiten Gruppe gehört (es gehört zur zweiten, wird also nicht wie -sentir- konjugiert). Wenn ich allerdings immer mal wieder die Listen studieren würde, würde mir sowas leichter fallen.

Hm, vielleicht mache ich da aber auch grade ein wenig zu sehr Wissenschaft draus :-). Wenn aber jemand eine Liste der Verben hat, die nach den Verbparadigmen 7-18 aus dem Bescherelle konjugiert werden - immer her damit!

Beste Grüße,

Jan

#2 Verfasser JanHase (247478) 18 Aug 10, 10:55

Kommentar Und vielen Dank nochmal für deine Ausführungen Joseph, die ich hier im Eifer des Gefechts nicht richtig gewürdigt habe!

#3 Verfasser JanHase (247478) 18 Aug 10, 11:00

Abb. 5b: <http://dict.leo.org/forum/viewGeneralDiscussion.php?idThread=970363&idForum=10&lp=frde&lang=de>  
(Stand: 13.04.2013)

Eine Erweiterung des Wörterbuchangebots, die von den Usern geleistet wird, betrifft außerdem dialektale Elemente, insbesondere Lexeme bzw. Phraseologismen, die den primären Dialekten entstammen. Solche Elemente finden keinen Platz in einem auf die Gemeinsprache orientierten Wörterbuch, können jedoch für die Nutzer (und auch für die User) durchaus von Bedeutung sein.<sup>23</sup> Die Beiträge im Forum – insbesondere in den Rubriken, die dafür vorgesehen sind<sup>24</sup> – können das Bedürfnis der Nutzer und User nach Äquivalenten bzw. nach geeigneten Übersetzungen befriedigen, ohne dass im Wörterbuch selbst dialektale Einträge vorhanden sind. Die User, die vielfach Sprecher der jeweiligen Dialekte sind, können in solchen Fällen Hilfe leisten – was dem jeweiligen Lexikograph nicht immer möglich wäre. In Abbildung 6 ein Beispiel aus dem Italienisch-Deutsch-Wörterbuch:

<sup>23</sup> Viele Nicht-Muttersprachler, vor allem wenn mit der jeweiligen Fremdsprache nicht sehr vertraut, können häufig zwischen Elementen der primären, der tertiären Dialekte bzw. der Gemeinsprache nicht unterscheiden, sie suchen daher auch dialektale Elemente im Wörterbuch.

<sup>24</sup> Die jeweils „Deutsch gesucht“ bzw. „Englisch gesucht“, „Französisch gesucht“ usw. heißen.

Betreff **tropo fore te capu**

Quellen Hallo,  
kann mir jemand helfen und diesen Satz übersetzen, bitte!?  
"tropo fore te capu"  
Ich wäre euch wirklich sehr dankbar!!!  
Danke...

Verfasser letizia 06 Jan 10, 15:41

Übersetzung tropo fore te capu

Quellen es ist dialekt, aber "fuori di testa" ist Italienisch, es bedeutet "Verrückt" aber in eine gute sinn, keine böse sache.

#1 Verfasser Maty 06 Jan 10, 16:19

Abb. 6: <http://dict.leo.org/forum/viewUnsolvedquery.php?idThread=861770&idForum=29&lp=itde&lang=de>  
(Stand: 14.04.2013)

Darüber hinaus leisten die User ihren Beitrag auch beim Auffinden von Neologismen, die noch nicht ins Wörterbuch aufgenommen worden sind. Die Tatsache, dass diese – versehen mit Quellenangaben – in den Foren gepostet werden, erlaubt es dem Wörterbuchmacher, auch abzuwarten, ob sie sich durchsetzen oder ob sie schnell aus dem Gebrauch verschwinden. Dies ist z.B. bei Bezeichnungen für neue außersprachliche Referenten (also Denotate) der Fall, die häufig aus dem Anglo-Amerikanischen stammen, und für die in der Presse unterschiedliche Übersetzungsalternativen vorgeschlagen werden (z.B. für *swine flu*, das in Zeiten der Pandemiegefahr im Italienischen unterschiedliche Übersetzungen fand, von denen einige sich nicht durchsetzen konnten, vgl. dazu Melchior 2012, S. 363). Das Forum als Wörterbuchaußentext bietet den Vorschlägen für die Übersetzung einen Platz, bevor die neuen Einträge ins Wörterbuch selbst Eingang finden.

#### 4.1.4 Weiterentwicklung des Projekts

Auch im Bereich der Weiterentwicklung des lexikographischen Projektes können die User auf unterschiedliche Art und Weise einen wertvollen Beitrag leisten. Dazu zählen zuerst einmal das Testen neuer Versionen des Wörterbuchs (bzw. neuer Wörterbuchprojekte), wozu die User ausdrücklich im Forum bzw. auf der Ausgangsseite des Projektes aufgefordert werden. Das Userfeedback wird ausgewertet bevor ältere Versionen ersetzt und für die breite Öffentlichkeit zugänglich gemacht werden, wie Abbildung 7 zu entnehmen ist, oder auch nach dem Launch neuer Wörterbuchversionen, damit eventuell verbliebene Probleme gemeldet werden können (vgl. Abbildung 8).

**Neue Version des Wörterbuchangebots** 21. Januar 2013

Nach längerer Pause melden wir uns mit einer in den vergangenen Monaten gereiften neuen Version unseres Angebots zurück, in den nächsten Tagen erst einmal auf unserem **Testserver**. Unsere Bitte: Testen Sie die neue Version und schicken Sie uns Kommentare dazu, bevorzugt über die dafür eingerichtete **Forumsdiskussion**, gerne aber auch per **E-Mail**. So können wir bis zur geplanten Online-Stellung Anfang Februar noch Störendes und Fehlerhaftes beseitigen.

Die wichtigsten Änderungen in Schlagworten:

- **flotteres Erscheinungsbild**, statt des in die Jahre gekommenen, leicht "kantig" wirkenden; die Eingabemaske ist größer geworden und zentraler platziert, die weiteren Funktionen haben wir versucht "aufgeräumter" anzurichten,
- während des Eintippens eines Suchbegriffs werden nach Häufigkeit gewichtete **Wortvorschläge** angezeigt (wie sie den Nutzern unserer Smartphone-Apps bereits vertraut sind),
- zwei neue Wörterbücher erweitern das bisherige Angebot: **Deutsch-Portugiesisch** und **Deutsch-Polnisch**,
- bei den Suchergebnissen werden Begriffe, die allgemein als **Grund- und Aufbauwortschatz** angesehen werden, **am Anfang der Ergebnisliste** angezeigt,
- für **flektierte Formen** von Wörtern wird bei den Ergebnissen nicht wie bislang nur auf deren mögliche Grundformen verwiesen, es werden zusätzlich **Treffer für Grundformen** mit ausgegeben,
- **Audio und Flexionstabellen** können direkt **per Klick** abgerufen werden, nicht mehr nur über die **Zusatzinformationen**; letztere werden nun **für jeden Treffer insgesamt** (linke und rechte) Seite angezeigt, nicht mehr für jede Seite getrennt,
- die für eine Suche ausgelieferten **Forumartikel** werden **mit Kontext** ausgeliefert, über den sich schneller erkennen lässt, ob eine Diskussion von Interesse ist.

Die Palette der weiteren Neuerungen ist groß und betrifft ebenso das Forum und den Trainer. Wir hoffen aber, dass wir den geneigten Nutzer mit den oben aufgeführten Aspekten so neugierig gemacht haben, dass sie oder er bei Gelegenheit einen Blick darauf wirft.

Viel Spaß beim Erkunden  
Ihr LEO-Team

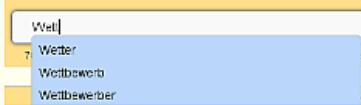
Abb. 7: <http://dict.leo.org/frde?lp=frde&search=> (Stand: 23.01.2013)

**Neue Wörterbuchversion**

Um weitere Meinungen einzuholen zur neuen Version, haben wir am 8. Februar eine letzte Testphase gestartet. Zufallsgesteuert wurden dafür Adressen ausgewählt, deren Nutzer bereits die neue Wörterbuchversion erhalten.

Um den größtmöglichen Nutzen für Sie und uns zu erreichen, bitten wir Sie, uns Fehler oder Sie störendes Verhalten zu melden damit wir ggf. Nachbesserungen bis zum eigentlichen Start vornehmen können. Sollte die neue Version aus irgend einem Grund mit Ihrem Browser nicht funktionieren, so können Sie über [dict.leo.org](http://dict.leo.org) zur vertrauten Version zurückgelangen.

Die wichtigsten Änderungen in Schlagworten:

- moderneres Erscheinungsbild statt des in die Jahre gekommenen, leicht "kantig" wirkenden; die Eingabemaske ist größer geworden und zentraler platziert, die weiteren Funktionen haben wir versucht "aufgeräumter" anzutragen
-  während des Eintipps eines Suchbegriffs werden nach Häufigkeit gewichtete Wortvorschläge angezeigt (wie sie den Nutzern unserer Smartphone-Apps bereits vertraut sind),
- zwei neue Wörterbücher erweitern das bisherige Angebot:  Deutsch-Portugiesisch  und  Deutsch-Polnisch 
- bei den Suchergebnissen werden Begriffe, die allgemein als Grund- und Aufbauwortschatz angesehen werden, am Anfang der Ergebnisliste angezeigt,
- für flektierte Formen von Wörtern wird bei den Ergebnissen nicht wie bislang nur auf deren mögliche Grundformen verwiesen, es werden zusätzlich Treffer für Grundformen mit ausgegeben,
-  Audio und  Flexionstabellen können direkt per Klick abgerufen werden, nicht mehr nur über die  Zusatzinformationen; letztere werden nun für jeden Treffer insgesamt (linke und rechte) Seite angezeigt, nicht mehr für jede Seite getrennt,
- die für eine Suche ausgelieferten Forumsartikel werden mit Kontext ausgeliefert, über den sich schneller erkennen lässt, ob eine Diskussion von Interesse ist.

Die Palette der weiteren Neuerungen ist groß und betrifft ebenso das Forum und den Trainer.

Viel Spaß beim Erkunden wünscht

Ihr LEO-Team

Abb. 8: <http://dict.leo.org/frde?lp=frde&search=> (Stand: 27.02.2013)

Durch die Testarbeit, die die User leisten, können eventuelle Fehler und Bugs, fehlende oder nicht funktionierende Features, Probleme usw. entdeckt und beseitigt werden. Dadurch tragen die User zur steten Weiterentwicklung und Verbesserung des Wörterbuchs bei (zum Konzept des 'Simultaneous Feedback' vgl. de Schryver/Prinsloo 2000a, 2000b). Das lexikographische Team erhält Feedback zum Produkt Wörterbuch und zu seinen Features sowohl vor der Herausgabe neuer Versionen als auch nachdem diese schon online sind.

Feedback können die User auch zum Wörterbuch selbst abgeben und eventuelle Änderungen, Ergänzungen und neue Features vorschlagen. Dies kann sowohl in der vorgesehenen Rubrik im Forum (im konkreten Fall in der Rubrik „Betrifft LEO“) geschehen als auch häufig durch Meldungen an das Team via E-Mail. Solche Meldungen helfen den Lexikographen und dem technischen Team, Bedürfnisse und Anforderungen der User zu erfahren und diesen, soweit es möglich ist, entgegen zu kommen. So war das der Fall als sich die User die Möglichkeit wünschten, IPA-Transkriptionszeichen im Forum zu verwenden – ein Bedürfnis, das sie schon sehr früh ausdrückten und das schließlich durch das technische Team implementiert wurde.

Vorschläge zur Gestaltung des Wörterbuchs betreffen die unterschiedlichsten Bereiche: Sie reichen vom Layout bis hin zu inhaltlichen Fragen, von der Darstellung und Hierarchisierung der Übersetzungsäquivalente bis zum Wunsch nach neuen Features wie die Auto vervollständigung in der Suchmaske oder die Ergänzung von Markierungsangaben, vom Geschäftsmo-

dell bis hin zur konzeptionellen Ausrichtung des Wörterbuchs selbst – als präskriptives, enzyklopädisches, multilinguales usw. Wörterbuch. All diese Vorschläge stellen wichtige Anregungen zur Weiterentwicklung des Wörterbuchs selbst sowohl auf inhaltlicher als auch auf technischer Seite dar, die vom lexikographischen und vom technischen Team ernst zu nehmen sind. Sie werden jedoch nur dann umgesetzt, wenn sie mit dem Projekt kompatibel sind, und dies wird grundsätzlich von der Wörterbuchredaktion und von den Informatikern, die am Wörterbuch arbeiten, bestimmt.

## 5. Vorteile und Probleme des halbkollaborativen Ansatzes

Der halbkollaborative Ansatz, den ich im vorliegenden Beitrag durch theoretische Überlegungen und anhand praktischer Beispiele zu illustrieren versucht habe, kann in der lexikographischen Praxis, aber auch in der metalexikographischen Theorie weitreichende Konsequenzen haben.<sup>25</sup> Dieser Ansatz ist mit vielerlei Vorteilen verbunden, die im Folgenden zusammengefasst werden sollen:

- Der User als Sprecher rückt ins Zentrum der Wörterbucharbeit: Seine Sprache und sein metasprachliches Wissen bilden einen wichtigen Teil der Arbeit. Dies ermöglicht die Überwindung der Unidirektionalität mancher lexikographischen Praxis zugunsten eines zirkulären Modells, das in Melchior/D'Agostini (2011) wie in Abbildung 9 gezeigt schematisiert wurde.
- Der Sprecher ist damit nicht mehr nur Rezipient eines fertigen, abgeschlossenen und „fremdbestimmten“ Produktes, sondern aktiv (und zumindest partiell richtungsbestimmend) an seiner Gestaltung beteiligt, ohne dass dadurch die wissenschaftliche Qualität des Projektes gefährdet würde, da die endgültige Entscheidungshoheit weiterhin beim ausgebildeten Lexikographen liegt;
- die Authentizität der beschriebenen Sprache, die nicht kanonisierten Modellen entstammt und welche jenseits von präskriptiven Tendenzen aufgefasst wird, stellt einen weiteren Vorteil dar;
- die ständige Verfügbarkeit unterschiedlicher Quellen ist für die lexikographische Praxis ein klarer Vorteil: Kontinuierlich wird den Wörterbuchmachern immer neues sprachliches und metasprachliches Material zur Verfügung gestellt;
- für die Wörterbuchmacher ergibt sich die Möglichkeit, die Bedürfnisse der User direkt zu erfahren und das Produkt Wörterbuch an diese – soweit wie möglich – anzupassen.<sup>26</sup> Auch hier zeigt sich die Zentralität des Users als Sprecher.

<sup>25</sup> Ohne an dieser Stelle der Frage nachzugehen, ob eine lexikographische Theorie überhaupt möglich bzw. nötig ist (zur Diskussion diesbezüglich vgl. u.a. Rundell 2012), wird hier davon ausgegangen, dass (auch linguistisch) fundierte metalexikographische Reflexion die lexikographische Praxis leiten soll.

<sup>26</sup> Es handelt sich dabei um Bedürfnisse und Wünsche für Features und Angebote, die das Projekt allgemein oder in seinen Einzelteilen betreffen, jedoch (noch) nicht um Tools, die eine Personalisierung und Individualisierung des Angebots bedeuten (wie z.B. bei der Menge an Informationen, die angezeigt werden/werden sollen), wie z.B. von Bergenholz (2011, 2013), Bothma (2011), Spohr (2011), Tarp (2011, 2012) illustriert und von Rundell (2012, S. 79) kurz diskutiert. Ziel ist, dass die User sich an der Gestaltung des Angebots mitbeteiligen, indem sie ihre teils äußerst unterschiedlichen Wünsche und Bedürfnisse einfließen lassen, damit das (polyfunktionale) Wörterbuch diesen auch gerecht werden kann.

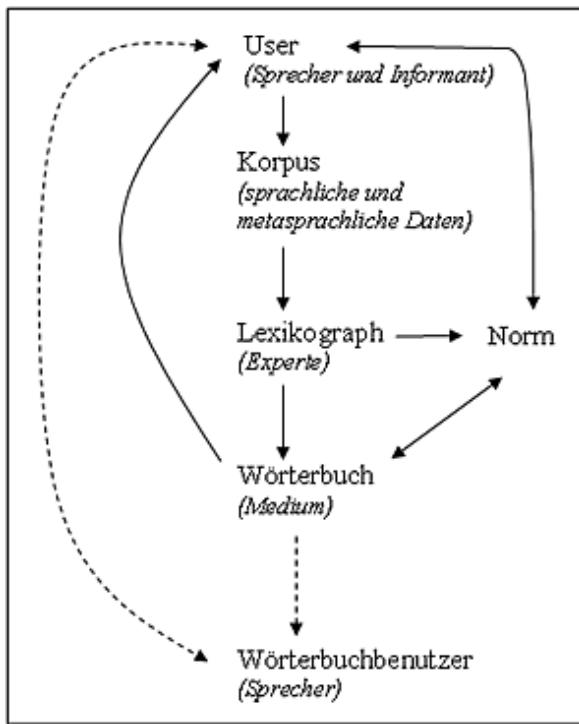


Abb. 9: Zirkuläres Modell (Melchior/D'Agostini 2011, S. 269)

Mit dem halbkollaborativen Ansatz sind allerdings auch Probleme verbunden, darunter vor allem folgende:

- die (zumindest partielle) Asystematizität der möglichen Beobachtungen: Da sich nicht alle Diskussionen in den Foren gleich entwickeln, nicht alle Themen mit der gleichen Ausführlichkeit und dem gleichen Tiefgang behandelt werden, besteht das Risiko, dass zu manchen Aspekten nur unsystematische, ja sogar rein punktuelle Beobachtungen gemacht werden können. Das Risiko ist noch größer, wenn Themen, die für den Lexikographen von Interesse wären, bei der *community* auf wenig Interesse stoßen;
- aus konzeptioneller Sicht besteht darüber hinaus das Risiko, dass jeder User das Wörterbuch nach den eigenen Bedürfnissen „(um)formen“ will, während das Wörterbuch letzten Endes doch das kohärente Produkt der Redaktion und ihrer Entscheidungen sein sollte. Damit verbunden ist das Risiko, dass die User enttäuscht bis verärgert sind, wenn ihre Vorschläge nicht umgesetzt werden. Daher ist es notwendig, dass die Redaktion ihre Entscheidungen, eventuelle Vorschläge abzulehnen – weil sie technisch nicht machbar, lexikographisch nicht sinnvoll, konzeptionell nicht vereinbar mit dem Projekt usw., den Usern begründet mitteilt;
- auf praktischer Ebene verlangt der halbkollaborative Ansatz von der Wörterbuchredaktion sowie vom technischen Team eine arbeitsintensive und zeitaufwändige Betreuung des Forums. Dieses wird zwar nicht moderiert, eine zeitnahe Rückmeldung der Lexikographen bzw. der Techniker auf die Fragen, Wünsche und Anregungen der User sowie eine rasche Überprüfung und ggf. Integration ihrer Vorschläge sind aber notwendig, um zu vermeiden, dass die User das Interesse an der aktiven Mitgestaltung verlieren bzw. dass die von ihnen gebrachten Vorschläge (z.B. für neue Einträge) (zu) lange Zeit nur im Forum zu finden sind, ohne dass sie in das eigentliche Wörterbuch integriert werden;<sup>27</sup>

<sup>27</sup> Der Zeit- und Arbeitsaufwand, den eine konsequente Anwendung des halbkollaborativen Ansatz nötig macht und der nicht immer vom lexikographischen bzw. technischen Team leistbar ist, ist der Grund, warum dieser auch in den konkreten Teilprojekten der vorgestellten Wörterbücher LEO auf unterschiedliche Weise und in unterschiedlichem Maß verfolgt wird.

- zeitaufwändig und arbeitsintensiv ist die Betreuung des Forums auch in der Hinsicht, dass die den Usern zur Verfügung gestellten Plattformen durch die Verbreitung illegaler bzw. beleidigender Inhalte oder durch Verletzung von Urheberrechten usw. missbraucht werden können. Jegliche strafrechtlich relevante, beleidigende oder in ähnlicher Weise unangemessene Inhalte müssen rechtzeitig entfernt und ggf. an die dafür zuständige Behörde gemeldet werden, da die Betreiber der Seite für deren Inhalte bzw. Qualität (auch rechtlich) verantwortlich sind. Dies bedeutet für das lexikographische bzw. technische Team eine weitere, vielfach schwierige Aufgabe, die neben den anderen zu erledigen ist.

Solche Probleme sind sicherlich ernst zu nehmen, stellen aber den Nutzen des halbkollaborativen Ansatzes nicht in Frage, da dieser auch traditionellere Arbeitsweisen in der Wörterbucherstellungspraxis nicht ausschließt, sondern bewusst und konsequent anwendet. Der halbkollaborative Ansatz bietet aber eine sinnvolle und gewinnbringende Ergänzung zu diesen, indem er die Möglichkeiten des Webs 2.0 und die neuen Wege der Wissenskonstituierung auszuschöpfen versucht. Die Vorteile, die die Zentralität des Nutzers als Sprecher mit seiner Sprache sowie der fruchtbare Austausch zwischen „Experten“ und „Laien“<sup>28</sup> mit sich bringen, sind ein großer Gewinn für die Lexikographie.

## 6. Literatur

- Abel, Andrea (2006): Elektronische Wörterbücher: Neue Wege und Tendenzen. In: San Vicente, Félix (Hg.): *Lessicografia bilingue e traduzione. Metodi, strumenti, approcci attuali*. Monza, S. 35-55.
- Bergenholtz, Henning (2011): Access to and presentation of needs-adapted data in monofunctional Internet dictionaries. In: Fuertes-Olivera, Pedro A./Bergenholtz, Henning (Hg.): *e-Lexicography. The Internet, digital initiatives and lexicography*. London/New York, S. 30-53.
- Bergenholtz, Henning (2013): Wortbildungsangaben als Hilfe für den Zugriff auf andere Datentypen und als Hilfe bei kommunikativen und kognitiven Informationsbedürfnissen. In: Klosa, Annette (Hg.): *Wortbildung im elektronischen Wörterbuch*. (= Studien zur Deutschen Sprache 63). Tübingen, S. 133-155.
- Bergenholtz, Henning/Mugdan, Joachim (1990): Formen und Probleme der Datenerhebung II: Gegenwartsbezogene synchronische Wörterbücher. In: Hausmann, Franz Josef/Reichmann, Oskar/Wiegand, Herbert Ernst/Zgusta, Ladislav (Hg.): *Wörterbücher. Ein internationales Handbuch zur Lexikographie*. Zweiter Teilband. Berlin/New York, S. 1611-1625.
- Bothma, Theo J.D. (2011): Filtering and adapting data and information in an online environment in response to user needs. In: Fuertes-Olivera, Pedro A./Bergenholtz, Henning (Hg.): *e-Lexicography. The Internet, digital initiatives and lexicography*. London/New York, S. 71-102.
- Carr, Michael (1997): Internet dictionaries and lexicography. In: *International Journal of Lexicography* 10, 3, S. 209-230.
- de Schryver, Gilles-Maurice (2003): Lexicographer's dreams in the electronic-dictionary age. In: *International Journal of Lexicography* 16, 2, S. 143-199.
- de Schryver, Gilles-Maurice/Prinsloo, Daan J. (2000a): The concept of „Simultaneous Feedback“: Towards a new methodology for compiling dictionaries. In: *Lexikos. Official Journal of the Africa Association for Lexicography* 10, S. 1-31.
- de Schryver, Gilles-Maurice/Prinsloo, Daan J. (2000b): Dictionary-making process with ‘Simultaneous Feedback’ from the target users to the compilers. In: Heid, Ulrich/Evert, Stefan/Lehmann, Egbert/Rohrer, Christian (Hg.): *Proceedings of the Ninth EURALEX International Congress*. Bd. 1. Stuttgart, S. 197-209.

<sup>28</sup> Auch die Gegenüberstellung zwischen „Experten“ und „Laien“ darf nicht als dichotomisch angesehen werden, sondern als ein Kontinuum zwischen den zwei Polen. Im konkreten Fall der User der LEO-Wörterbücher kann man vielfach beobachten, dass sie relativ nahe dem Expertenpol anzusiedeln sind (zumindest als *Linguisten*, wenn vielleicht nicht als *Lexikographen*), da sie eine hohe sprachliche aber auch metasprachliche Kompetenz aufweisen.

- Engelberg, Stefan/Lemnitzer Lothar (2009): Lexikographie und Wörterbuchbenutzung. 4., überarb. u. erw. Aufl. Tübingen.
- Fuertes-Olivera, Pedro A. (2009): The Function Theory of Lexicography and Electronic Dictionaries: Wiktionary as a prototype of collective free multiple-language Internet dictionary. In: Bergenholz, Henning/Nielsen, Sandro/Tarp, Sven (Hg.): Lexicography at a crossroads. Dictionaries and encyclopedias today, lexicographical tool tomorrow. Bern u.a., S. 99-134.
- Fuertes-Olivera, Pedro A. (2012): Lexicography and the Internet as a (re-)source. In: Lexicographica 28, S. 49-70.
- Gouws, Rufus H. (2011): Learning, unlearning and innovation in the planning of electronic dictionaries. In: Fuertes-Olivera, Pedro A./Bergenholz, Henning (Hg.): e-Lexicography. The Internet, digital initiatives and lexicography. London/New York, S. 17-29.
- Haß, Ulrike/Schmitz, Ulrich (2010): Lexikographie im Internet 2010 – Einleitung. In: Lexicographica 26, S. 1-17.
- Hausmann, Franz Josef/Werner, Reinhold Otto (1991): Spezifische Bauteile und Strukturen zweisprachiger Wörterbücher: eine Übersicht. In: Hausmann, Franz Josef/Reichmann, Oskar/Wiegand, Herbert Ernst/Zgusta, Ladislav (Hg.): Wörterbücher. Ein internationales Handbuch zur Lexikographie. Dritter Teilband. Berlin/New York, S. 2729-2769.
- Heid, Ulrich (2012): Introductory note. In: Lexicographica 28, S. 1-2.
- Heid, Ulrich/Prinsloo, Daan J./Bothma, Theo J.D. (2012): Dictionary and corpus data in a common portal: state of the art and requirements for the future. In: Lexicographica 28, S. 269-291.
- Klosa, Annette (2013): The lexicographical process (with special focus on online dictionaries). In: Gouws, Rufus H./Heid, Ulrich/Schweickard, Wolfgang/Wiegand, Herbert Ernst (Hg.): Dictionaries. An international encyclopedia of lexicography. Supplementary volume: Recent developments with focus on electronic and computational lexicography. Berlin/Boston, S. 517-524.
- Lemberg, Ingrid (2001): Aspekte der Online-Lexikographie für wissenschaftliche Wörterbücher. In: Lemberg, Ingrid/Schröder, Bernhard/Storrer, Angelika (Hg.): Chancen und Perspektiven computergestützter Lexikographie. (= Lexicographica Series Maior 107). Tübingen, S. 71-91.
- Marcoccia, Michel (2011): Sur l'internet, personne ne sait que tu es un linguiste: problèmes méthodologiques et éthiques de l'analyse des discours médiatisés par ordinateur. In: Calabrese, Laura (Hg.): L'internet, corpus sauvage. Nouvelles ressources, nouveaux problèmes? (= Le discours et la langue. Revue de linguistique française et d'analyse du discours 2.1 (2010 [2011])). Bruxelles, S. 55-72.
- Markham, Annette/Buchanan, Elizabeth (2012): Ethical decision-making and Internet research. Recommendations from the AoIR Ethics Working Committee (Version 2.0). Internet: <http://aoir.org/reports/ethics2.pdf> (Stand: 08.04.2013).
- Martin, Robert (2012): Présentation. Internet: <http://www.atilf.fr/dmf/PresentationDMF2012.pdf> (Stand: 02.01.2013).
- Melchior, Luca (2009): „Frocio“, „checca“, „morosa“ e... un problema lessicografico. In: Italienisch 62, S. 67-88.
- Melchior, Luca (2012): Halbkollaborativität und Online-Lexikographie. Ansätze und Überlegungen zu Wörterbuchredaktion und Wörterbuchforschung am Beispiel LEO Deutsch-Italienisch. In: Lexicographica 28, S. 337-372.
- Melchior, Luca/D'Agostini, Fabio (2011): Halbkollaborative Wörterbücher – Zur Darstellung von Perzeptionsdaten in bilingualen Wörterbüchern. In: Kittler, Judith/Nickenig, Annika/Siebenborn, Eva/del Valle, Victoria (Hg.): Repräsentationsformen von Wissen. Beiträge zum XXVI. Forum Junge Romanistik in Bochum (26.-29. Mai 2010). München, S. 249-273.
- Müller-Spitzer, Carolin (2003): Ordnende Betrachtungen zu elektronischen Wörterbüchern und lexikographischen Prozessen. In: Lexicographica 19, S. 140-168.
- Nesi, Hilary (2009): Dictionaries in electronic form. In: Cowie, Anthony Paul (Hg.): The Oxford history of English lexicography. Bd. 2: Specialized dictionaries. Oxford, S. 458-478.
- Rakotonolaina, Florimond (2011): Analyse du discours, corpus internet et recherche qualitative: objets d'étude et objets de recherche. In: Calabrese, Laura (Hg.): L'internet, corpus sauvage. Nouvelles ressources, nouveaux problèmes? (= Le discours et la langue. Revue de linguistique française et d'analyse du discours 2.1 (2010 [2011])). Bruxelles, S. 33-53.
- Rundell, Michael (2012): ‘It works in practice but will it work in theory?’ The uneasy relationship between lexicography and matters theoretical. In: Vatvedt Fjeld, Ruth/Torjusen, Julie Matilde (Hg.): Proceedings of the 15th EURALEX International Congress (EURALEX 2012). Oslo, S. 47-92. Online unter: [http://www.euralex.org/elx\\_proceedings/Euralex2012/pp47-92%20Rundell.pdf](http://www.euralex.org/elx_proceedings/Euralex2012/pp47-92%20Rundell.pdf) (Stand: 08.04.2013).
- Schröder, Martin (1997): Brauchen wir ein neues Wörterbuchkartell? Zu den Perspektiven einer computerunterstützten Dialektlexikographie und eines Projektes „Deutsches Dialektwörterbuch“. In: Zeitschrift für Dialektologie und Linguistik 64, 1, S. 57-65.

- Spoehr, Dennis (2011): A multi-layer architecture for ‘Pluri-Monofunctional’ dictionaries. In: Fuertes-Olivera, Pedro A./Bergenholtz, Henning (Hg.): e-Lexicography. The Internet, digital initiatives and lexicography. London/New York, S. 102-120.
- Storrer, Angelika/Freese, Katrin (1996): Wörterbücher im Internet. In: Deutsche Sprache 24, S. 97-153.
- Tarp, Sven (2011): Lexicographical and other e-tools for consultation purposes: Towards the individualization of needs satisfaction. In: Fuertes-Olivera, Pedro A./Bergenholtz, Henning (Hg.): e-Lexicography. The Internet, digital initiatives and lexicography. London/New York, S. 54-70.
- Tarp, Sven (2012): Online dictionaries: today and tomorrow. In: Lexicographica 28, S. 253-267.
- Wiegand, Herbert Ernst (1998): Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband. Berlin/New York.
- Zgusta, Ladislav (1971): Manual of lexicography. (= Ianua linguarum. Series maior 39). Prag.

## ***Duden online* und seine Nutzer**

*Karin Rautmann, Bibliographisches Institut GmbH, Mannheim*

### **1. Zur Genese von *Duden online***

Im Mai 2011 ging unter [www.duden.de](http://www.duden.de) mit *Duden online* ein neues Internetwörterbuch an den Start. Es handelt sich um ein umfassendes Wörterbuch der deutschen Gegenwartssprache, das alle wesentlichen Aspekte der Sprache beschreibt: Rechtschreibung, Grammatik, Bedeutung, Aussprache, Etymologie, Synonyme, Beispiele, Redewendungen und typische Verbindungen in Form von Wortwolken.

*Duden online* stellt bis heute das Herzstück der ebenfalls im Mai 2011 relaunchten Verlagswebsite [www.duden.de](http://www.duden.de) dar: Rund 80 % aller Besuche der Website gelten dem Wörterbuch. Es wird aktuell von mehr als 5 Millionen Besuchern<sup>1</sup> jeden Monat genutzt (Stand: Januar 2013).

Das Wörterbuch basiert auf Daten des Duden-Wissensnetzes, der zentralen Quelle der von Duden vorgehaltenen Sprachdaten (vgl. dazu ausführlicher Alexa 2011 und Konopka/Wind/Alexa in Vorbereitung). Da bei der Erstellung von *Duden online* auf bestehende Wörterbuchsubstanzen zurückgegriffen werden konnte, war es möglich, das Projekt nach Freigabe durch die Verlagsleitung des Bibliographischen Instituts relativ schnell umzusetzen. Die Realisierung erfolgte durch ein fachübergreifendes Team: Lexikographen, Computerlinguisten und Onlinespezialisten haben aufs Engste zusammengearbeitet.

Primäre Aufgabe der Lexikographen war es, das Wörterbuchkonzept zu entwickeln und die für die Umsetzung des Konzepts von *Duden online* relevanten Daten im Wissensnetz zu definieren. Der wesentliche Beitrag der Computerlinguisten bestand darin, die Daten im Wissensnetz möglichst automatisch auszuzeichnen, sie anzureichern und den Export der Duden-online-Daten aus dem Wissensnetz einschließlich der Anbindung von Bild- und Audiodaten zu programmieren. Die Onlinespezialisten verantworteten den Import der Daten in ein Content-Management-System (ursprünglich *Liferay*, seit Juli 2012 *Drupal*), darüber hinaus das Webdesign gemäß den Markenrichtlinien, die Usability sowie die Auffindbarkeit der Inhalte für Suchmaschinen. 80 % der Nutzer von *Duden online* kommen zurzeit über die Suchmaschine *Google* auf [www.duden.de](http://www.duden.de). Deshalb ist die Suchmaschinenoptimierung (SEO, aus engl. search engine optimization) in Bezug auf *Google* besonders wichtig für den Erfolg des Wörterbuchs.

### **2. Marktforschung im Dudenverlag (2009)**

In regelmäßigen Abständen, insbesondere bei der Konzeption neuer Werke oder bei wichtigen Neuauflagen, werden im Dudenverlag in größerem Umfang Nutzermeinungen eingeholt. Das geschieht in der Regel dadurch, dass entweder externe Firmen mit einer Marktforschung beauftragt werden oder der Verlag selbst „mit Bordmitteln“ Umfragen durchführt. Gern genutzt wird bei letzterem Vorgehen die Möglichkeit, mittels eines Umfragetools eine Onlineumfrage auf [www.duden.de](http://www.duden.de) zu verankern und die Ergebnisse selbst auszuwerten. Das spart nicht nur

<sup>1</sup> Wegen der besseren Lesbarkeit wird in dieser Publikation das generische Maskulinum verwendet. Damit sind ausdrücklich Frauen und Männer gemeint.

Ressourcen, es liegen auch sehr schnell Ergebnisse vor. Bei der Konzeption und Weiterentwicklung von *Duden online* sind beide Vorgehensweisen zum Einsatz gekommen.

In der Phase der Konzeptentwicklung von *Duden online*, inhaltlich aber nicht nur zum Thema *Duden online*, hat der Verlag eine groß angelegte Marktforschung bei einer externen Firma in Auftrag gegeben (Ende 2009/Anfang 2010). Gegenstand der Forschung waren u.a. die Einstellung der Befragten zum Thema Rechtschreibung sowie zur Marke Duden, die Nutzung von Rechtschreibhilfen (Wörterbücher, Korrektursoftware) und die Anforderungen an ein Onlineangebot unter der Marke Duden (Bezahlbereitschaft, Konzept). Es handelte sich um zwei repräsentative Studien, eine qualitative und eine quantitative.

Im Ergebnis haben die Studien in Bezug auf die Nachschlagefunktionalität gezeigt, dass es zum Zeitpunkt der Erhebung, also Ende 2009, im Internet nur eine sehr geringe Zahlungsbereitschaft für Wörterbuchinhalte gab, dass hingegen Werbung als Finanzierungsmodell akzeptiert war, solange sie dezent ist und zum Angebot passt.<sup>2</sup> Auf inhaltlicher Ebene haben die Ergebnisse der Studien gezeigt, dass aus Nutzersicht neben Rechtschreibung, Grammatik und Bedeutung vor allem Synonyme wichtige Bestandteile des Angebotsspektrums sind. Die Frage, ob die Nutzer – unabhängig von der Marke Duden – auch Interesse an einer aktiven Beteiligung an einem sprach- bzw. rechtschreibbezogenen Internetangebot hätten, wurde zwar als interessant bewertet, aber mehrheitlich mangels individueller Relevanz negativ beantwortet.

Die Ergebnisse der Studien hatten entscheidende Folgen für die bis dahin vorliegende Konzeption von *Duden online*, und zwar sowohl in Bezug auf das Geschäftsmodell von [www.duden.de](http://www.duden.de) als auch in Bezug auf die Inhalte. Die Verlagsleitung traf die Entscheidung, die Nachschlagemöglichkeit auf [www.duden.de](http://www.duden.de) den Nutzern künftig kostenfrei anzubieten. Im Gegenzug räumte man der Werbung für Dudenprodukte größeren Raum ein. In Bezug auf die Inhalte wurde der ursprünglich veranschlagte Stichwortbestand etwas reduziert und bestimmte Informationstypen, z.B. Zitate, ganz aus dem Angebot genommen. Der Aufbau eines Wörterbuchartikels wurde noch einmal dahin gehend überarbeitet, dass z.B. die von den Nutzern relativ hoch geschätzten Synonyme näher an den Anfang des Artikels gerückt sind. Die Informationen, die für weniger Besucher interessant waren, z.B. Herkunft oder Aussprache, rückten weiter ans Artikelende.

### 3. Onlineumfrage auf [www.duden.de](http://www.duden.de) (2011)

Im November 2011, also rund ein halbes Jahr nach dem Start von *Duden online*, wollte das Duden-online-Team erstmals mehr über die Nutzer und die Nutzung von *Duden online* in Erfahrung bringen und die Planung für die Weiterentwicklung von *Duden online* absichern. Zu diesem Zweck wurde eine insgesamt 13 Fragen umfassende Onlineumfrage konzipiert und auf den Wörterbuchseiten von [www.duden.de](http://www.duden.de) eingebaut. Sie lief insgesamt knapp vier Wochen (28.11.2011 bis 19.12.2011). Für die Teilnahme an der Umfrage wurde ein Verlagsprodukt als Preis ausgelobt. Es haben sich insgesamt 626 Personen an der Umfrage beteiligt. Auch wenn die Ergebnisse nicht repräsentativ sind, waren sie für das Duden-online-Team doch sehr aufschlussreich und haben zu Konsequenzen in Bezug auf die weitere Planung von *Duden online* geführt. Im Folgenden werden die wichtigsten Ergebnisse der Onlineumfrage vorgestellt und kommentiert.

<sup>2</sup> Auf [www.duden.de](http://www.duden.de) wurden Wörterbuchinhalte bis zu diesem Zeitpunkt nur zum Teil kostenfrei (die ersten 150 Zeichen eines Artikels) angeboten. Vollständige Artikel wurden nur nach Abschluss eines Abonnements angezeigt.

Bezüglich der Altersstruktur der Umfrageteilnehmer ist festzustellen, der mehr als die Hälfte (54,5 %) unter 30 Jahre alt war, ein weiteres knappes Viertel (23,0 %) unter 45. Das restliche Viertel teilen sich die über 45-Jährigen (21,9 %) und die Unter-18-Jährigen (18,3 %).

In Bezug auf die Nutzungssituation stellte sich heraus, dass *Duden online* von knapp einem Viertel (24,4 %) der Befragten eher beruflich genutzt wurde und von 16,9 % für die Schule. 40,4 % der Befragten nutzten das Onlinewörterbuch beruflich und privat gleichermaßen. Nur 17,9 % gaben an, *Duden online* eher privat zu nutzen. Der sich hier abzeichnende Schwerpunkt im Bereich der beruflichen Nutzung wird auch von der Verteilung des Traffics auf [www.duden.de](http://www.duden.de) über den Tag und über die Woche bestätigt: An Wochenenden sowie unter der Woche abends und nachts sinken die Besuchszahlen auf der Website in der Regel drastisch.

Um die Zweckmäßigkeit des Angebots zu prüfen, wurde den Nutzern die Frage „Konnten Sie bei Ihrem letzten Besuch von *Duden online* Ihre Frage schnell und sicher klären?“ gestellt. Diese beantworteten 90,6 %, wie die folgende Grafik zeigt, mit Ja, 9,3 % mit Nein, 0,2 % machten keine Angaben:

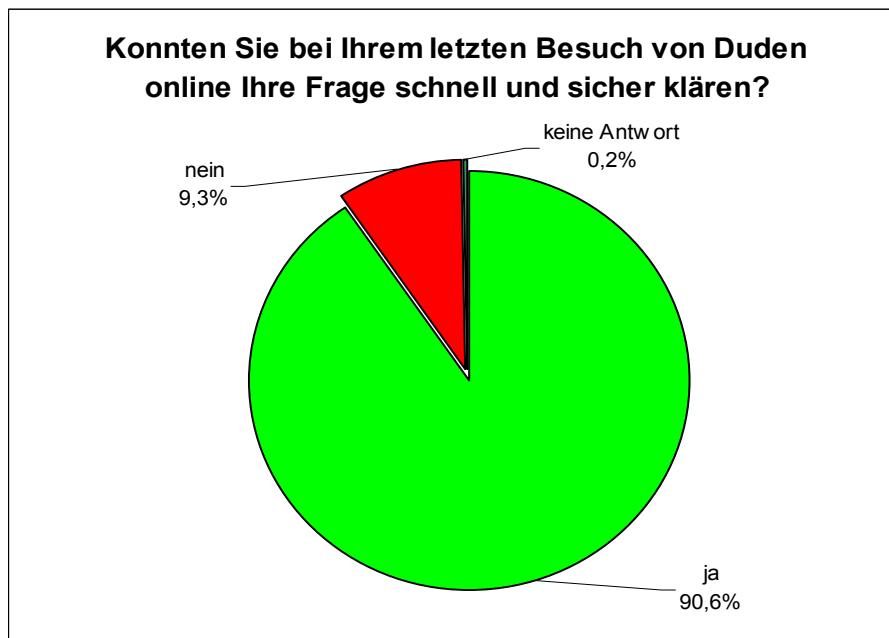
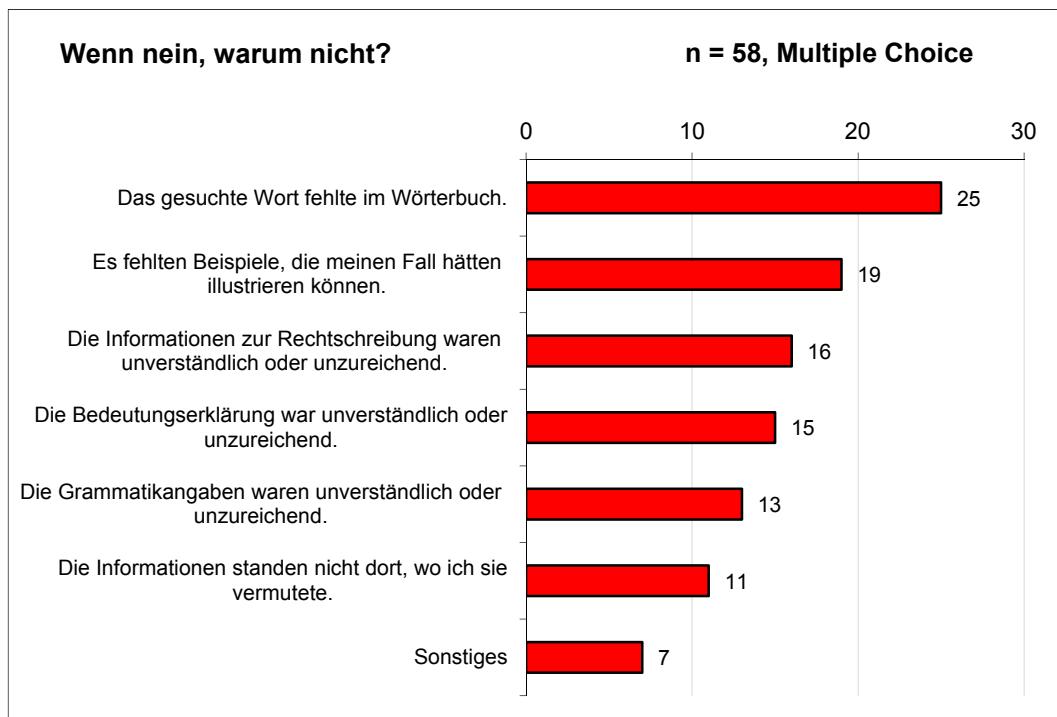


Abb. 1: Zweckmäßigkeit von *Duden online*, Teil 1

Dass das Urteil so positiv ausgefallen ist, hat das Duden-online-Team überrascht. Wenig überraschend waren dagegen die Gründe, warum ein Besuch von *Duden online* nicht zum Erfolg geführt hatte. Am häufigsten wurde die Lemmalücke genannt (25 % der insgesamt 58 Antworten), gefolgt von dem Mangel an passenden Beispielen (19 %) und dem Mangel bzw. der Unverständlichkeit der Rechtschreibinformation (16 %):

Abb. 2: Zweckmäßigkeit von *Duden online*, Teil 2

Auch wenn nicht jedes vom Nutzer gesuchte, aber nicht gefundene Wort faktisch im Wörterbuch fehlt – vielleicht hat er es auch in einer falschen Schreibung eingegeben und weder die Fehlertoleranz in der Suche noch die Meinten-Sie-Funktionalität führen ihn zur richtigen Schreibung –, ist als Konsequenz aus diesem Befund unter dem Aspekt der Weiterentwicklung von *Duden online* sicherlich die Erweiterung des Stichwortbestands vorrangig, aber auch die vermehrte Aufnahme von Beispielen.

Mit dem Ziel, mehr über die inhaltlichen Interessen der Besucher zu erfahren, wurden den Befragten bestimmte Angabetypen zur Bewertung vorgelegt (Frage: „Wie wichtig sind Ihnen folgende Informationen in *Duden online*?“). Nach welchen Angaben gefragt wurde und wie die Besucher die einzelnen Angabetypen bewertet haben, zeigt die folgende Grafik:

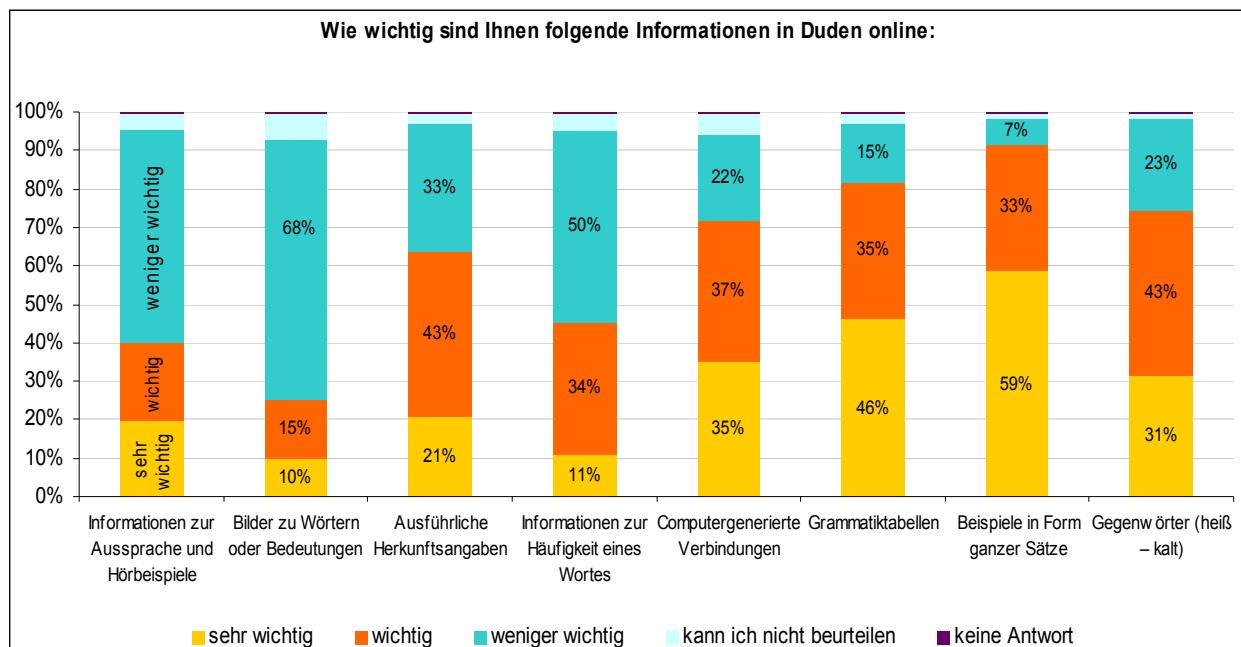


Abb. 3: Inhaltliche Interessen der Duden-online-Nutzer

Auf alle Ergebnisse detailliert einzugehen, würde den Rahmen dieser Publikation sprengen. Interessant sind vor allem die beiden Extrempositionen, und zwar zum einen die Einschätzung der Wichtigkeit von „Bildern zu Wörtern oder Bedeutungen“ (2. Säule), zum anderen die Wichtigkeit von „Beispielen in Form ganzer Sätze“ (vorletzte Säule). Eher unwichtig war für die Nutzer das Bildmaterial in *Duden online*: Lediglich 10 % der Nutzer fanden sie „sehr wichtig“, 68 % „weniger wichtig“. Die Annahme des Duden-online-Teams, dass die Onlinewörterbuch-Nutzer auch Bildmedien erwarten, hat sich damit als falsch erwiesen.<sup>3</sup> Als Konsequenz daraus wurde bei der Planung zur Weiterentwicklung von *Duden online* auf die Aufnahme neuer Bilder vorerst verzichtet.

Ganz im Gegensatz zu den Bildern wurden „Beispiele in Form ganzer Sätze“ von vielen Duden-online-Nutzern als hilfreich empfunden: 59 % fanden sie „sehr wichtig“, nur 7 % halten sie für „eher unwichtig“. Da es bereits verschiedene Onlinewörterbücher gibt, die vollständige Beispielsätze anbieten (z.B. *Wiktionary*) oder Korpora einbeziehen (z.B. *DWDS*), und natürlich, da sich jeder beispielsweise über *Google* selbst ganz einfach Satzkontexte verschaffen kann, hat sich hier offensichtlich schon eine Gewohnheit herausgebildet, die die Nutzer bedient sehen wollen. Aus Dudensicht würde sich daher die Einbindung des Dudenkorpus anbieten.

<sup>3</sup> Dieses Ergebnis entspricht im Übrigen den Ergebnissen von Müller-Spitzer/Koplenig/Töpel (2011), die u.a. feststellten, dass gerade das Besondere von Onlinewörterbüchern, nämlich (u.a.) die Multimedialität, von den Nutzern von Onlinewörterbüchern als relativ unwichtig eingestuft wird.

In Zeiten des Web 2.0 und des damit verbundenen Konzepts des „user generated contents“ (UGC) wollte das Duden-online-Team 2011 prüfen, wie groß das Interesse der Duden-online-Nutzer ist, an *Duden online* mitzuwirken. Die Marktforschung von 2009 hatte ja gezeigt, dass sich nur die wenigsten Personen vorstellen konnten, an der Gestaltung von Sprach- bzw. Wörterbuchinhalten im Internet aktiv mitzuwirken (vgl. Kapitel 2). Konkret wurde im Rahmen der Onlineumfrage auf [www.duden.de](http://www.duden.de) 2011 danach gefragt, ob die Nutzer sich wünschen, Feedback zu den einzelnen Wörterbuchartikeln zu geben und am Wörterbuch selbst mitszuschreiben. Die Ergebnisse von 2011, die die folgenden Grafiken zeigen, haben die Ergebnisse der Marktforschung von 2009 alles andere als bestätigt.

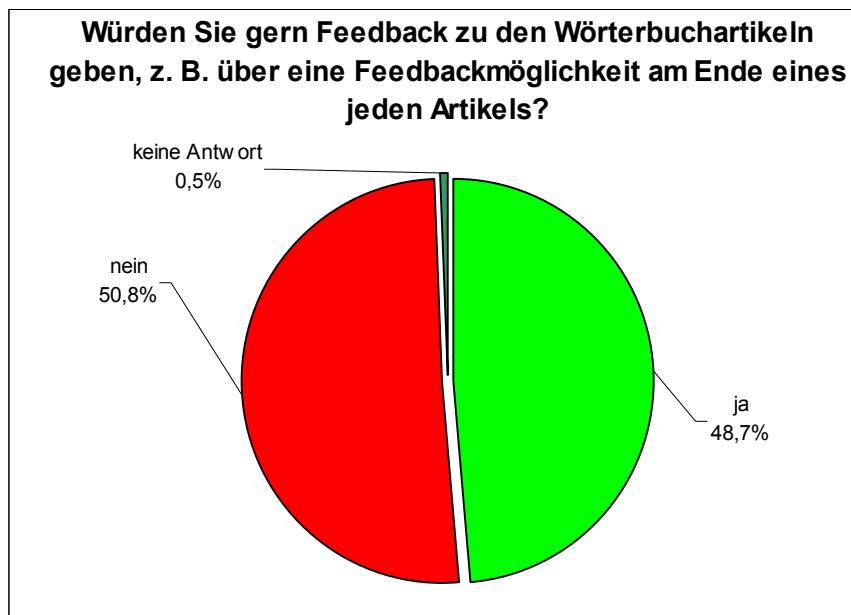


Abb. 4: Einschätzung einer Feedbackmöglichkeit am Ende von Duden-online-Artikeln

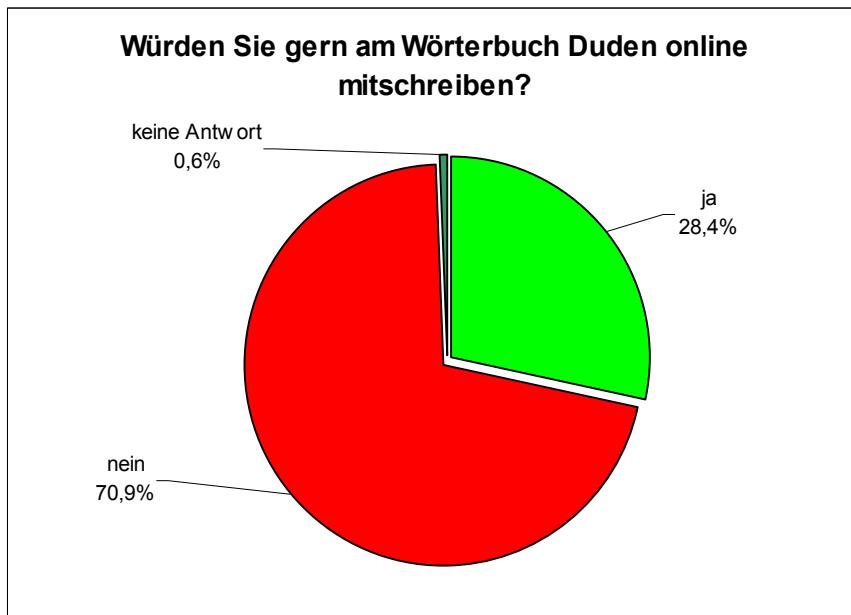


Abb. 5: Einschätzung des Wunsches nach redaktioneller Mitgestaltung von *Duden online*

Knapp die Hälfte aller Nutzer würde gern Feedback zu den Wörterbuchartikeln geben und knapp ein Drittel aller Nutzer würde gern selbst an *Duden online* mitschreiben. Nun sind die Ergebnisse der Onlineumfrage einerseits nicht repräsentativ. Sicherlich haben, bedingt durch

das Onlinemedium, viel mehr internetaffine Personen teilgenommen als bei der Marktforschung aus dem Jahr 2009. Andererseits existierte *Duden online* 2009 nur auf dem Reißbrett, sodass die Fragen zum Wörterbuch 2009 nur abstrakt beantwortet werden konnten. 2011 wurde hingegen im direkten Wörterbuchkontext gefragt, sodass die Befragten eine deutlich konkretere Vorstellung davon entwickeln konnten, was gefragt wurde. Auch wenn es also problematisch ist, die beiden Ergebnisse miteinander zu vergleichen, hat das Thema UGC dennoch für das Duden-online-Team stark an Bedeutung gewonnen. Da die Marke Duden aber gleichzeitig als Instanz in Sachen deutsche Sprache gilt und mit einem hohen Qualitätsversprechen verbunden ist, müssen konzeptuelle Überlegungen in dieser Richtung sorgfältig geprüft und abgewogen werden.<sup>4</sup> In welcher Weise die Nutzer von *Duden online* bereits heute ganz direkt an der Gestaltung der Wörterbuchinhalte mitwirken, ist Gegenstand des folgenden Kapitels.

#### 4. Direktes Nutzerfeedback

Seit dem Start von *Duden online* haben Benutzer die Möglichkeit, dem Duden-online-Team über ein Webformular Vorschläge für Neuaufnahmen zukommen zu lassen. Das Webformular lässt sich über einen Button „Wortvorschlag“ am Kopf eines jeden Wörterbuchartikels öffnen. Der Nutzer wird aufgefordert, seinen Namen sowie eine gültige E-Mail-Adresse anzugeben. Über den eigentlichen Wortvorschlag hinaus kann er optional noch eine Beschreibung des vorgeschlagenen Wortes geben. Unabhängig davon nutzen die User auch andere Wege, um dem Duden-online-Team etwas mitzuteilen. Sie schreiben z.B. an den Webmaster von [www.duden.de](http://www.duden.de), den Endkundenservice des Verlages oder die technische Hotline für die Duden-Softwareprodukte, um inhaltliche Fragen zu stellen oder Hinweise zu geben.

Das Duden-online-Team erhielt in der ersten Jahreshälfte 2012 über die gerade skizzierten Wege pro Monat im Durchschnitt rund 120 Mails. Von diesen 120 Mails waren rund zwei Drittel reine Wortvorschläge, also Vorschläge für Wörter, die bis zu dem Zeitpunkt noch nicht in *Duden online* verzeichnet waren.

Hier ein Beispiel für einen Wortvorschlag:

Anrede = Frau  
Vorname = xxxxxxxx  
Nachname = xxxxxxxx  
E-Mail =xxxxxxxxxxxxxxxxxxxxxx  
Vorgeschlagenes Wort = Filesharing  
Beschreibung des Worts = englisch für Dateien teilen, Dateifreigabe oder gemeinsamer Dateizugriff

Rund ein Drittel der 120 Mails beinhaltete sog. Ergänzungsvorschläge oder Fehlerhinweise. Ergänzungsvorschläge sind Vorschläge zu bereits existierenden Artikeln. In den meisten Fällen schickten die Nutzer Vorschläge für die aus ihrer Sicht fehlenden Bedeutungen oder Synonyme. Seltener wurden grammatische Angaben oder idiomatische Wendungen vermisst, ganz selten Ausspracheangaben oder Bilder.

<sup>4</sup> Ein Wörterbuch, bei dem [jugendliche] Nutzer Wortvorschläge unterbreiten und Wörter definieren können, betreibt der Dudenverlag bereits seit 2009 zusammen mit dem Hamburger Trendbüro: *Duden – Neues Wörterbuch der Szenesprachen*, siehe <http://szenesprachenwiki.de>.

Hier ein Beispiel für einen – in diesem Fall schon sehr umfangreichen – Ergänzungsvorschlag:

Anrede = Herr

Vorname = xxxxxxxx

Nachname = xxxxxxxxx

E-Mail =xxxxxxxxxxxxxxxxxxxxxx

Vorgeschlagenes Wort = Rolle rückwärts machen

Beschreibung des Worts = Rolle rückwärts machen (Redewendung) (fig.)

Jemand macht eine Rolle rückwärts, wenn er plötzlich eine andere Meinung vertritt, von einem Vorhaben abrückt o.Ä. (wohl synonym mit einer Kehrtwendung machen)

Belege:

„... zeigt, dass er wohl selbst von seinem Konzept nicht mehr überzeugt ist. Mit der Verkündung der zusätzlichen Standorte macht der Minister jetzt eine Rolle rückwärts‘, meinen Razavi und Birk.“

(<http://www.swp.de/goeppingen/lokales/goeppingen/CDU-Politiker-Minister-macht-Rolle-rueckwaerts;art5583,1576698>)

„Regelhüter machen Rolle rückwärts

Die FOTA, die Interessenvertretung aller Formel-1-Teams, hat es geschafft. Die geplante Regeländerung des Internationalen Automobil-Verbandes, wonach der Fahrer mit den meisten Siegen Weltmeister wird, wurde gekippt...“

(<http://www.stern.de/sport/formel1/rolle-rueckwaerts-90832345t.html>)

Alle Wort- und Ergänzungsvorschläge werden vom Duden-online-Team täglich gesichtet, sorgfältig geprüft und im positiven Fall direkt im Duden-Wissensnetz (vgl. Kapitel 1) umgesetzt. Für die Aufnahme von Wörtern in *Duden online* gelten grundsätzlich die allgemeinen Richtlinien der Dudenredaktion zur Aufnahme eines Wortes in ein Wörterbuch. Das heißt, die Wörter müssen im Dudenkorpus in einer bestimmten Häufigkeit belegt sein, es darf sich also nicht um Augenblicksbildungen handeln, sie müssen textsortenübergreifend belegt sein usw. (siehe ausführlicher dazu Scholze-Stubenrecht 2002 u. Kunkel-Razum 2003). Das Argument der Platznott entfällt natürlich (s. ausführlicher dazu Konopka/Wind/Alexa in Vorbereitung). Die Kunden erhalten in der Regel binnen weniger Tage per Mail einen Hinweis darauf, ob ihr Vorschlag angenommen wurde oder nicht. Mit dem nächsten Export der Wissensnetzdaten ins Livesystem werden die Veränderungen dann auch auf der Website sichtbar.

Von den rund 80 Wortvorschlägen, die in der ersten Jahreshälfte 2012 beim Duden-online-Team eingingen, wurde mehr als jeder zweite (52 %) aufgenommen. Stichwörter wie *Kuchendiagramm*, *Burgerbude*, *Fiskalpakt*, *Kassenpraxis* oder *eifersüchteln* sind aufgrund von Nutzervorschlägen aufgenommen worden. Nicht geschafft in *Duden online* haben es Erfindungen oder Eintagsfliegen wie *Tellerrandbestauner*, *herumjürgen* (herumnögeln), *pdffieren* (in PDF umwandeln) oder *kuschelliberal*.

Mit Fehlerhinweisen verfährt das Duden-online-Team ähnlich wie mit Wort- oder Ergänzungsvorschlägen, wobei die Prüfung in der Regel schneller vorstatten geht. Oft sind es Formalia, die bemängelt werden, z.B. fehlende oder zu große Abstände, manchmal Widersprüche in der Pluralbildung zwischen *Duden online* und einem Printwerk oder auch fehlerhafte Verweise wie im folgenden Fall.

Hier ein Beispiel für einen Fehlerhinweis:

Liebes Duden-Team,

ein Hinweis zum Eintrag „Ulknudel“: Meiner Ansicht nach muss hierin auf die fünfte Bedeutung von „Nudel“ anstelle der vierten verwiesen werden.

Mit freundlichen Grüßen

XXXX

Die Korrektur solcher Fehler erfolgt ebenfalls direkt im Wissensnetz, sodass sie nicht nur *Duden online*, sondern gleichzeitig auch anderen Werken im Wissensnetz zugutekommt. Deshalb kann der Wert dieser Hinweise für die Verbesserung der Qualität der Dudensubstanzen kann gar nicht hoch genug eingeschätzt werden.

## 5. Webanalyse

Die Nutzer von *Duden online* treten nicht nur per Mail mit dem Duden-online-Team in Kontakt, sondern sie hinterlassen durch die Nutzung von *Duden online* auch ihre Spuren im Internet. Diese kann man mithilfe eines Webanalysetools, z.B. *Google Analytics*, sichtbar machen. So kann man beispielsweise sehen, wann die Besucher über den Tag, die Woche, den Monat verteilt auf [www.duden.de](http://www.duden.de) gehen, welche externen Seiten sie vorher besucht haben oder wie sie sich auf der Duden-Website selbst verhalten. Die Auswertung von Daten der Website Nutzung zielt auf ein besseres Verständnis des Besucherverhaltens ab, was wiederum Voraussetzung für eine Optimierung des Angebots ist.

Aus redaktioneller, aber auch aus Nutzersicht ist es in diesem Zusammenhang vor allem interessant zu sehen, welche Artikel die Nutzer von *Duden online* am meisten interessieren bzw. welche sie am häufigsten aufrufen. Der Aufruf eines Duden-online-Artikels erfolgt in über 80 % der Fälle nach dem Auslösen einer Suche in *Google*, 18 % kommen über die Eingabe der URL [www.duden.de](http://www.duden.de) in den Browser bzw. haben diese Seite gebookmarked.

Die zehn meistgelesenen Artikel von *Duden online* werden täglich auf der Startseite von [www.duden.de](http://www.duden.de) in der rechten Randspalte publiziert, wobei sich der Messzeitraum immer auf den Vortag (0 bis 24 Uhr) bezieht. Für Interessierte findet sich am Ende der Top Ten ein Link auf die Top 100 des Vortages. Die zehn am häufigsten aufgerufenen Artikel beispielsweise des 14. Juni 2012 sind folgende:

The screenshot shows a yellow header bar with the text 'Meistgelesene Artikel'. Below it is a list of ten most-read articles, each preceded by a small red arrow icon. The list includes: 'des Weiteren', 'selbstständig', 'sodass', 'kennenlernen', 'zum einen ... zum anderen', 'Bescheid', 'im Voraus', 'unter anderem', 'Status', and 'aufwendig'. At the bottom of the list is a link 'Liste der 100 meistgelesenen Artikel'.

Meistgelesene Artikel
› des Weiteren
› selbstständig
› sodass
› kennenlernen
› zum einen ... zum anderen
› Bescheid
› im Voraus
› unter anderem
› Status
› aufwendig
› Liste der 100 meistgelesenen Artikel

Abb. 6: Die 10 meistgelesenen Artikel auf [www.duden.de](http://www.duden.de) vom 14. Juni 2012

Auf einer Website der Marke Duden überrascht es nicht, dass es sich bei den meistgelesenen Artikeln primär um rechtschreiblich schwierige Wörter handelt. Nach ihnen wird auch oft in der Duden-Sprachberatung gefragt. Die Liste der Top Ten ist relativ stabil. „Ausreißer“ finden nur Eingang, wenn z.B. in einer großen TV-Quizsendung eine Frage nach der Herkunft oder

Bedeutung eines Wortes oder einer Redewendung gefragt wird und die Zuschauer massenweise das gesuchte Wort in das Sucheingabefeld ihres Browsers eingeben oder direkt auf die Duden-Website kommen.

An besagtem 14. Juni 2012 wurden auf [www.duden.de](http://www.duden.de) insgesamt rund 100.000 Artikel aufgerufen, wobei auf die Nummer 1 (*des Weiteren*) rund 1.200 Aufrufe entfielen und auf Nummer 10 (*aufwendig*) rund 400 Aufrufe. Ganz allgemein kann für *Duden online* konstatiert werden, dass sehr wenige Wörterbuchartikel sehr häufig aufgerufen werden und sehr viele Artikel sehr selten.

Interessant ist in diesem Zusammenhang aber nicht nur, welche Artikel die Nutzer aufrufen, sondern auch, was sie suchen und vor allem wie sie suchen, d.h., was genau sie in das Suchfeld von *Duden online* oder *Google* eingeben. Das wird Gegenstand des folgenden Kapitels sein. Wenn man weiß, was die Nutzer eingeben, kann man das Angebot optimal auf ihr Verhalten abstimmen und dadurch auf Nutzerseite größtmögliche Zufriedenheit erzielen.

## 6. Analyse der Sucheingaben

In diesem Kapitel stehen die Sucheingaben der Nutzer von [www.duden.de](http://www.duden.de) im Vordergrund. Um den Rahmen dieser Arbeit nicht zu sprengen, fokussiert die folgende Analyse die 2.500 häufigsten Sucheingaben eines ganz normalen (Arbeits-)Tages. Der Datenabzug stammt vom 12. Juni 2012. Angesichts des enormen Traffics, der täglich auf [www.duden.de](http://www.duden.de) stattfindet, ist das nur ein sehr kleiner, punktueller Ausschnitt, umfangreichere Analysen wären äußerst wünschenswert.

Alle der im Folgenden angeführten Sucheingaben wurden an dem besagten Tag mindestens 370-mal eingegeben (es handelt sich also nicht um Einzelfälle), wobei nicht unterschieden werden kann zwischen der Eingabe in das Sucheingabefeld von *Google*, der Duden-Website oder einer anderen Quelle. Alle folgenden Schreibungen sind Originalschreibungen von Nutzern, Falschschreibungen sind mit \* gekennzeichnet.

Betrachtet man die Top 2.500 der Sucheingaben unter formalen Aspekten, ist zunächst festzustellen, dass die Nutzer in mehr als 90 % der Fälle Lemmata eingegeben haben (*aufgrund* [Top 1], *explizit* [Top 8], *widerspiegeln* [Top 13]). In knapp 5 % der Fälle gaben sie statt der Grundform eine flektierte Form ein. Am häufigsten eingegebene Flexionsformen sind Infinitivkonstruktionen mit „zu“, allen voran *kennenzulernen* (Top 4), *weiterzuentwickeln* (Top 77) und *zukommen zu lassen* (Top 80). Darüber hinaus finden sich unter den Top 2.500 zahlreiche

- Plurale (z.B. *Unannehmlichkeiten* [Top 241], *hobbies\** [Top 1537], *kenntnisse\** [Top 699]),
- Partizipien (z.B. *kennengelernt* [Top 57], *bekanntgegeben* [Top 1358], *zugrundeliegende\** [Top 1029]),
- Verbformen (z.B. *heisst\** [Top 2033], *hinzukommt* [Top 2056], *vertan* [Top 2039]),
- sonstige flektierte Formen (*am besten* [Top 12], *ähnliches* [Top 1065], *weitere* [Top 1453]).

Was die Komplexität der Eingaben betrifft, so wurden bevorzugt Einwortausdrücke eingegeben (*zurzeit* [Top 2], *mithilfe* [Top 7], *dekadent* [Top 17]), es finden sich aber auch komplexere Konstruktionen, angefangen bei Mehrwortlexemen (z.B. *des weiteren\** [Top 7], *zum einen* [Top 10], *vor allem* [Top 14], *unter anderem* [Top 19], *nach wie vor* [Top 71], *im Folgenden* [Top 107]) über feste bzw. idiomatische Wendungen (z.B. *auf dem laufenden halten\** [Top 608], *alles in allem* [Top 741]), *herzlich willkommen* [Top 802], *ohne Schmarrn* [Top 1903]) bis hin zu ganzen Sätzen (z.B. *gib mir bescheid\** [Top 1010], *Guten Morgen Zusammen\** [Top 1104], *es tut mir leid* [Top 186], *ich weis\* bescheid\** [Top 422]).

Unter rechtschreiblichen Gesichtspunkten ist zunächst festzustellen, dass die Groß- und Kleinschreibung von den Nutzern weitestgehend ignoriert wird bzw. sie auch Substantive bevorzugt in Kleinschreibung eingeben (*email\** [Top 37], *im voraus\** [Top 46], *status\** [Top 49]). Davon abgesehen finden sich in der Liste alle möglichen Arten von Falschschreibungen, darunter z.B.

- falsche Zusammenschreibungen (z.B. *desweiteren\** [Top 5], *immernoch\** [Top 41], *vorallem\** [Top 72]),
- falsche Getrenntschreibungen (z.B. *zur zeit\** [Top 85], *zu zusenden\** [Top 108]),
- Fehler in der Laut-Buchstaben-Beziehung (z.B. *strasse\** [Top 200], *wiederspiegeln\** [Top 31], *email\** [Top 37] *weist\** [Top 682]),
- alte Rechtschreibung (z.B. *bißchen* [Top 301], *Tip* [Top 284], *muß* [Top 899], *daß* [Top 1306]).

Die Liste zeigt sehr deutlich, worin die rechtschreiblichen Unsicherheiten der Nutzer von *Duden online* bestehen. Sie hilft dem Duden-online-Team aber auch, festzustellen, welche Wörter bzw. Kombinationen in *Duden online* gesucht, aber nicht oder nicht leicht gefunden werden. Diese besonders häufig vorkommenden Fälle werden redaktionell einzeln geprüft und optimiert, sodass die Nutzer sie besser finden können. Optimierung heißt dabei, dass das fragliche Wort bzw. der betreffende Artikel so behandelt wird, dass er durch das Zusammenspiel zwischen der auf [www.duden.de](http://www.duden.de) standardmäßig ausgelösten Stichwortsuche und der nachgeschalteten Volltextsuche unter den drei ersten Treffern der Trefferliste landet oder über die „Meinten-Sie“-Funktionalität, die im Fall von null Treffern Alternativen aus der Substanz vorschlägt, direkt angesteuert werden kann. Konkrete Beispiele für die Optimierung der Substanz finden sich im nachfolgenden Kapitel.

Zum Abschluss dieser exemplarischen Analyse sei noch drauf hingewiesen, dass einige Nutzer eine Wortsuche mit einer Themensuche kombinieren, z.B. *index\* plural* (Top 2302) oder *Komma vor sowie* (Top 1265). Eine solche Suche kann natürlich nur dann zum Erfolg führen, wenn die Informationen auch explizit in den Daten vorhanden sind. Auch diesbezüglich könnte die Substanz optimiert werden.

Interessant sind Listen dieser Art natürlich auch, um etwaige Lemmalücken aufzuspüren. Die Top-2.500-Liste ist allerdings keine gute Quelle für Neuaufnahmen, da die meisten darin enthaltenen Fälle bereits in *Duden online* verzeichnet sind – oder gar nicht hinein sollen, z.B. die Fälle in alter Rechtschreibung. Unter den Top 2.500 findet sich nur ein einziger guter Neuaufnahmekandidat für *Duden online*, und zwar *heidewitzka* (Top 842). Um weitere interessante Kandidaten zu finden, muss man die Sucheingaben jenseits der Top 2.500 durchgehen.

## 7. Konsequenzen aus der Analyse der Sucheingaben

Die Analyse der häufigsten Duden-online-Sucheingaben in Kapitel 6 hat unter anderem gezeigt, dass viele Nutzer Mehrwortlexeme wie *des Weiteren* (Top 7), *zum einen* (Top 10), *vor allem* (Top 14), *unter anderem* (Top 19), *nach wie vor* (Top 71), *im Folgenden* (Top 107) o.Ä. suchen. Der lexikographischen Praxis der Dudenredaktion entsprechend zeigte *Duden online* diese Mehrwortlexeme zum Start im Mai 2011 nur als Beispiele eines bestimmten Lemmas und nicht eigens als Lemma, also *des Weiteren* unter dem Stichwort „weiter“, *zum einen ... zum anderen* unter „ein“, *unter anderem* unter „ander...“, *im Folgenden* unter „folgend“ usw. Anfangs hatten die Artikel „weiter“, „ein“, „ander...“ und „folgend“ einen festen Platz unter den Top 100 der meistgelesenen Artikel. Die genauere Analyse der Sucheingaben (nicht der

meistgelesenen Artikel) brachte das Duden-online-Team auf die Spur, was genau die Nutzer beispielsweise in dem (sehr langen) Artikel „ein“ gesucht haben. Und natürlich auch darauf, wie lange die Nutzer im Zweifel innerhalb des Artikels suchen mussten, um zur gewünschten Information zu gelangen. Um den Nutzerinteressen hier möglichst weit entgegenzukommen, hat das Team beschlossen, den Lemmaansatz ähnlich weit zu fassen, wie es bereits im Rechtschreibduden der Fall ist, und auch solche stark nachgefragten Mehrwortlexeme wie die genannten als Lemmata zu führen, damit sie gut gefunden werden. Die folgenden beiden Artikel sind entstanden, indem ein Beispiel der ursprünglichen Substanz noch einmal als Lemma angesetzt wurde (das an der ursprünglichen Stelle angebrachte Beispiel blieb dort erhalten). Sie sind innerhalb kürzester Zeit in die Top 100 der meistgelesenen Artikel geklettert.

**im Folgenden**

**Häufigkeit:** ■■■■

**Dieser Artikel gehört zu den 100 meistgelesenen Artikeln.**

[› Liste der 100 meistgelesenen Artikel](#)

**Rechtschreibung** [↑ Nach oben](#)

**Worttrennung:**  
im Fol|gen|den

**Bedeutung**

im später Erwähnten, Geschehenden; in den nachfolgenden Ausführungen  
**Beispiel**  
im Folgenden sehen Sie ...

Abb. 7: Beispiel 1 einer Umwandlung vom Beispiel zum Lemma

**zum einen ... zum anderen, zum einen ... zum andern**

**Häufigkeit:** ■■■■■

**Dieser Artikel gehört zu den 100 meistgelesenen Artikeln.**

[› Liste der 100 meistgelesenen Artikel](#)

**Rechtschreibung** [↑ Nach oben](#)

**Worttrennung:**  
zum ei|nen ... zum an|de|ren, zum ei|nen ... zum an|dern

**Beispiel:**  
[Regel 77](#): zum einen hat es geregnet, zum ander[e]n war es zu kalt

Abb. 8: Beispiel 2 einer Umwandlung vom Beispiel zum Lemma

Die Analyse der Sucheingaben hat auch gezeigt, dass viele Nutzer gar keine Lemmata suchen, sondern flektierte Formen eingeben. Auch in dieser Hinsicht hat das Duden-online-Team die ursprüngliche Substanz verbessert, indem es viele häufig gesuchte Flexionsformen, sofern sie in der Substanz nicht schon vorhanden waren und über die Volltextsuche gefunden wurden, systematisch nachgetragen hat. Die Abbildungen 9 und 10 zeigen die alte und die neue Version des Artikels *kennenlernen*. Letztere wurde durch die Aufnahme der extrem häufig angefragten Formen *kennengelernt* und *kennenzulernen* optimiert:

Sie betrachten Version 2.1 des Artikels vom 01.03.2012.

## kennenlernen, kennen lernen

**Wortart:** schwaches Verb

**Häufigkeit:** ■■■■

**Dieser Artikel gehört zu den 100 meistgelesenen Artikeln.**

› Liste der 100 meistgelesenen Artikel

**Rechtschreibung** ↑ Nach oben

**Von Duden empfohlene Schreibung:**

kennenlernen

**Alternative Schreibung:**

kennen lernen

**Worttrennung:**

ken|nen|ler|nen, ken|nen ler|nen

Abb. 9: Der Artikel *kennenlernen* vom 1. März 2012

Sie betrachten Version 2.2 des Artikels vom 13.04.2012.

## kennenlernen, kennen lernen

**Wortart:** schwaches Verb

**Häufigkeit:** ■■■■

**Dieser Artikel gehört zu den 100 meistgelesenen Artikeln.**

› Liste der 100 meistgelesenen Artikel

**Rechtschreibung** ↑ Nach oben

**Von Duden empfohlene Schreibung:**

kennenlernen

**Alternative Schreibung:**

kennen lernen

**Worttrennung:**

ken|nen|ler|nen, ken|nen ler|nen

**Inhalte:**

- › Rechtschreibung
- › Bedeutungsübersicht
- › Wussten Sie schon?
- › Synonyme zu kennenlernen
- › Aussprache
- › Grammatik
- › Typische Verbindung (computergeneriert)
- › Bedeutungen, Beispiele und Wendungen
- › Blättern

**Beispiele:**

es freut mich, Sie **kennenzulernen** oder kennen zu lernen; ich habe ihn gestern **kennengelernt** oder kennen gelernt

**Inhalte:**

- › Drucken
- › Zitieren
- › Wortvorschlag
- › Hilfe zum Wörterbuch
- › Weitersagen

Abb. 10: Der Artikel *kennenlernen* vom 13. April 2012

Die händische Anpassung von Wörterbuchartikeln in der skizzierten Art und Weise lohnt sich, da die Nutzer schneller zum Ziel kommen, dadurch zufriedener sind und die Duden-Website häufig wieder besuchen. Die monatlich steigende Zahl der Besucher auf [www.duden.de](http://www.duden.de) bestärkt das Duden-online-Team darin, an der Kombination aus Marktforschung, Onlineumfragen, Webanalyse und direktem Nutzerkontakt festzuhalten, um auf der Basis der daraus gewonnenen Erkenntnisse die Substanz von *Duden online* kontinuierlich zu verbessern und weiterzuentwickeln.

## 8. Literatur

- Alexa, Melina (2011): Modellierung eines semantischen Wissensnetzes für lexikographische Anwendungen am Beispiel der Duden-Ontologie. Mannheim: Institut für Deutsche Sprache. In: Klosa, Annette/Müller-Spitzer, Carolin (Hg.): Datenmodellierung für Internetwörterbücher. 1. Arbeitsbericht des wissenschaftlichen Netzwerks „Internetlexikografie“. (= OPAL – Online publizierte Arbeiten zur Linguistik 2/2011). Mannheim, S. 61-70.
- Duden – Neues Wörterbuch der Szenesprachen. Internet: <http://szenesprachenwiki.de/> (Stand: Januar 2013).
- DWDS – Das digitale Wörterbuch der deutschen Sprache. Internet: <http://www.dwds.de/> (Stand: Januar 2013).
- Konopka, Anja/Wind, Lutz/Alexa, Melina (in Vorbereitung): *Duden online*: Gibt es einen medienspezifischen lexikografischen Prozess für das Onlinewörterbuch? In: Hildenbrandt, Vera (Hg.): Der lexikografische Prozess bei Internetwörterbüchern. 4. Arbeitsbericht des wissenschaftlichen Netzwerks „Internetlexikografie“. Mannheim: Institut für Deutsche Sprache. (= OPAL – Online publizierte Arbeiten zur Linguistik).
- Kunkel-Razum, Kathrin (2003): Steht das Wort *Muggel* bald im „Duden“? In: Der Sprachdienst 5, S. 161-166.
- Müller-Spitzer, Carolin/Koplenig, Alexander/Töpel, Antje (2011): What makes a good online dictionary? – Empirical insights from an interdisciplinary research project. In: Proceedings of eLex 2011. Trojina, S. 203-208. Internet: <http://www.trojina.si/elex2011/Vsebine/proceedings/eLex2011-27.pdf> (Stand: Januar 2013).
- Scholze-Stubenrecht, Werner (2002): „Die Auswahl der Einträge ist äußerst beliebig.“ Warum *Jagdherr* und *Pokémon* nicht im Duden stehen. In: Sprachwissenschaft 27, 2, S. 225-248.
- Wiktionary – Das freie Wörterbuch. Internet: <http://de.wiktionary.org/wiki/Wiktionary:Hauptseite> (Stand: Januar 2013).

## Das *Oxford English Dictionary* und seine Nutzer

Katrin Thier, *Oxford English Dictionary*, Oxford University Press, Oxford, UK

### 1. Einleitung

#### 1.1 Aufrufe und Leseprogramme

Das *Oxford English Dictionary* (OED) hat eine lange Tradition der Zusammenarbeit mit der Öffentlichkeit, die bis weit vor die Veröffentlichung der ersten Lieferungen zurückgeht. Im Jahre 1857 wurden von der *Philological Society of London* die ersten Konzepte für ein neuartiges Wörterbuch des Englischen vorgestellt, das neben der Gegenwartssprache auch die Sprachgeschichte bearbeiten sollte, das *New English Dictionary on Historical Principles* (NED). Schon zu dieser Zeit wurden erste Aufrufe in Großbritannien, in die Kolonien und nach Nordamerika verschickt, in denen die Öffentlichkeit gebeten wurde, systematisch Belege für den englischen Wortschatz zu sammeln. Diese freiwilligen Helfer wurden gebeten, Bücher zu lesen, Belege auf Zetteln in einem bestimmten Format zu notieren und diese Zettel bei der *Philological Society* einzureichen.

1879 übernahm *Oxford University Press* (OUP) die Publikation des geplanten Werks, und James Murray wurde angestellt, um mit konkreten Arbeiten zu beginnen. In dieser Zeit wurden neue Aufrufe verschickt, vor allem zusammen mit dem Abonnement für das neue Wörterbuch. Um Duplikation zu vermeiden, wurden auch Listen bereits gelesener Bücher zur Verfügung gestellt. Diesmal wurde auch darauf geachtet, eine weitere Auswahl zeitgenössischer Texte mit einzubeziehen, z.B. wissenschaftliche Literatur und Tageszeitungen. Murray richtete in seinem Gartenhaus ein Büro ein, das sogenannte *Scriptorium*, in dem die Zettel alphabetisch sortiert und in kleinen Fächern abgelegt wurden.<sup>1</sup> Sie bildeten die Grundlage für das *New English Dictionary*.

Während der gesamten Zeit, in der das NED publiziert wurde, von 1884 bis 1928 und darüber hinaus bis zur Publikation des ersten Supplements 1933, wurde ein solches Leseprogramm aufrecht erhalten, sodass Murray und seine Nachfolger in der Redaktion immer aktuelles Material zur Verfügung hatten. 1933 wurde das gesamte Werk unter dem Namen *Oxford English Dictionary* neu aufgelegt.

Als in den 1950er Jahren Pläne für ein weiteres Supplement gemacht wurden, ließ man das Leseprogramm wieder auflieben, um Material für dieses Werk zu sammeln, das schließlich 1972-86 erschien. Danach wurde zunächst der gesamte Wörterbuchtext – NED und Supplements – in eine einzige alphabetische Folge gebracht, die mit minimalen Änderungen und Ergänzungen 1989 als *Oxford English Dictionary – Second Edition* (OED2) veröffentlicht wurde. Eine vollständige Überarbeitung des Textes war jedoch bereits vorgesehen, und das Leseprogramm wurde – und wird – weiter geführt. Die Revision begann 1990, und seit März 2000 wird vierteljährlich neues und überarbeitetes Material online für Abonnenten veröffentlicht (OED3).

---

<sup>1</sup> Murray korrespondierte so intensiv mit seinen Helfern, dass vor seinem Haus ein eigener Postbriefkasten aufgestellt wurde. Während das *Scriptorium* heute nicht mehr existiert, ist der Briefkasten nach wie vor in Gebrauch.

Das Leseprogramm ist inzwischen in verschiedene Spezialgebiete aufgespalten: aktuelle britische Texte, aktuelle amerikanische Texte, historische Texte, naturwissenschaftliche Literatur und Sekundärliteratur zur englischen Sprachgeschichte. In allen Bereichen arbeiten inzwischen auch freiberufliche, bezahlte Leser, aber freiwillige, unbezahlte Mitarbeiter sind nach wie vor beteiligt – zum Teil auf gelegentlicher und zum Teil auf regelmäßiger Basis. In den meisten Fällen werden jedoch heute keine Zettel mehr geschrieben, sondern die Zitate werden direkt in eine XML-Datenbank eingetippt, in einem Format, das es erlaubt, sie später direkt in das Wörterbuch zu importieren.<sup>2</sup>

Neben den ersten allgemeinen Aufrufen und den Leseprogrammen, die sich aus ihnen entwickelten, sind auch in der gesamten Geschichte des Wörterbuchs immer wieder spezifischere Aufrufe an die weitere Öffentlichkeit gemacht worden, um Informationen zu bestimmten Wörtern und Ausdrücken zu finden. Diese Aufrufe hatten zunächst die Form von Pamphleten, Flugblättern und Zeitschriftenartikeln; später wurden auch neuere Medien mit einbezogen. Die jüngste Kampagne begann im Herbst 2012 und erscheint im Internet im Format eines Blog, der regelmäßig ergänzt wird (*OED Appeals*: <http://public.oed.com/appeals/>). Darin bittet die Redaktion um Nachweise oder Information zu bestimmten Wörtern, zu denen die weitere Bevölkerung möglicherweise besseren Zugang hat, zum Beispiel in Lokalzeitungen, Zeitschriften für bestimmte Interessensgruppen, Romanen, Briefen usw., die nicht alle systematisch gelesen werden können. Die einzelnen Blog-Einträge beschreiben die jeweiligen Wörter, und Leser können in Kommentarfeldern neue Informationen einreichen; wenn nötig, kann die Redaktion auch auf diese Kommentare antworten.

## 1.2 Öffentlichkeitsarbeit

Heute genießt das *Oxford English Dictionary* ein hohes Ansehen in der Öffentlichkeit. Es ist eine zentrale Institution in den Augen vieler englischsprachiger Menschen, nicht nur in Europa, und wird häufig als Autorität für „richtiges“ Englisch angesehen – obwohl es rein deskriptiv ist, also diese Autorität nicht für sich beansprucht. Dieses Ansehen geht zum Teil darauf zurück, dass Kontakt zur weiteren Öffentlichkeit ständig aufrecht erhalten wird, sowohl durch die Presse als auch über neuere Medien.

Jedes Vierteljahr, wenn ein neuer Teil des revidierten Textes veröffentlicht wird, wird eine Pressemitteilung über neue und interessante Details herausgegeben, die oft in Berichten in Zeitungen, Radio und Fernsehen aufgegriffen werden. Andersherum treten die Medien auch von sich aus an OUPs *Publicity Department* heran, wenn sie Information oder Experten zu sprachlichen Themen suchen. Mitglieder der Redaktion geben regelmäßig Interviews im Radio und gelegentlich auch im Fernsehen. Im Internet hat das *Oxford English Dictionary* eine eigene Präsenz auf Twitter (vgl. Abb. 1) und, zusammen mit anderen Wörterbüchern des Verlags, auf Facebook. Beiträge zum OED finden sich auch auf einem YouTube-Kanal von *Oxford University Press*, und einige Redakteure schreiben Beiträge für das *Oxford Words Blog* der Wörterbuchprojekte. Diese Dienste können ein unterschiedliches Publikum erreichen; sie sind aber auch miteinander verlinkt, sodass alle Nutzer die Möglichkeit haben, auf dieselben Inhalte zuzugreifen:

---

<sup>2</sup> Mehr Information zu den Leseprogrammen findet sich hier: <http://public.oed.com/history-of-the-oed/reading-programme/>.



Abb. 1: Das Profilbild des OED auf Twitter

Über Twitter (<http://twitter.com/OED>) wird regelmäßig ein *Word of the Day* publik gemacht und auch kurze Informationen zu anderen Wörtern, manchmal im Zusammenhang mit aktuellen Themen (etwa zum Tagesgeschehen, zu Jubiläen oder zur Jahreszeit). Leser können diese Nachrichten kommentieren und auch selber Fragen stellen. Auch zu den Aufrufen der neuen *Appeals*-Liste gibt es hier Verweise (vgl. Abb. 2). Nicht ganz so kurze Texte zu Themen aus der Sprachgeschichte und zum Sprachgebrauch werden im *Oxford Words Blog* (<http://blog.oxforddictionaries.com/>) publiziert; auch sie sind oft von Twitter aus erreichbar. Hier haben Leser ebenfalls die Möglichkeit, Kommentare zu hinterlassen. OUP hat auch einen weiteren, allgemeineren Blog (<http://blog.oup.com/>), mit dem das OED jedoch relativ selten direkt in Berührung kommt. Im YouTube-Kanal von OUPs akademischer Abteilung (<http://www.youtube.com/user/OUPAcademic>) stehen eine Anzahl von Videos zum OED. Viele davon sind für diesen Zweck gedreht (z.B. zur Publikation neuer Wörter), andere sind Gebrauchsanleitungen, und einige waren ursprünglich ungesendete Fernsehbeiträge. Die Facebookseite von OUPs Wörterbüchern (<http://www.facebook.com/OxfordDictionaries>) ist hauptsächlich ein Portal mit Hyperlinks zu den anderen Angeboten. All diese Dienste werden von der Marketingabteilung betrieben; viele der Inhalte werden jedoch von Wörterbuchmitarbeitern zur Verfügung gestellt.

**The OED @OED**

**OED** Calling all #snowboarding and #skiing fans: can you help us with a word hunt? [oxford.ly/13SQIJU](http://oxford.ly/13SQIJU) #OEDAppeals

[Expand](#)

Abb. 2: Ein Tweet mit einem Aufruf (das Link führt zum Eintrag für *Superpipe* in der Appeals-Liste)

Ein wichtiges, wenn auch weniger interaktives Kommunikationsmittel ist natürlich auch die Webseite des OED selber (<http://www.oed.com>). Sie enthält neben dem kostenpflichtigen und passwortgeschützten Wörterbuchtext auch eine Anzahl frei zugänglicher Seiten mit Material

zur Geschichte der englischen Sprache und zur Geschichte des Wörterbuchs, aber auch z.B. mit Informationen und Lehrplanvorschlägen für Lehrer, die das Wörterbuch im (muttersprachlichen) Englischunterricht in der Schule verwenden wollen.

Der Wörterbuchtext selber ist zwar im Prinzip kostenpflichtig, ist aber dennoch vielen Menschen frei zugänglich. Weltweit haben viele Institutionen, wie etwa Schulen und Universitäten, ein Abonnement für ihre Mitglieder, und in Großbritannien zusätzlich auch die meisten öffentlichen Bibliotheken. Dieser Zugang kann sowohl innerhalb der Institution als auch extern vom eigenen Computer aus genutzt werden. Da die Mitgliedschaft in kommunalen öffentlichen Bibliotheken in Großbritannien kostenlos ist, heißt das, dass die meisten Bürger dort das gesamte OED umsonst nutzen können.

Schließlich nehmen manche Redaktionsmitglieder an Konferenzen teil, halten Vorträge und veröffentlichen Artikel, nicht nur in wissenschaftlichen Publikationen. Einige Mitarbeiter haben auch Bücher geschrieben, sowohl zur Geschichte des Wörterbuchs als auch zu ihrem eigenen Arbeitsfeld. Um nur zwei sehr unterschiedliche Beispiele zu nennen (beide von OUP publiziert): *The Ring of Words* (von P. Gilliver, J. Marshall & E. Weiner 2006) berichtet über J. R. R. Tolkiens Arbeit als OED-Redakteur in den 1920er Jahren, während das *Oxford Handbook of Etymology* (von P. Durkin 2009) auf die lange Erfahrung des Autors als *Principal Etymologist* in der Wörterbuchredaktion zurückgreift.

## 2. Beispiele

Seit den ersten Aufrufen im 19. Jahrhundert hat sich manches geändert, und viele Werke sind bereits durch Leseprogramme ausgewertet worden. Zusätzlich ist inzwischen auch eine Vielfalt von Texten aller Art und fast jeden Alters elektronisch zugänglich und kann von Redakteuren einfach und schnell durchsucht werden. Die Mitarbeit der Öffentlichkeit ist jedoch nach wie vor wichtig, schon allein weil nie alle Texte systematisch gelesen oder in Datenbanken durchsucht werden können. Zusätzlich ist die Auswahl von Texten in Datenbanken oft beschränkt auf Bücher und größere Zeitungen oder Zeitschriften; unkonventionell publiziertes Material und Ephemera wie etwa Speisekarten werden seltener berücksichtigt. Die zur Zeit rein elektronische Publikationsweise des OED3 ermöglicht es, auch zu schon revidierten Einträgen jederzeit zurückzukehren und gegebenenfalls neues Material hinzuzufügen.

### 2.1 Fernsehen: *Wordhunt*

In den Jahren 2006 und 2007 wurden im BBC zwei Staffeln einer Fernsehserie über Wörter und das OED gesendet, mit insgesamt 16 Folgen, unter dem Namen *Balderdash & Piffle*,<sup>3</sup> Ein zentraler Teil dieser Serie war ein Aufruf an die Öffentlichkeit, frühere Belege für bestimmte Wörter zu finden; unter dem Titel *Wordhunt* wurde die erste Liste von Aufrufen schon 2005 von der Pressestelle der BBC veröffentlicht.<sup>4</sup> Diese Wörter waren von der Redaktion ausgesucht worden, weil aus verschiedenen Gründen die Vermutung nahelag, dass sie älter sind als zu jener Zeit im OED belegt, und dass sie gerade in solchen Texten vertreten sein können, die nicht in den großen Datenbanken erscheinen, aber Menschen im alltäglichen Leben begegnen. Auf der Liste standen unter anderem umgangssprachliche Ausdrücke (z.B.

<sup>3</sup> Zwei Wörter für (gesprochenen) Unsinn (etwa ‘Quatsch’), die auch gelegentlich zusammen verwendet wurden.

<sup>4</sup> [http://www.bbc.co.uk/pressoffice/pressreleases/stories/2005/06\\_june/08/wordhunt.shtml](http://www.bbc.co.uk/pressoffice/pressreleases/stories/2005/06_june/08/wordhunt.shtml).

*Old Bill*, kollektiv ‘Polizei’) und Slang (z.B. *dosh* ‘Geld’), aber auch z.B. die Namen von beliebten Gerichten (z.B. *ploughman’s lunch*). In den einzelnen Sendungen wurden die Ergebnisse präsentiert (zusammen mit einigem Hintergrundmaterial) und einer Jury aus OED-Mitarbeitern vorgelegt, die entscheiden sollten, ob der neue Beleg für das Wörterbuch geeignet ist. Dazwischen wurden kleine Filme gezeigt, in denen Prominente die Geschichte bekannter und interessanter Wörter verfolgten (z.B. *cocktail*). In der ersten Staffel war die Auswahl von Wörtern für jede Folge relativ locker; in der zweiten Staffel hatte jede Folge ein bestimmtes Thema. Zu der Sendereihe wurden 2006 auch von BBC Books (nicht von der Wörterbuchredaktion) zwei Begleitbücher publiziert.

Als Beispiel mag das Substantiv *domestic* ‘(gewalttätige) häusliche Auseinandersetzung’, ‘Zwischenfall mit häuslicher Gewalt’<sup>5</sup> dienen, das 2005 noch nicht mit dieser Bedeutung im OED verzeichnet war. Der Ausdruck stammt aus dem Polizeislang und ist in der weiteren Bevölkerung vor allem durch Krimalliteratur und -filme bekannt. Als Adjektiv geht *domestic* in die frühe Neuzeit zurück; es kann vieles bezeichnen, das mit dem Haushalt oder dem häuslichen Leben zu tun hat, etwa *domestic service* ‘Dienst im Haushalt’, oder *domestic science* ‘Hauswirtschaftslehre’. Entsprechend ist das Wort auch früher schon als Substantiv gebraucht worden, z.B. in den Bedeutungen ‘Mitglied eines Haushalts’ (1539) und ‘Bediensstter im Haushalt’ (1623). In ähnlicher Weise kommt der Substantivgebrauch in der Polizeisprache von Ausdrücken wie *domestic violence* ‘häusliche Gewalt’ (1891) her.

Das Substantiv *domestic* in diesem Zusammenhang war der Redaktion durch die ehemalige britische Krimiserie *Z-Cars* bekannt, die 1962 anlief. Der Erstbeleg allerdings kam nicht aus einer britischen Quelle, sondern aus einer australischen Fernsehzeitschrift von 1963, in der für australische Zuschauer britisches Vokabular erklärt wurde. Als Ergebnis der *Wordhunt* erhielt das OED einen Beleg aus einem Manuscript des Drehbuchs zur ersten Folge, datiert 1961, also sogar früher als die erste Sendung:

1961 T. K. MARTIN *Z Cars* (B.B.C. TV script) Episode 1 (2nd draft) 21 You just sent Seaport Crime Car to that domestic, here at Newtown?

Es ist allerdings nicht gelungen, das Wort über fiktive Texte hinaus in die Sprache der wirklichen Polizei hin zu verfolgen, wahrscheinlich weil es doch im Urprung der rein gesprochenen Sprache angehört. Für die Sendung wurde auch ein (datierbares) Polizeinotizbuch eingereicht, es ist allerdings letztendlich nicht als Nachweis für dieses Lemma zugelassen worden, da das Wort ohne Zusammenhang gebraucht war, und nicht zu entscheiden war, ob es sich nur um das Adjektiv (als verkürzte Notiz oder Gedankenstütze) oder schon um das Substantiv handelte.<sup>6</sup>

<sup>5</sup> OED3 DOMESTIC adj. & n., Draft additions June 2007: (n.) “orig. and chiefly Police slang (orig. Brit.). A domestic disturbance, spec. one meriting police intervention; a (violent) quarrel between members of the same household, esp. between husband and wife.”

<sup>6</sup> <http://www.experts123.com/q/why-didnt-the-oed-accept-the-police-notebook-pictured-above-as-evidence-of-domestic.html>  
– die Information stammt ursprünglich von einer Seite des BBC, die nicht mehr im Internet verfügbar ist.

## 2.2 Internet: *Science Fiction Citations*

Während die *Wordhunt* gezielt eine kleine Zahl bestimmter Wörter zum Thema hatte, zeigt das folgende Beispiel einen weniger strukturierten Aufruf, der dadurch eine größere Anzahl Wörter erfassen konnte, wenn auch weniger intensiv.

Das Internet ermöglicht weltweiten Kontakt zu Menschen mit vielen unterschiedlichen Interessen und aus unterschiedlichen Fachgebieten, die bei weitem nicht nur akademischer Natur sind. 2001 wurde von Science-Fiction-Fans in Zusammenarbeit mit dem amerikanischen Zweig der OED-Redaktion die Webseite *Science Fiction Citations* eingerichtet (<https://gabdictonaries.oup.com/sf/>); die ursprüngliche Idee stammte von einem Mitarbeiter eines Leseprogramms. Hier können Leser von Science-Fiction-Literatur interessante Wörter einreichen, die ihnen bei der Lektüre oder Diskussion von Texten dieser Gattung mehr oder weniger zufällig begegnet sind. Anders als bei den Leseprogrammen werden hier nicht systematisch Bücher auf relevantes Material hin durchsucht, vielmehr sollen Interessierte, die mit dem Genre vertraut sind, ihr Wissen aus ihrer alltäglichen Unterhaltung teilen. Das Ziel ist vor allem, frühere Erstbelege für bekannte Wörter zu finden, aber mit der Zeit sind der Seite auch neue Wörter hinzugefügt worden. Die Wortliste ist in drei Kategorien unterteilt: *Science Fiction* (Primärtexte), *SF Criticism* (literaturwissenschaftliche Sekundärliteratur), und *SF Fandom* (informeller Diskurs unter Fans). In jedem Bereich steht eine alphabetische Wortliste mit Informationen zur Beleglage, z.B. das Datum des zur Zeit frühesten bekannten Belegs, der vordatiert werden soll, und eine kurze Beschreibung der Bedeutung. Die Stichwörter in diesen Listen sind Hyperlinks, die jeweils zu einer Seite mit mehr Informationen führen, wie z.B. dem Datum des Erstbelegs im OED. Hier werden auch alle inzwischen eingereichten Belege zitiert. Eine eigene Einführungsseite erklärt, welche Informationen ein Beleg enthalten muss, und alle neuen Belege werden per E-Mail bei den Moderatoren eingereicht (die Seite ist kein Wiki). Diese Belege erscheinen dann auf der Webseite in einer Form, die OED-Redakteure ohne große Probleme in den Zitierstil des Wörterbuchs „übersetzen“ können.

Als Beispiel kann hier der Eintrag für FANZINE dienen, eine Art Zeitschrift von und für Fans, die nicht professionell publiziert ist. Das Wort ist aus *fan* ‘Fan’ und dem zweiten Teil von *magazine* ‘Zeitschrift’ zusammengesetzt. Im OED erschien es zum ersten Mal 1972 im Supplement und daher auch im OED2; der Erstbeleg ist von 1949. Auf der *SF Citations*-Webseite steht das Wort unter *SF Fandom*, da Fanzines ein Medium des Fandiskurses sind. Benutzer der Webseite haben eine Reihe früherer Belege gefunden, der erste ist jetzt von 1940: Hier wird diskutiert, wie so eine Publikation genannt werden soll; es ist also wahrscheinlich einer der frühesten Belege überhaupt (vgl. Abb. 3). Es gibt auch eine verkürzte Form dieses Wortes, ZINE n., mit einem Erstbeleg von 1965 im Supplement und OED2; die *SF-Citations*-Seite zeigt inzwischen Nachweise von 1946. Diese illustrieren zusätzlich, dass das Wort auch *'zine* geschrieben werden kann (mit Apostroph am Anfang), eine Form, die das OED zur Zeit noch nicht erwähnt. Beide Einträge sind noch nicht revidiert, aber wenn die Zeit dafür kommt, kann die Redaktion auf die Webseite zurückgreifen und somit auf Quellen, die durch Leseprogramme oder Datenbankrecherche nicht oder nur schwierig hätten erschlossen werden können.

<b>Definition</b>	a magazine for fans, esp. those of science fiction
<b>OED requirements</b>	antedating 1940
<b>Earliest cite</b>	L. Chauvenet 'Detours'
<b>Comment</b>	Mike Christie submitted a 1952 cite from a letter to Planet Stories by Algis Budrys. Geri Sullivan submitted a 1944 cite from the Fancyclopedia. Ruth Berman submitted a 1942 cite from H.H. Holmes "Rocket to the Morgue". Keith Stokes submitted a 1941 cite from Bob Tucker and Everett Evans' fanzine "Le Zombie". Cory Panshin submitted a 1970 cite from Rolling Stone.  Hal Hall located a quote in Sam Moskowitz' research files from the October 1940 issue of Louis Russell Chauvenet's fanzine, "Detours". Alistair Durie located a copy and provided a cite from it. This appears to be the coinage of the term.  Earliest cite in the OED: 1949.
<b>Last modified</b>	6 July, 2008
<b>1940 C. CHAUVENET in Detours Oct. 6</b>	We hereby protest against the un-euphonious word 'fanag' and announce our intention to plug <i>fanzine</i> as the best short form of 'fan-magazine'.

Abb. 3: Von der Webseite *Science Fiction Citations*: Kurzbeschreibung für den Eintrag *fanzine* und (aus demselben Eintrag) der neu gefundene Erstbeleg

Aus dem Projekt *Science Fiction Citations* ist auch ein Buch hervorgegangen: Prucher, J. (Hg.), *Brave New Words – The Oxford Dictionary of Science Fiction* (OUP 2007), das einen *Hugo Award* als bestes Sachbuch zum SF Genre gewann. Die *SF-Citations*-Webseite ist inzwischen weniger aktiv, da schon für viele Einträge die Grenzen des Projekts erreicht sind, sie wird aber nach wie vor unterhalten.

## 2.3 Spontane Beiträge

Da das OED im öffentlichen Bewusstsein präsent ist, gibt es auch viele Menschen, die außerhalb spezifischer Projekte spontan Beiträge einreichen. Oft handelt es sich dabei um frühere Belege, manchmal aber auch um Vorschläge zur Etymologie. Die Beiträge kommen sowohl aus der breiteren Öffentlichkeit als auch von Akademikern, die in ihrem Forschungsgebiet Zugang zu ungewöhnlichen Quellen haben. Damit solche Nutzervorschläge möglichst gut genutzt werden können, steht auf der Webseite von OUP ein detailliertes Formular mit Anleitung zur Verfügung, in das so viel Information wie möglich eingetragen werden kann (<http://www.oup.com/uk/oedsubform/>). Daneben erhält die Redaktion auch informelle Zuschriften in Form von Briefen und E-Mails.

Als Beispiel für einen solchen spontanen Beitrag kann die Vordatierung von *pirate* im Zusammenhang mit Radiosendern dienen. Die revidierte Fassung von *PIRATE* n. wurde 2006 publiziert, und der früheste Beleg für eine Zusammensetzung mit der Bedeutung 'illegaler Rundfunksender, Piratensender'<sup>7</sup> kam aus der *Times* von 1928, die von einer '*pirate*' broadcasting station spricht (vgl. Abb. 4). Nicht lange davor hatte das *Museum of the History of Science* der Universität Oxford den Nachlass von Guglielmo Marconi, dem Erfinder des Radios, erhalten, und die Papiere werden seitdem aufgearbeitet.<sup>8</sup> 2010 bekam die OED-Redaktion eine E-Mail von einer beteiligten Wissenschaftlerin mit einem Scan eines Briefs von 1903 im Anhang. Dieser Text enthielt das Wort *pirate-station*, wenige Jahre nach Einführung

<sup>7</sup> OED3 *PIRATE* n. C3.c. "Designating a radio or television station that broadcasts without a licence, esp. from a position outside the territorial waters of the receiving country[...]".

<sup>8</sup> <http://www.mhs.ox.ac.uk/marconi/>.

der Technologie, und ein Vierteljahrhundert früher als der oben erwähnte Zeitungsartikel. Da der Brief nicht veröffentlicht ist, wäre es der Redaktion nahezu unmöglich gewesen, diesen Beleg unabhängig zu finden.

Das erste Zitat für diesen Gebrauch von PIRATE n. lautet jetzt:

1903 H. C. Hall *Let.* 5 Sept. (MS Marconi 175, Marconi Coll., Mus. Hist. Sci., Oxf.) 4 “So long as pirate-stations exist, we cannot even guarantee the Admiralty against our own Stations.” (vgl. Abb. 4)

- 1903 H. C. HALL *Let.* 5 Sept. (MS Marconi 175, Marconi Coll., Mus. Hist. Sci., Oxf.)  
4 So long as pirate-stations exist, we cannot even guarantee the Admiralty  
against our own Stations.
- 1928 *Times* 18 Dec. 14/2 Officials of the Post Office are taking immediate steps to  
locate the unknown ‘pirate’ broadcasting station.

Abb. 4: Im OED-Eintrag jetzt nebeneinander: der frühere und der neue Erstbeleg für Komposita mit der Bedeutung ‘Piratensender’

### 3. Zum Schluss

Seit den Anfängen in der Mitte des 19. Jahrhunderts hat das *Oxford English Dictionary* mit der Öffentlichkeit zusammengearbeitet, und die Beiträge vieler freiwilliger, meist unbezahlter und oft nur gelegentlicher Helfer bilden die Basis des Werkes. Die Beiträge stammen bei weitem nicht nur von Akademikern, sondern von Menschen aus allen Teilen der Bevölkerung, die sich für ihre Sprache interessieren. Die technologischen Entwicklungen der letzten Jahrzehnte haben einerseits durch Datenbanken und ähnliche Ressourcen viel Quellenmaterial einfacher und direkter zugänglich gemacht, sie haben andererseits aber auch den Kontakt mit der Öffentlichkeit erleichtert, die, wie oben gezeigt, nach wie vor wichtige Beiträge leistet. Ohne die Mitarbeit dieser Menschen wäre das *Oxford English Dictionary* heute nicht das, was es ist.