Online Dictionaries as Emergent Archives of Contemporary Usage and Collaborative Codification

Colleen Cotter
Queen Mary, University of London

John Damaso
Arizona State University

February 2007

Abstract

Within the history of modern English lexicography, individual dictionary editors have had ultimate control over the selection, meaning, and illustration of words; extensive collaboration with contributors has been limited. However, Internet technologies that easily permit exchanges between a user and a database have allowed a new type of dictionary online, one that is built by the collaboration of contributing end-users, allowing ordinary users of dictionaries who are not trained lexicographers to engage in dictionary-making. We discuss a popular online slang dictionary called UrbanDictionary.com (UD) to illustrate how lexicographic principles are joined with Web-only communication technologies to provide a context for collaborative engagement and meaning-making; and to note the many characteristics and functions shared with the traditional print dictionary. Significantly, UD captures what most traditional English dictionaries fall short of: recording ephemeral quotidian spoken language and representing popular views of meaning. By relying on the users of language to select and define words for a dictionary, UD, which defines more than 1 million words, has in effect influenced access to and formulation of the lexis.

Keywords computer-mediated communication, lexicography, slang, youth language; English
1 Introduction

English lexicography stems from a tradition of relatively limited functional collaboration, beginning with Samuel Johnson's dictionary in 1755, in which editors overseeing numerous contributors held the ultimate authority over the selection, meaning, and illustration of words. Today's online dictionaries of contemporary English usage develop and expand this collaboration, and in the process extend the parameters of the dictionary genre and the channels of transmission through which codification of contemporary usage occurs.

We discuss an online collaborative slang dictionary called UrbanDictionary.com (UD), identifying this “new populist dictionary” (Damaso 2005) as an emergent dictionary genre that joins lexicographic principles with Web-only communication technologies to provide a context in which users collaborate, cooperate, and compete for meaning-making. The collaborative opportunities inherent in dictionaries like UD distinguish them from traditional print dictionaries in that an authoritative editor is replaced by what can be seen as a large-scale usage panel. At the same time, their creation shares many characteristics with the traditional dictionary.

2 UrbanDictionary.com

UrbanDictionary.com (UD) is an online dictionary of contemporary English slang usage, a collaborative project of over 1 million definitions for over 400,000 unique headwords. It was created in 2000 by then-college student Aaron Peckham, who described the origins of the project in a radio interview in 2004:

UrbanDictionary was originally conceived as a parody of dictionary.com because I noticed that dictionary.com didn't emphasize slang words and the origination of slang words – like which part of the country they came from […] Originally I just put in words I was hearing among my friends, and then I sent the link to my friends; they sent it to their friends’ friends, and eventually it spread around the world. (National Public Radio, Jan 17, 2004; emphasis added)

Peckham was originally challenging the authority paradigm of lexicographic tradition although soon his project would become collaborative and its Web interface and interactive features would change. UD, like most online dictionaries (save OED Online and special expanded editions of free-use dictionaries), is a free service and available online.

A single description of UD is complicated by its transitory form and function and underlying social dimensions, characteristics it shares with that which it seeks to record, English slang. Although a comprehensive profile of UD would include a range of data-gathering methods involving site interface and design evaluation, a corpus of dictionary entries themselves, interviews with the Moderator and volunteer Editors, and an ethnographic account of various participant practices involving word selection and deletion (cf. Damaso 2005), we focus here on the way UD compares with general-purpose dictionaries and privileges the user-author.

While UD can be differentiated from paper dictionaries not only in form and function and the way it is compiled and written, it nonetheless assumes many of the same methodological strategies of traditional dictionaries and reproduces elements from several traditional lexicographic genres, adding its own features involving immediacy and group action that derive from its communicative technologies. Significantly, UD captures what most traditional dictionaries fall short of: recording ephemeral quotidian spoken language, and representing popular and divergent, as opposed to authorized and uniform,
views of meaning. Additionally, the functionality of UD – namely its various Web-based communicative channels, dictionary additions occurring nearly in real-time, its user feedback and control capacity (through its Edit feature), and its non-specialist lexicographic team (of self-appointed users) – contributes to its uniqueness.

Other factors make UD itself worth examining: First, while there are other accessible online slang dictionaries, UD is one of the most popular. It ranks consistently in the top 1,500 websites visited each day with about thirty times the number of page views as Wiktionary, another collaboratively authored dictionary. For example, in 2006, UD received 50 million site visits. Second, UD places an emphasis on democracy and equal access to meaning-making rights: from its inception in 2000 (although less so now) anyone could contribute, and anyone could edit or have a say in the formulation of the dictionary and the inclusion of words. By relying on the users of language to select and define words for a dictionary, UD has equalized access to and formulation of the lexis.

Third, while the notion of collaboration in lexicography is not new, most histories of dictionaries pay little critical attention to the functions of the communities behind them. Individual labor is emphasized in the scholarship on early English lexicography and collaboration is described as limited – most often by financial resources, deadlines, and physical space (Reddick 2005). Instead, we have only anecdotal reports of Dr. Johnson’s scribes (Reddick 1990, 2005), or the relationship of James Murray and his prodigious contributor William Chester Minor for the *Oxford English Dictionary (OED)* (Winchester 1998). More often, accounts emphasize the individual labor or personality of the editor, such as *Webster’s Third* editor Philip Gove’s “formidable presence” (Morton 1994: 72). For his nine-year, 40,000-word dictionary, Samuel Johnson oversaw eight amanuenses, scribes charged with copying passages marked by Johnson and arranging them in his ever-expanding manuscripts. OED editor James Murray’s famed Mill Hill scriptorium housed thousands of pigeonholes he used to file the illustrative sentences amassed during his editorship. With UD, the medium of the Web allows unlimited contribution and expansion – computer speed and storage being the only potential obstacles. And in the case of UD, its authors are its users, site owner-moderator Peckham playing a background role.

3 UD as dictionary

General-purpose dictionaries serve several functions which UD shares, such as authorizing usage, storing vocabulary, improving communication, strengthening the language, and affording metalinguistic reflections on language (cf. Hartmann 1987). In terms of function, UD users visit UD to determine how people use the language and often find the information credible and even authoritative. For usage judgments in the 18th century, Samuel Johnson relied on exemplary writers, and for UD in the 21st century it is anyone who accesses the Internet. As a vocabulary repository, UD stores the words of contemporary popular culture, the retention and application of which empower UD users with a culturally relevant tool to communicate, especially in other online domains such as chat rooms, discussion boards, and blogs. In terms of its relevance to enhancing language vitality, some volunteer Editors found UD to be a celebration of the proliferation of meaning while others viewed it as a dilution. When UD users read others’ definitions and decide to recommend them for deletion or to add their own (or to make a Thumbs Up/Down vote), UD serves as a metalinguistic prompt, a feature evident at all levels of UD practice, from user definition-submissions to Editor Talk discussions online.

When compared historically to other dictionaries, UD, which is “encyclopedic” in form (cf. Algeo 1990) because it includes proper names and images as well as the
technological potential to contain other non-lexical items such as maps and sounds, is similar in less obvious ways. Users often use UD to **scold, preach, mock, and fight** just like the idiosyncratic lexicographers of the 19th and early 20th centuries such as Samuel Johnson, Pierre Larousse, and Émile Littré (Bejoint 1994). Just as **dictionary plagiarism** influenced the attitudes and practices of Webster and Worcester (Micklethwait 2000), UD Editors note the numerous submissions which excerpt (“copy-and-paste”) from other online resources. The perception of **anonymously written definitions** is somewhat preserved on UD. In the same way that the Thompson sisters desired anonymity when they contributed tens of thousands of slips to the **OED** (Winchester 2003: 214), UD contributors and Editors can have anonymity behind screennames and pseudonyms. **OED** Editor James Murray’s *Appeal for Readers* to investigate words is analogous to the UD moderator’s request for Words of the Day suggestions. Murray listed readers’ contributions in order of how many entries they edit. In both cases, **volume is valued**.

There is similarity even to the extent of collaboration despite the noted differences. The interactions of UD Editors in the Talk and Chat online spaces are comparable to those of earlier dictionary makers who relied on correspondence with editors to communicate frequently. **OED**’s Murray did not often meet face-to-face with his collaborators, Bradley and Craigie; they, like their predecessors, worked apart and communicated by letter about their editorial work (Burchfield 1987). Irrespective of UD, even today’s general-purpose dictionaries are written on-line, and editors do not need to work together in the same place (Landau 2001).

UD distinguishes itself by virtue of the nature of its collaboration via the online medium. Dictionary editors often receive letters from readers including lists of suggested neologisms written by the readers themselves (Landau 1999: 294); UD removes this letter-writing step and users can freely upload these neologisms individually. Whereas traditional lexicographers apply defining principles (Zgusta 1971) or engage in “good lexicographic practice” (Landau 1999: 124) as a definition is authored, UD applies guidelines _after_ authoring definitions in an interactive editorial process. Béjoint contends that “every lexicographer knows that true exhaustiveness is impossible” (Béjoint 2001: 180), but this may not be the case with UD, which, because of its low maintenance costs, free staff, and simple interface, can evolve indefinitely as words change, acquire new meanings, or drop out of the lexicon. In fact, online dictionaries have “the potential of never being out of date, and can as such represent the ultimate dynamic repository of knowledge” (de Schryver 2003: 157).

### 4 Slang lexicography

For the most part, general-purpose dictionaries have focused on “hard words” and “common words,” and the words of contemporary usage, often considered fleeting vestiges of fashion, are usually relegated to the status of “slang.” As a result, few “slang” words are knowingly accepted by editors for inclusion into general-purpose dictionaries (Algeo 1989). For example, James Murray insisted in the 19th century that there be “no slang, no dialect, no coarseness, no recent coinage…considered jargon” in the **Oxford English Dictionary** (Skelton-Foord 1989: 37). A century earlier, Dr. Samuel Johnson cited in both his 1747 *Plan* and 1755 *Preface of the Dictionary of the English Language* that he would avoid including “low bad words” in his work.

Nonetheless, the application of lexicographic principles to the creation of slang dictionaries has a long history. The cant dictionaries of the 17th-19th centuries recorded the often-secret codes of societal misfits and underground criminal networks (cf.
Coleman 2004, 2005). With the diffusion of English, some lexicographers have compiled local lexicons such as Boontling in pioneer California (Adams 1971), Pittsburghese in contemporary Pennsylvania (Johnstone and Baumgardt 2004), or Cockney Rhyming Slang in London (Lillo 2001), while others have focused on various segments of the populace in recording jargon (e.g., cowboys, hippies, the military, ham radio operators).

Each UD contributor works from a potentially different conception of what slang is, thus allowing for wide variation among UD entries, whereas in traditional print lexicography, each slang dictionary posits a definition of “slang” in its preface and follows it when evaluating words for inclusion. Eric Partridge’s A Dictionary of Slang and Unconventional English, first published in 1937, divides words into six categories. Other editors favor a narrower definition of slang, minimizing the appearance of jargon (Green 1984), or stress the individual psychology of slang speakers (Chapman 1986), or claim that understanding that “a definition of slang that confines itself to stylistic traits […] will necessarily remain inadequate” (Lighter 1994: xi).

The flexibility afforded by UD’s online format and idiosyncratic contribution methods of the thousands of often-disparate users makes a rigorous definition of slang a moot point. While traditional print slang dictionaries determine a definition of “slang,” isolate a corpus, and abide by physical parameters determined by publishers and their deadlines, UD grows with the language and with the evolving notion of what the contributors themselves consider slang.

5 UD Word-formation processes

The same word-formation processes for general vocabulary, such as compounding, doubling, shortening (initialism, acronym), blending, and allusion, have been similarly established for slang (Eble 1996). Table 1 displays examples of these processes as they occur throughout the UD corpus, in addition to contemporary borrowings and allusions to cultural artifacts relevant to UD users.

The examples in Table 1 reveal a group of users interested in lexical innovation, achieved through the same processes of word-formation available within Standard English.1 The online setting produces many examples of shortening (val pal, TMI, WTF, BTW), often derivative of abbreviated forms suited for quick exchanges on the Web, especially in IRC or IM. Especially popular throughout UD are blends (folex, mesbian, askhole); they often receive the greatest Thumbs Up/Down ratio because wordplay, especially resulting from the astute combination of morphemes, is highly valued. The metaphors used by UD contributors (break the glass, circle the drain) often belong to the slang semantic field called “destruction” (Eble 1996: 44) while allusions are typically drawn from the entertainment fields, such as sports (full court press), film (Death Star), and gaming (dish jenga), the strong presence of allusion in UD entries signaling group membership. To understand that a “Death Star” is an impenetrable building, UD users, for example, must know that the “Death Star” was the immense space station in the film series Star Wars.

---

1 Within the examples in Table 1, mechanical errors in spelling, spacing, and punctuation belong to the data as created by UD users. Examples in Section 6 are also excerpted verbatim from UD and mechanical errors remain.
Table 1: Slang-Formation Processes (adapted from Eble 1996) in UrbanDictionary.com

<table>
<thead>
<tr>
<th>Process</th>
<th>UD Example</th>
<th>UrbanDictionary.com Author's Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compounding</td>
<td>earjacking</td>
<td>• Eavesdropping on a conversation that you have no business hearing.</td>
</tr>
<tr>
<td></td>
<td>wet backs</td>
<td>• Mexicans coming into the U.S.A by swimming</td>
</tr>
<tr>
<td>Doubling</td>
<td>job job</td>
<td>• 1 aka a real job; a job with health insurance; as opposed to an internship or the wonderful world of retail; 2 a primary job; a day job</td>
</tr>
<tr>
<td>Blending</td>
<td>Chevrolegs</td>
<td>• The kind of vehicle you own when you can’t afford a car. Your feet.</td>
</tr>
<tr>
<td></td>
<td>icescape</td>
<td>• Ice cube that ends up on the floor when you break a new tray of ice.</td>
</tr>
<tr>
<td></td>
<td>folex</td>
<td>• Fake expensive watch as in, faux (french for fake)Rolex, faux-lex.</td>
</tr>
<tr>
<td></td>
<td>lesriage</td>
<td>• When two women get ‘married’ it is a lesriage.</td>
</tr>
<tr>
<td></td>
<td>mesbian</td>
<td>• A man with feminine qualities who still likes women.</td>
</tr>
<tr>
<td></td>
<td>shiim</td>
<td>• A he-she, she-male, transvestite</td>
</tr>
<tr>
<td></td>
<td>askhole</td>
<td>• Someone who asks too many stupid, pointless, obnoxious questions.</td>
</tr>
<tr>
<td>Affixation</td>
<td>broughten</td>
<td>• Purposefully incorrect use of brought. Also: broughted</td>
</tr>
<tr>
<td>Shortening</td>
<td>keep it posi</td>
<td>• Short for “keep it positive.” to maintain a positive outlook.</td>
</tr>
<tr>
<td></td>
<td>val pal</td>
<td>• A valentines day partner, someone to exchange gifts with.</td>
</tr>
<tr>
<td></td>
<td>bis cas fri</td>
<td>• This is how we say business casual friday around the office, because, you know, we have to abbrev everything.</td>
</tr>
<tr>
<td>Initialisms</td>
<td>TMI</td>
<td>• Too much information</td>
</tr>
<tr>
<td></td>
<td>DTR</td>
<td>• Define the relationship</td>
</tr>
<tr>
<td></td>
<td>ADF</td>
<td>• On company time</td>
</tr>
<tr>
<td></td>
<td>WTF</td>
<td>• What the fuck</td>
</tr>
<tr>
<td></td>
<td>BTW</td>
<td>• By the way</td>
</tr>
<tr>
<td>Acronyms</td>
<td>UDWOD</td>
<td>• Urban Dictionary Word of the Day</td>
</tr>
<tr>
<td></td>
<td>POATEW</td>
<td>• Angry; Irate; woke up on the wrong side of the bed. From the acronym Pissed Off At The Entire World. Pronounced poe-AH-too.</td>
</tr>
<tr>
<td>Metaphor</td>
<td>circle the drain</td>
<td>• to gradually die (literally or figuratively)</td>
</tr>
<tr>
<td></td>
<td>trowler</td>
<td>• A woman wearing a huge amount of makeup</td>
</tr>
<tr>
<td></td>
<td>break the glass</td>
<td>• When everything you’ve tried has failed, and you’re resorting to your emergency plan. From breaking the glass of fire extinguishers to put out a fire.</td>
</tr>
<tr>
<td>Borrowing</td>
<td>seme</td>
<td>• Japanese for the partner on top</td>
</tr>
<tr>
<td>Allusion</td>
<td>Death Star</td>
<td>• A building which electronic signals (i.e. cellphone) cannot penetrate.</td>
</tr>
<tr>
<td></td>
<td>full court press</td>
<td>• Named after the play in basketball, it means to aggressively put the moves on, or to hit on someone.</td>
</tr>
<tr>
<td></td>
<td>dish jenga</td>
<td>• The pile of precariously balanced dishes in a dishrack that cannot be disturbed lest there be an avalanche of china, crockery, and silverware.</td>
</tr>
<tr>
<td>New Coinage</td>
<td>snarf</td>
<td>• to pilfer, to take something that one perceives as off-limits, especially as related to food or beverage consumption.</td>
</tr>
</tbody>
</table>
Speakers of a language often consult a dictionary to solve disputes of meaning, spelling, etymology, pronunciation, and sense. Some speakers refer to “the dictionary” as they do “the Bible” or “the Koran” or “the Upanishads” or other religious tomes. The high status given to “the dictionary” has been called lexicographolatry by Algeo, who points out that “[a]s the Bible is the sacred Book, so the dictionary has become the secular Book, the source of authority, the model of behavior, and the symbol of unity in language” (Algeo 1989: 29). Such reverence is afforded to UD as well. In fact, the sixth most popular UD definition for “urban dictionary” makes this association explicit in (1).²

(1) **urban dictionary**
The slang Bible
*Urban Dictionary rox my fuckin sox*

Dictionary editor Wilson Follett believed that lexicographers should know that their work is “received by millions as the Word on high” (1962: 77). Although there are many, varied dictionary types, “the hypostatization of an archetypal Book lying behind actual published volumes, coupled with the respect accorded the archetype, leads to a kind of unconscious lexicographolatry that has shaped public attitudes to dictionaries in large segments of the English-speaking world” (Algeo 1989: 29).

The model of dictionary authority in the UD Web context, however, is different. UD’s authority resides in the fact that it challenges traditional dictionaries. For example, whereas no entry exists in UD for *The Merriam-Webster Dictionary*, ten entries exist for the misspelled *The Marriam Webster Dictionary*. The entry with the earliest date is (2).

(2) **the marriam webster dictionary**
the dictionary sitting by me that doesn't have half of the words and definitions on this site. I don't think it should.

This definition’s author sees UD as the authority on everything that is *not* in traditional print dictionaries by virtue of the absence of certain material from traditional print dictionaries (i.e., UD content). Subsequent definitions referred to the misspelling in the headword, although no user decided to redefine *Merriam-Webster* with the correct spelling. Authority is determined by popularity on UD and the most popular definition for *The Marriam Webster Dictionary* appears in (3).

(3) **The Marriam Webster Dictionary**
An imaginary dictionary, created by the inventive minds who cannot spell Merriam-Webster.

Horace Walpole wrote that societies, and not individuals, should publish a standard dictionary and that the authority of a dictionary is undermined without consensus (Reddick 1990). Although UD begins with individual voices defining their world, its design aims to find consensus or at least utilize social popularity to determine what a word means. A UD entry receives praise through its increasing popularity (i.e., users click the Thumbs Up button) and even notoriety. Definitions with the greatest number of Thumbs Up usually have several of the following traits: humor (often via language play), wisdom, polysemy, and linguistic

---

² Each UD entry is listed by line in this order: headword, definition, and example sentence.
compence for the desired voice of the entry. In the same way that Scrabble players utilize dictionaries to solve disputes, users often visit UD to solve disputes related to slang or, in rare cases, more academic disagreements. One definition of “UrbanDictionary” noted that “[i]t is a quite convenient tactic (one that is often employed by debaters on the high school circuit) to put a definition that they want on the site so that they may quote it in round.”

7 Linguistic innovation and ‘communicative risk’

UD entries are often used an instrument of competition or intimidation, following both “real-life” social practices or more recent online ones. There are high stakes in creating new slang: “Linguistic innovation is always a communicative risk. Slangisms – as part of the lexicon in the making – have to take the risk of sounding misplaced and unsuitable” (Sornig 1981: 75). When UD entries sound “misplaced and unsuitable,” battles erupt, manifested whenever a UD user fights another for meaning-making rights. This typically occurs when User A writes a definition; User B refutes or amends User A’s definition with her own; and other users join in by ranking the existing ones (with Thumbs Up/Down) and posting their own definitions.

UD entries also employ the tactics of bullying and name-calling, known online as flaming. Aware of UD’s function as a place to flame others, users have created entries for this online speech act, e.g., “Urbandictionary namecaller” (4) and “nameflamer” (5).

(4) **Urbandictionary namecaller**
Someone who feels the need to define someones name and call them bad things.

(5) **nameflamer**
One who insults (flames) another by using their name as a derogatory term in urbandictionary.com. A nameflamer would probably hesitate to insult the flamed person to their face.

Most nameflamers on Urbandictionary.com seem to be teenagers.

Sometimes the bullying is not visible in the headword itself but is deployed within the definition as in (6).

(6) **scene**
T. B. depicted above, is a faggot. He should kill himself.

*Someone should really murder that kid.*

Within this entry is an allusion to another online feature of UD, Images, which allows users to post pictures associated with the headword. In (6), one user recognized the photograph of another and antagonized the pictured user through a commentary unrelated to the headword “scene.” The example also includes an indirect threat and is an example of online bullying, a noted practice among adolescents on discussion boards, IRC, and blogs (cf. Simmons 2003).

8 ‘Empowered amanuenses’

Web technologies have allowed ordinary users of dictionaries who are not trained lexicographers to engage in dictionary-making. Traditional dictionaries, no matter how
authoritative, “are the products of human beings, [who] try as they may, bring their prejudices and biases into the dictionaries they make” (Green 1996: 11). In the case of UD users, they bring these “prejudices and biases” into every stage of this would-be amateur lexicography: wordlist compilation, orthography, definition, exemplification, organization, editing, and updating. To find the items for the wordlist, they cull them from quotidian speech or invent them in moments of on-the-spot coinage. Their spellings are variant and inconsistent as are their conventions for punctuation (e.g., overused hyphens, apostrophes) and capitalization. The definitions they write follow no formatting guidelines, and their examples sometimes use the headword and sometimes do not. Definitions are added below headwords ad infinitum, and they remain or are expunged based on popular vote, concerted effort of Editors, or trouble-seeking trolls (self-appointed online sabateurs).

Many metalexicographers agree that the dictionary user must be an important consideration in dictionary-making (Hartmann 1987; Hartman and James 1998; Knowles 1990; McDavid 1973). With UD (and wiki-based dictionaries to a somewhat lesser degree), the user is the primary consideration because the user is the one doing the considering, i.e., the user is contributor is Author-Editor. With the advent of collaboratively authored “populist” dictionaries such as UD (cf. Damaso 2005), the attitudes and practices of users need not be abstracted into principles and trends. Instead, users themselves, in this case UD contributors, merely act and make meaning, and use the dictionary as they wish.

We might say, then, that the process of creating UD is analogous (although by no means equivalent) to that of Johnson’s Dictionary or the OED. In a similar way to how Dr. Johnson established a system that eight amanuenses put into practice through repetitive, learnable tasks – copying citations and pasting them into a manuscript (Reddick 1990) – UD’s moderator Aaron Peckham established the system behind an online dictionary (e.g., creating an easy-to-use Add a Word feature) that thousands of users put into practice on a daily basis. The primary difference, however, is the spontaneity and individuality of UD users who actuate the moderator’s system. Whereas Johnson’s amanuenses, or word-copiers, would not stray from their tasks of mimesis, UD’s empowered amanuenses both abide the norms of the site one day and refute them the next. Indeed in the wider view taken by James Murray in The Evolution of English Lexicography, “the English Dictionary, like the English Constitution, is the creation of no one man, and of no one age; it is a growth that has slowly developed itself down the ages” (Murray 1900: 6-7).

Internet technologies that easily permit exchanges between a user and a database have allowed a new type of dictionary online, one that is built by the collaboration of contributing end-users. Since contemporary usage (“slang”) is often associated with the general populace – the end-users – an online slang dictionary can make meaning, codify meaning, and provide a rare auto-symbiosis between language user and lexicographer.

References


