

## Oxford Language Researchers All A-Twitter

**I wish the coffee machine at work was capable of dispensing an entire pot instead of 1 cup.**

**At the airport. Flight doesnt leave for another 3 hours. I CANT WAIT TO GO BACK TO MY DORM OMG**

What is President Obama using the Obamaberry for? Secrets leaked in LA Times. <http://ping.fm/fBaTz>

Here's something to get the twitterati tweeting: dictionary experts at Oxford University Press (OUP) have been monitoring 1.5 million tweets this year to explore the impact of the Twitter phenomenon on the English language.

Using the cutting-edge technology used for gathering data in OUP's two-billion-word Oxford English Corpus, lexicographers are able to track changes in the language across a wide variety of media including newspapers, magazines, blogs - and now Twitter.

"We've been adding tweets to our Corpus research since January this year, reflecting the fact that Twitter has become such a popular and dynamic form of communication," says Judy Pearsall, Senior Publishing Manager for English Dictionaries at OUP. "To really analyze the English language, compilers of current dictionaries need to incorporate all forms of communication into their research."

So, what does Twitter show us about how people are using the language? "Obviously, with just 140 characters in which to get your message across, we're seeing some very creative uses of language," says Pearsall.

Here are just some of the findings to date:

- **New words** are being coined by Twitters all the time. One obvious recent example is the proliferation of what we've come to call 'Obamologisms' - terms associated with Barack Obama - so we have expressions such as *Obamania*, *Obamaniacrazy*, *Obamacracy*, *Obamanomics*, *Obamafication*, *Obamalicious* and even *Obamaberry*.
- **Abbreviations** feature highly in our Top 500 list of frequent words. People are using "=" instead of "is", "+" (and), "u" (you), and "RT" (retweet), which incidentally is the third mostly likely way people are starting their tweets. Twitterers make lighter use of connector words, such as "the", "and" and "to", possibly because they don't want to use up precious characters on them, use abbreviated forms, instead - or just aren't bothered about writing these words in.
- **Gerunds** ("-ing" words e.g. "watching", "listening" etc) are popular, showing just how often people use Twitter to report on whatever they are experiencing or consuming at the time.
- Another fact that comes out of the research into the top 500 words used by Twitterers is

how often **adjectives** such as “cool” and “awesome” appear. People are also using exclamations e.g. “wow”, and “yay” a lot more than in other forms of communication.

- **“Tweet”** and **“twitter”** themselves also feature in the top 500 words in Twitter messages.

“We’ll be continuing to monitor Twitter, as well as any other significant forms of new communication methods, alongside our established programme of data-collecting into the future.” says Pearsall. “The ability to track language and reflect real language use – by real speakers - ensures that Oxford dictionaries remain at the forefront of dictionary publishing.”

Finally, will all this activity mean we're likely to see tweets appearing in Oxford dictionaries any time soon?

**Yay! = cool idea, working on it**

**For more information please contact Juliet Evans at Oxford University Press on 01865 353911, email [juliet.evans@oup.com](mailto:juliet.evans@oup.com) or via Twitter at: [www.twitter.com/jueEjue](http://www.twitter.com/jueEjue)**

#### **TWITTER FACTS FROM THE OXFORD ENGLISH CORPUS**

The Oxford English Corpus contains almost 1.5 million tweets, randomly selected from all public tweets between January and April 2009

#### **BASIC NUMBERS**

Total tweets = 1,496,981

Total sentences = 2,098,630

Total words = 22,431,033

Average words per tweet = 14.98

Average sentences per tweet = 1.40

Average words per sentence = 10.69

#### **QUESTION OR STATEMENT?**

Almost exactly 10% of Tweets contain a question.

#### **MOST FREQUENT WORDS**

Some of the items distinguishing the top words in Twitter from those in general English are:

Several very popular web addresses feature among the Twitter top 500: "tinyurl.com", "twitpic.com", "ff.im", "is.gd", "twurl.nl". These all appear because they offer services useful to twitterers. There is also a higher profile of computer-related terms such as "Google", "Facebook", "internet", "website", "blog", "Mac" and "app".

Both "Twitter" and "tweet" appear in the Top 500.

#### **MOST FREQUENT FIRST WORDS**

"I" is the top-ranking word that tweets begin with, showing that most people are

'twittering' about themselves.

The abbreviation "RT" (retweet) is extremely common, in third place.

Stephen Fry's tweet stream is so popular that "@stephenfry" is in 81st place, due to people using this label to indicate that they are replying to Fry's last tweet.

"Watching", "trying", "listening", "reading" and "eating" are all in the top 100 first words, revealing just how often people use Twitter to report on whatever they are experiencing at the time.

### **SHORT AND TWEET**

Various abbreviations appear in the Twitter top 500, though these are far less common in other forms of English: "=" (is), "RT" (retweet), "lol" (laughing out loud), "OMG" (Oh my God!), "u" (you), "+", "n" (and), "re" (regarding), "x" (kiss), "b" (be), "ur" (you're/your).

"Im" is often used without its apostrophe.

---

### **Notes to editors:**

#### **Powered by Oxford Corpus**

Our dictionary entries are powered by the Oxford English Corpus, part of the largest language research project in the world. Containing more than two billion words, collected daily from sources ranging from novels to newspapers to chat rooms and blogs, the Oxford English Corpus means that Oxford dictionaries empower you with the most authoritative and up-to-date information on the English language. For more information, please visit [www.askoxford.com/oec](http://www.askoxford.com/oec)