New Media & Society

Ambient affiliation: A linguistic perspective on Twitter

Michele Zappavigna

The online version of this article can be found at:

Published by:

$SAGE

Additional services and information for *New Media & Society* can be found at:

Email Alerts: http://nms.sagepub.com/cgi/alerts

Subscriptions: http://nms.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://nms.sagepub.com/content/13/5/788.refs.html

>> Version of Record - Jul 29, 2011

OnlineFirst Version of Record - May 31, 2011

What is This?

Ambient affiliation: A linguistic perspective on Twitter

Michele Zappavigna
University of Sydney, Australia

Abstract
This article explores how language is used to build community with the microblogging service, Twitter (www.twitter.com). Systemic Functional Linguistic (SFL), a theory of language use in its social context, is employed to analyse the structure and meaning of 'tweets' (posts to Twitter) in a corpus of 45,000 tweets collected in the 24 hours after the announcement of Barak Obama's victory in the 2008 US presidential elections. This analysis examines the evaluative language used to affiliate in tweets. The article shows how a typographic convention, the hashtag, has extended its meaning potential to operate as a linguistic marker referencing the target of evaluation in a tweet (e.g. #Obama). This both renders the language searchable and is used to upscale the call to affiliate with values expressed in the tweet. We are currently witnessing a cultural shift in electronic discourse from online conversation to such 'searchable talk'.

Orientation: language and social networking sites (SNS) such as Twitter

Electronic conversation lends itself to search in a way that face-to-face conversation, currently, does not. Search engine usage has become a ubiquitous process for locating information. It is, however, only as we have begun to leave more traces of social

Corresponding author:
Michele Zappavigna, Department of Linguistics, F12, University of Sydney, NSW 2006, Australia
Email: michele.zappavigna@sydney.edu.au

interaction online, via microblogging services such as Twitter, that a cultural shift toward a more interpersonal function for search has emerged. This is a shift from searching purely for content, to searching what other people are saying online and forming communities of shared value. In popular terms, it is becoming increasingly useful to search the 'hive mind': the stream of online conversation occurring across semiotic modes (e.g. blogs, online chat and social networking sites). For example, the kind of discourse-search that Twitter affords has been described as a rival to Google, with commentators claiming that searching Twitter may soon be one of the most effective ways to gather useful information, since returns capture what users are saying online in real-time (Rocketboom, 2009).

This cultural shift to interpersonal search has resulted in the emergence of *searchable talk*, that is, online discourse where the primary function appears to be affiliation via 'findability'. This kind of talk expands linguistic meaning potential by using punctuation to incorporate metadata into language so that online talk can be found. Taking a corpus of posts to Twitter as a case study, this article aims to suggest how search is beginning to function as a community-building linguistic activity. It will demonstrate how what are known as 'hashtags' function as linguistic markers enacting the following social relation: 'Search for me and affiliate with my value!'.

By enabling users to affiliate online, social networking sites (SNS),[1] accessed by millions worldwide, afford a new form of sociality in which language maintains a pivotal role. While studies of online discourse from a linguistic perspective are relatively established (Baron, 2008; Crystal, 2006; Herring, 1996), whether analysis of linguistic function and structure can serve as evidence for defining communities is an emergent area of inquiry. This is not to say that there has been little work on language in online communities, but rather that there is yet to be an accumulation of research providing linguistic models of online, and indeed offline, affiliation. The notion of online community was popularized in Rheingold's (1993) work on 'virtual community': 'virtual communities are social aggregations that emerge from the Net when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace' (Rheingold, 1993: 5).

Since the emergence of this definition, often criticized for its vagueness, there has been a debate surrounding the criteria for establishing the bounds of online communities, the structure of community and how communities are built or emerge (see for example Hagel and Armstrong, 1997; Jones, 1997; Burnett, 2000; Wellman, 2001; Herring, 2004, 2008). No stable definition of community has prevailed. However, a linguistic perspective on virtual community might aim to explicitly describe how people use language to construe social bonds and how they rally around, defer or reject different values construed in language (Knight, 2008). This type of work is emerging as linguists begin to expand their territory into different modes of communication, such as image, gesture and music, viewing these modes either as forms of semiosis that are 'parasitic' on language (Halliday and Matthiessen, 1999) or as themselves having a grammar that can be analysed (Kress and Van Leeuwen, 2006).

This article begins by introducing Twitter and a corpus of tweets (messages posted to Twitter) containing the keyword 'Obama' collected in the 24 hours after the announcement that Barak Obama had won the 2008 US presidential election. It then introduces Systemic Functional Linguistics (SFL), the theory of language used in the study. SFL is

a social semiotic theory that investigates discourse in context. The Appraisal framework (Martin and White, 2005), a model of evaluative language using the SFL approach, is then detailed. The aim of the appraisal analysis undertaken in this article is to show the kinds of interpersonal meanings made in the corpus of tweets and their role in ambient affiliation.

Constrained meaning-making on Twitter

Twitter, developed in 2006, is an example of a microblogging service. These services allow users to post character-constrained messages via a range of technologies such as mobile phone, instant messaging clients and the web. Tweets, messages posted to Twitter, are messages presented to a virtual audience who 'follow' by subscribing to another user's feed and, as the service has evolved, who search. Tweets unfold in time as a 'twitter stream' that is presented in reverse chronological order on a user's Twitter page and also, unless privacy selections are adjusted by the user, as part of the public feed. A tweet, in reference to its original conception as an SMS-based message, is constrained to no more than 140 characters in length and may incorporate links to micromedia, small-scale multimedia, and shortened aliases of longer hyperlinks (Tiny URLs). As such, they are interesting cases in making meaning within constrained environments.

   In contrast to other forms of communication, there is no communal expectation that anyone respond to a tweet, as the metaphor of 'twittering' continuously like a bird implies. There is, however, a social need among users to engage with other voices in public and private feeds. Hence we see creative use of punctuation to reference other users and tag common topics. These expansions in typographic meaning potential are part of a community-driven movement toward Twitter becoming a form of 'public conversation'. It is conversation, however, that is multiparty, temporarily fluid and highly intertextual.

   As Twitter has evolved, so have the resources for attributing and addressing other users. Linguistic markers have begun to populate tweets to facilitate heteroglossia (Bakhtin, 1981), in other words, to bring other voices into tweets by addressing other users, republishing other tweets, and flagging topics that may be adopted by multiple users. The first of these conventions is the @ character, used as a deictic marker as in the following example:

   '@username I didn't vote Obama. But I wasn't exactly for McCain either. I hate both, to be frank.'

   The @ character indicates that the username which follows it is addressed in the tweet and the structure functions like a vocative, that is, as a form of address. The @ character does not have to operate in this initial position in a clause but can also occupy a medial or final position. In these instances it is more likely to mark a user, flagging that they are being referred to, but not explicitly inscribing an address. For example, in the following tweet, the user is not directly addressing 'username1' although the use of the @ character means that this user, and users who follow him, are likely to see the message in their feed:

   'I'm joining @username1 in his four commitments to Pres.-elect Obama. Will you? http://tinyurl.com/579akg.'

Due to its electronic affordances, this form of punctuation renders the deictic marker aggregatable and searchable. So, for example, a user may track all @ references to themselves or to a user that they are following. Another way of bringing external voices into a tweet is to republish part or all of another tweet intact or modified by 'retweeting', using the character combination, RT. In most instances the RT will be followed by the @ character to 'source' the retweet:

'RT @username2 Absolutely bril http://tinyurl.com/2upsz4 Congratulations Obama!'

In this way the RT functions as a form of engagement realized by grammatical projection, in other words, it functions to indicate that the clause following '@username2' is a quotation, most often a direct quotation of that user's talk.

Hashtags on Twitter

The 'hashtag'[2] (#) has a different function to RT and @, functioning instead to mark the topic of a tweet. However, it is also broadly involved in construing heteroglossia in the sense that it presupposes a virtual community of interested listeners who are actively following this keyword or who may use it as a search term. This character usage derives from Internet Relay Chat (IRC) conventions for naming channels. Hashtags, as they are used on Twitter, are a form of 'inline' metadata, that is, 'data about data' that is actually integrated into the linguistic structure of the tweets.

Within a tweet, a tag marked by a # sets up an attributive relationship between the tweet as a tagged token and the tag as its type. Halliday and Matthiessen (2004: 219) define attributive relational processes as relationships where 'an entity has some class ascribed or attributed to it'. In other words, hashtags inscribe a keyword in a tweet as metadata referencing the topic of the message as assigned by the user. For example, the following tweet contains the tag '#obama' indicating the subject matter of the tweet:

'I just typed the words 'Obama presidency.' It felt good. **#obama**.'

The 'tag as type' relationship assumes that other users will also adopt this tag and use it as a keyword for a tweet on this topic. By generating keywords describing their discourse in this way, Twitter users enter into the social realm of collaborative tagging, or 'folksonomy' (Vander Wal, 2007). Collaborative tagging is a social form of verbal indexing involving a 'bottom-up' approach to the kind of classification previously achieved by reference librarians. It is, for example, used heavily on photosharing sites such as Flickr.
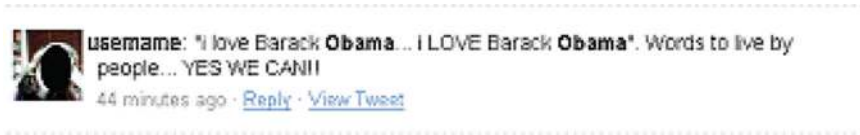
This type of inline metadata is different to how metadata is usually used because it is directly visible to the user as part of the text of the tweet. While metadata rendered in markup languages such as extensible markup language (XML) will typically separate form from content, social tagging on systems such a Twitter collapses this separation. Thus tweets of the following kind, where metadata is marked by a hashtag inside a clause, are possible:

'Organized the pics from last night's Obama rally **#Obama**.'

The hashtags can also mark functional roles in the linguistic structure. The tweets in Table 1 are examples of this potential where Classifiers, Things and Processes are

Table 1. Examples of the hashtags marking units with different grammatical functions in clauses

ID	Tweet	Function marked
1	Millions of pathetic **#obama** supporters woke this morning with nothing to talk about: http://twurl.nl/yzxt4t Funny stuff (h/t @User)	Classifier
2	my daughter is explaining to me why she finds **#obama** cool. it sounds something like: tatadatadatitadata. Can't argue.	Thing
3	Alright, let's all sign up. New meme on Election.twitter.com. Time for Obama. Volunteer to **#helpbushpack**	Process

username: "I love Barack **Obama**... I LOVE Barack **Obama**". Words to live by people... YES WE CAN!!
44 minutes ago · Reply · View Tweet

Figure 1. An example of a tweet in the Obama Win Twitter corpus

marked. These are linguistic terms from SFL used to describe the functions of a particular linguistic unit in a clause. They encode both the structure and the meaning of a particular unit. For example, in the following tweet the Process, typically realized by a verb, (woke) construes an action in the world, the Thing (supporter) encodes the agent that carries out the action, while the Classifier (obama) describes the type of social categories to which the Thing belongs: 'Millions of pathetic #obama (*Classfier*) supporters (*thing*) woke (*process*) this morning …'

Tweets and the hashtags which they contain may thus be thought of a two different orders of experience: a tweet is an instance of language use, while a tag is language about language, performing what this article will show is an affiliative function.

Data: the Obama Win Twitter corpus

The dataset analysed in this study is a corpus of tweets collected using a Python script and the Twitter API to scrape all tweets containing the string 'Obama' posted to Twitter in the 24 hours after the declaration of Barak Obama's victory in the 2008 presidential elections. The corpus contained 45,290 tweets (813,310 words). The aim was not to construct a representative corpus of the linguistic activity on Twitter, but instead to conduct a case study in which field variables, that is, the topic of the tweets, was held relatively constant to afford a rich investigation of meaning-making in a single domain on Twitter.

Figure 1 is an example of a tweet from the corpus. It is presented as it appears on the web interface to Twitter. The visual appearance would, however, be different depending on the channel via which the tweet is syndicated and displayed. In this way, a tweet is a text with multiple expression plane realizations or, in other words, with no single stable visual or typographic form. Thus tweets are likely to be particularly capricious objects for multimodal discourse analysis which considers how meanings are made with both language and paralanguage.

Table 2. The 10 most frequent # tags in the Obama Win Twitter corpus

ID	# tag
1	#Obama
2	#ElectionWrap
3	#TwitVote
4	#election08
5	#web2summit
6	#3News
7	#election
8	#haiku's
9	#puppy
10	#moc2008

The corpus contained 721 instances of re-tweeting and 7100 instances of @ characters used for attribution and addressivity. There were also 770 tweets employing the hashtag to refer to the 'topic' that the user wished to ascribe to the tweet. Eighty-five unique hashtags were used. The twenty most common hashtags in the corpus are presented in Table 2, with the most common tag being #Obama. Two hundred and thirty-four tweets in the corpus contained this tag. An example is the following: 'Its still surreal #Obama.'

Theoretical framework: Systemic Functional Linguistics and a language-based theory of affiliation

Systemic Functional Linguistics (SFL) is a theory of language that is tailored to answering questions about how meanings function within the particular contexts in which they are made. SFL is aligned with the functional tradition in linguistics manifest in the Prague School (Jakobson, 1971) and arose out of the linguistic school known as Firthian Systemics (Firth, 1957). It is a functional theory because it is targeted at answering questions about how meanings operate within the particular contexts in which they are created. As a method for managing the high dimensionality of language, SFL stratifies language into phonology (systems of sounds/writing), lexicogrammar (systems of wording), discourse semantics (systems of meaning), and context (genre and register). The strata are related to each other in terms of emergent complexity: as patterns of patterns (Lemke, 1984). For example, discourse semantics is a higher order patterning of lexicogrammatical patterns, which are in turn patterns of phonological patterns.

A key functional unit in SFL is the clause. A clause corresponds to the notion of a simple sentence. For example, 'I love Obama' is a clause. A complex sentence can contain multiple clauses. For example, 'I love Obama and I voted for him' contains the additional clause 'and I voted for him'. This is what is referred to as a clause complex because it contains two clauses that enter into a logical relationship. Halliday
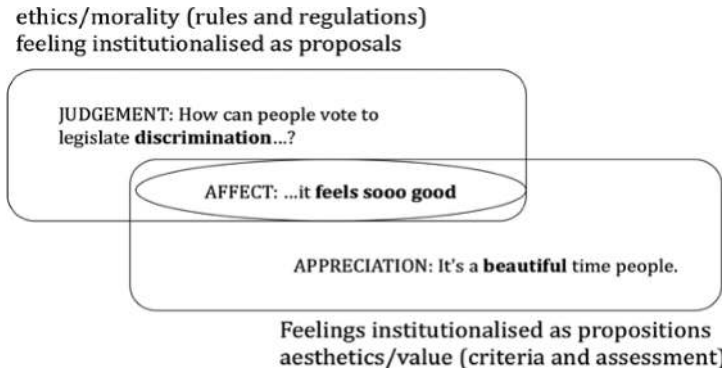
ethics/morality (rules and regulations)
feeling institutionalised as proposals

JUDGEMENT: How can people vote to
legislate **discrimination**...?

AFFECT: ...it **feels sooo good**

APPRECIATION: It's a **beautiful** time people.

Feelings institutionalised as propositions
aesthetics/value (criteria and assessment]

Figure 2. The institutionalization of affect (adapted from Martin and White, 2005: 45)

(1994: 106) proposes that the clause is central in construing experience, claiming that 'it embodies a general principle for modelling experience – namely, the principle that reality is made up of PROCESSES'. A process might be thought of as the heart of the clause. In lay terms, a process is a verb, that is, something that happens such as 'eating', 'thinking', 'saying', and 'playing', or something that 'is' such as 'being' or 'existing'.

According to the theory, language enacts three simultaneous functions, referred to as metafunctions: an *ideational* function of enacting experience, an *interpersonal* function of negotiating relationships, and a *textual* function of organizing information (Halliday and Matthiessen, 2004). A linguist using this theory will attempt to consider these three functions when analysing any instance of linguistic meaning.

This article focuses on relationships set up between ideational and interpersonal meanings in a corpus of tweets. Evaluation is a domain of interpersonal meaning where language is used to build power and solidarity by adopting stances and referring to other texts. In order to analyse evaluative meanings it draws on a theory of appraisal developed within the SFL paradigm (Martin and White, 2005). This theory considers how the linguistic patterning of a text construes emotional language in three areas: ATTITUDE[3] (making evaluations), ENGAGEMENT (bringing other voices into the text) and GRADUATION (scaling up or down evaluations). Figure 2 provides examples of each of these kinds of evaluation (shown in bold). This figure suggests, following Martin and White (2005: 45), that the protolinguistic expression of personal reactions via AFFECT that we may see in an infant, develops as we are socialized into a culture and into cultural institutions. These feelings become institutionalized as ethics or morality, forming the JUDGEMENT system, and as aesthetics or value, forming the APPRECIATION system.

For example, the evaluative lexical item, 'beautiful' in Figure 2 is an example of APPRECIATION because it makes a value-based assessment. This system is a resource for expressing attitudes about objects, states and processes, in contrast to the JUDGEMENT system, which typically assesses human behaviour and what people say and believe.

Recent work in SFL has considered the 'coupling' (Martin, 2000; Zhao, 2010) of evaluation with other kinds of linguistic meanings in texts as a way of tracking the kinds of values construed in the process of affiliation; that is, the process by which people involve themselves in social bonds (Knight, 2008). A coupling is a binding of two meanings across paradigmatic systems of potential and may be involved in larger syndromes of meaning in a text (Zappavigna et al., 2008). It is related to the simpler notion of collocation, two linguistic items occurring near each other in text. Knight (2008) suggests that couplings of interpersonal and ideational meanings have an affiliative function. She develops this idea in the context of conversational humour, which she argues is characterized by the tendency of participants to 'laugh off' couplings that cause 'wrinkles' in their network of values:

> 'to construe affiliation, conversational participants present couplings that represent what ties them together as members of particular communities, and variously commune around, reject, or laugh off these couplings as more or less "acceptable" bonds between them' (Knight, forthcoming).

Culture, when adopting this perspective, becomes a semiotic network of social bonds. We may investigate these bonds by considering how meanings unfold in texts as they are instantiated in particular contexts of use, for example, users twittering about Obama's election win.

Method

The Obama Win Twitter corpus was analysed using the metafunctional approach of Systemic Functional Linguistics, described in the previous section. In terms of close text analysis, this involved coding the corpus for evaluative language using the system of appraisal introduced in the previous section. The schema underlying this analysis was a system network for evaluative language. Systems networks conform to the SFL modelling strategy that conceives of language as a system of meaning potential, where that potential is realized as the particular configurations of linguistic choices which can be identified in texts. This is a perspective that considers language paradigmatically in terms of 'what could go instead of what' (Halliday and Matthiessen, 2004: 22) rather than modelling language as a catalogue of structures.

Figure 3 is the system network based on Martin and White (2005) that acted as the schema for annotating the evaluative language in tweets. In this network, a square bracket represents a choice between two options (an 'or' relation), while a brace represents simultaneous choices (an 'and' relation). For example, the ATTITUDE system involves three simultaneous systems: TYPE, EXPLICITNESS and POLARITY. Within the type system, there are three possible choices: AFFECT, JUDGEMENT or APPRECIATION. The extracts from tweets in the boxes are example realizations of the features in the network.

The software application, UAM Corpus Tool (O'Donnell, 2008), was used to annotate the data and AntConc (Anthony, 2009) was used for concordancing. The analysis presented in the section that follows draws upon concordance lines, n-grams and close text analysis. Since it was beyond the scope of manual analysis to analyse the entire corpus of tweets using the close text analysis involved in appraisal analysis, a random sample of 100 tweets was annotated for appraisal.
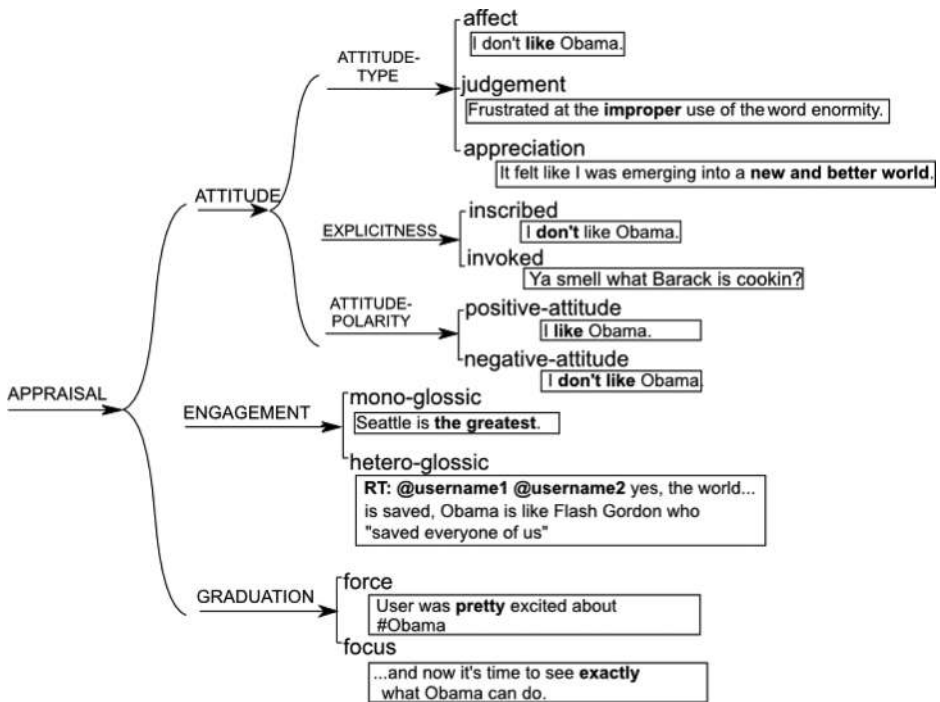
Figure 3. System network used a schema for appraisal analysis

Twitter and interpersonal meaning

Despite being parodied as a service that allows users to give status updates about their activities with irritating frequency, Twitter, as some commentators have noted, may be performing a more interpersonal function: 'The best Tweets tend to make an observation, take a stance, or crack a joke–none of which fall under the umbrella of using Twitter to tell the world what you're doing' (McCracken, 2009).

This intuition is supported by the findings of Mischaud's (2007) study of the extent to which Twitter users respond to Twitter's question, 'What are you doing?'. This content-based study found that 65 percent of tweets do not directly reply to this question by describing activities, although it should be noted that Mischaud's criteria that such tweets do not use 'ing' verbs is not the most rigorous linguistic basis upon which to make this kind of assertion. Instead topics such a family and friends were the focus of many tweets. The present study seeks to complement this initial content-based analysis with close linguistic analysis using the metafunctional lens offered by Systemic Functional Linguistics. This section will explore how tweets in the Obama Win corpus deploy resources for making interpersonal meaning.

Consider for example the following tweet in the corpus:

'HOLY CRAP. OBAMA WON HE WON!!!! IM SO HAPPY!!!'.

In this example, the expletive in initial position, the modifier 'so' and the choice of large-caps and repeated exclamation marks realizes INCREASED GRADUATION of the interpersonal meanings expressed, as annotated below. Obama is coupled with POSITIVE JUDGEMENT about winning the election and in turn with POSITIVE AFFECT (appraisal coding in square brackets, appraisal item in bold):

HOLY CRAP [increased graduation: force]. OBAMA **WON** [positive judgement] HE **WON** [positive judgement]**!!!!** [increased graduation: force] IM **SO** [increased graduation: force] HAPPY [positive affect]**!!!** [increased graduation: force]

Other evaluatively-charged tweets include the following (appraisal coding in square brackets, appraisal item in bold):

@username It's not **crying** [POSITIVE AFFECT] over Obama so much as it's **crying** [POSITIVE AFFECT] for what that means. We moved closer to being **what we aspire to be** [POSITIVE JUDGEMENT]. #obama

I **don't like** [NEGATIVE AFFECT] Obama and I am **not happy** [NEGATIVE AFFECT] about his election. **Almost half** [INCREASED FORCE] of the US agrees, but I guess they're not on twitter.

while we breathe we **hope** [POSITIVE AFFECT]**!!!!!** [INCREASED GRADUATION] **love love love** [POSITIVE AFFECT] [INCREASED FORCE] obama... he brought **hope** [POSITIVE AFFECT]!

#electionwrap NGO report: McCain's coverage was **much more** [INCREASED GRADUATION] **negative** [NEGATIVE JUDGEMENT] than Obama's

The process 'won' was the most frequent evaluative item in the corpus, occurring 2,764 times. 'Won' is an instance of POSITIVE JUDGEMENT, positively appraising the outcome of

Really <u>happy</u> that Obama	**won**	by such a margin! But wtf happend to prop 8 :/
Obama	**won!**	Yay! Lol. And I wanted to tell you something *funny* but I forgot.
I still can't believe it! Obama	**won!**	I knew it! -he's a modern day **heroe**, I think and I <u>hope</u>. Now we'll just have to wait and see..
it's pretty *great* that obama	**won,**	i <u>hope</u> he can deliver on his **promises**. now someone **decent** needs to replace gordon brown.
<u>happy</u> and <u>relieved</u> that obama	**won**	the election
i woke up today and Obama had still	**won.**	This is *awesome*
Still <u>stunned</u> Obama	**won.**	It's like *Christmas*.
@username <u>cheer up</u>... OBAMA	**WON!**	:)
Still watching msnbc so <u>happily</u> Obama	**won.**	Wish I'd be in the States to witness this *historical* day.
Thank god Obama	**won!**	I don't think the world would have lasted with another oil-hungry, war-loving, country-invading Republican

Figure 4. Concordance lines for 'won' in the Obama Win Twitter corpus

the election process. Examples are shown in the concordance lines presented in Figure 4. Instances of POSITIVE AFFECT (underlined), POSITIVE JUDGEMENT (bold) and POSITIVE APPRECIATION (bold italics) that are co-realized are also shown. While these concordance lines are only a random sample of the evaluation surrounding 'won', they show a general prosody of positive evaluation co-occurring with this item. Negative appraisal was possible but was generally used in a contrastive relation to Obama's victory.

Some examples of related evaluative 3-grams[4] include:

Glad Obama won

Obama has won

won the presidency

Given that the news of Obama's victory was widely known, at least after the first hours post-announcement, these tweets seems to be performing a function beyond informing other 'twits' of the news. The kind of evaluative language that sampled above suggests that the tweets may be forming a more interpersonal social function in which users are affiliating around values relating to the election result.

In the corpus these values are often construed as couplings of 'Obama' with appraisal. For example, a common evaluative 4-gram that coupled POSITIVE AFFECT and POSITIVE JUDGEMENT was:

happy that Obama won

This included tweets such as those shown in Figure 5.

Clearly, an intervening variable in the argument being made here about the evaluative charge and interpersonal-focus of tweets is the historical and political nature of the event to which these tweets are responding. However, the tendency manifest in this corpus

is really, really, really	**happy that Obama won.**	
is at work til 5 and is	**happy that obama won**	the election....GO OBAMA!!!!
I'm happy and sad.	**Happy that Obama won,**	but sad that Prop 8 passed.
is it weird to feel	**happy that obama won**	the presidency campaign? obama changes america, america changes the world! hurrah!
I've talked with a number of people today who supported McCain but are still pretty	**happy that Obama won.**	
I'm	**happy that Obama won!**	it's very emotional to witness history firsthand
Very	**happy that Obama won!**	

Figure 5. Concordance lines for 'happy that Obama won' in the Obama Win Twitter corpus

toward presenting evaluative couplings, while perhaps more pronounced due to the political context, seems characteristic of the kind of interpersonal patterning seen throughout Twitter discourse on a range of topics. However, a large-scale quantitative study, beyond the scope of this article, would be required to verify this claim.

Hashtags: labelling the ideation we are going to axiologize around

The affiliative function of the evaluative language seen in the tweets will now be considered by investigating the role of hashtags in coupling. Hashtagging is an emergent activity. Hashtags are a typographic convention used to mark the topic of a tweet. We may also think of them as indicating the target of the appraisal in the tweet. For example, tweets containing couplings of 'Obama' and positive evaluation sometimes included the tag '#Obama':

> 'From fear, hatred and economic collapse to hope, search for common ground and prosperity again. Good change! #obama.'

In these instances the hashtag seems to intensify a 'call' to affiliate with the values in the tweet by rendering the tweet more 'searchable'. It is more searchable because a user can find this tweet by searching for 'Obama' even when the body of the tweet does not contain the string 'Obama'.

Consider the following tweet that couples positive evaluation and Obama without a hashtag:

> 'Life is Good! And a Good Morning to all in Obama country!'

Similar tweets were possible using a hashtag: 'username: 01:42:23 CET bedtime. go'nite #obama country, land of hope & change!'

The hashtag in this example is an instance of 'inline' metadata as the marker operates within the clause attached to a classifier in the nominal group. This is a classifier as it is defining the 'country' as an 'obama' type of country. The presence of the hashtag expands the meaning potential of the tweet when compared with the non-hashtag example by making the coupling 'louder' in the sense that it is more available for search and more likely to be automatically followed by those subscribing to this tag.

While Twitter has developed a search facility whereby a user can search via a range of parameters, this does not override current hashtag practice. Twitter has in fact incorporated a hashtag field into its search form. The utility of the hashtag can be seen when considering searches, for example, for 'election 08' which will only return results containing this exact string. A tag such as #Election08 becomes very useful for quickly retrieving a range of tweets about a topic that do not necessarily reference the topic as a word in the body of the tweet. In addition, it inscribes a form of 'intentionality' in the tweet via the token–type relationship described earlier, because the user has explicitly indicated this tweet is a token of the category labelled by the tag. Hashtags allow users to search for potential 'targets' of appraisal and find out, for example, whether Obama is subject to prosodies of positive or negative evaluation as tweets unfold in time.

Interpersonal as well as ideational meanings could occur in a hashtag. For example in the following tweet POSITIVE AFFECT (*hope*) is used as a tag:

'Prediction: Obama as powerful leader starts within month, big speech, call for econmic legislation #hope #change.'

Thus we have a classificatory system that is very different to the principles that would be employed to categorize a reference collection in library using subject keywords. From the perspective of information science, this may be perceived as a degradation of classification. However, such a criticism ignores the social function of the classification which is to not only facilitate efficient relevance and recall, but to make possible what may be termed 'ambient affiliation', a concept discussed later in this article.

The inline nature of #tag usage opens up the possibility of play with users creating tags that are unlikely to be used as search terms and which instead seem to function to intensify the evaluation made in the tweet. Instances include:

#racialjokeswecanmakenow

#presidentsIhavetheslightestshredofrespectfor

#finallyicansleepatnight

#americastillhasabrainandaconscience

This play is possible because of the close relationship of hashtagging to evaluation. For example, the final two instances above are part of the following tweet:

'Yes! #f*ckin-A #obama #finallyicansleepatnight #byebyebush #americastillhasabrainand aconscience #whew!'

Here the repetition of hashtags functions to scale up the positive evaluation of Obama's victory via the system of GRADUATION. Without the hashtags this tweet would lose some of its humourous hyperbole.

Ambient affiliation

Markup is typically hidden by browsers and other display devices; however, hashtags on Twitter are visible within clauses in tweets and no technology is used to obscure them. In this way users can choose to mean in an explicitly searchable manner by integrating metadata into their talk through typographic conventions, such as the hashtag, that increase the 'loudness' of their discourse by increasing the likelihood that their words will be found. This, in turn, increases the probability that a user's production of texts over time will be actively 'followed' by others. In other words, it creates the possibility of *ambient affiliation.* Here we affiliate with a copresent (Goffman, 1963), impermanent, community by bonding around evolving topics of interest. This function is directly inscribed in the web interface to users' Twitter accounts as 'trending topics', a list of
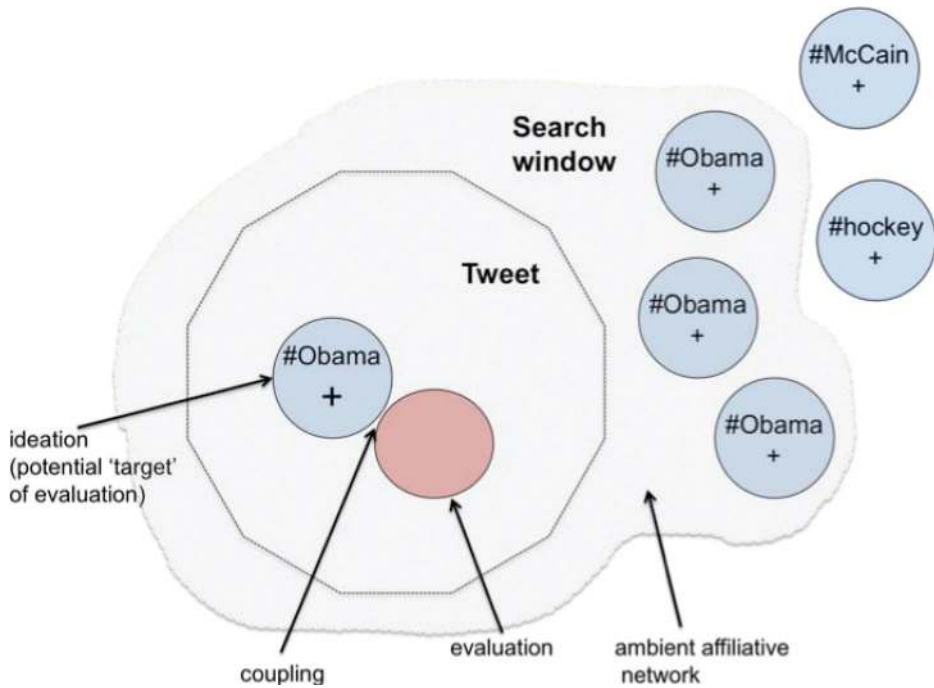
Figure 6. Search creating an ambient affiliative network for the '#Obama' tag

keywords occurring with high frequency in current posts, displayed under the search box. In increasing 'loudness' in this way, hashtags identify meanings that have become 'hyper-charged' with an additional semiotic pull that may be likened to a gravitational field. They act as both a label for the potential discourse community that they establish and render searchable the coupling that occurs in the tweet.

The expansion of meaning potential seen in the hashtag usage is both a product of the reduced affordances of the character-constrained mode and part of a 'multiplication' (Lemke, 1998) of what it means to talk online. It is electronic discourse explicitly encoded as searchable. Using the '#obama' tag is, for example, analogous to saying 'If you are interested in values about Obama search for me'. The social function of the hashtag is to provide an easy means of grouping tweets, and in turn, creating ad hoc social groups or sub-communities. Being searchable opens up a new kind of sociality where microbloggers engage in ambient affiliation. The affiliation is ambient in the sense that the users may not have interacted directly and likely do not know each other, and may not interact again. It also could not occur without adequate search functionality. Users searching to explore online conversations produced on social networking sites in this way is a new cultural process.

Interpersonally-charged tweets invite with their hashtags an ambient audience to align with their bonds. The 'hypercharge' of the hashtag involves the tweet in a larger bond network of values. For example, we might think of the search window for the tag '#Obama' as creating a momentary, ambient affiliative network of tweets in which this
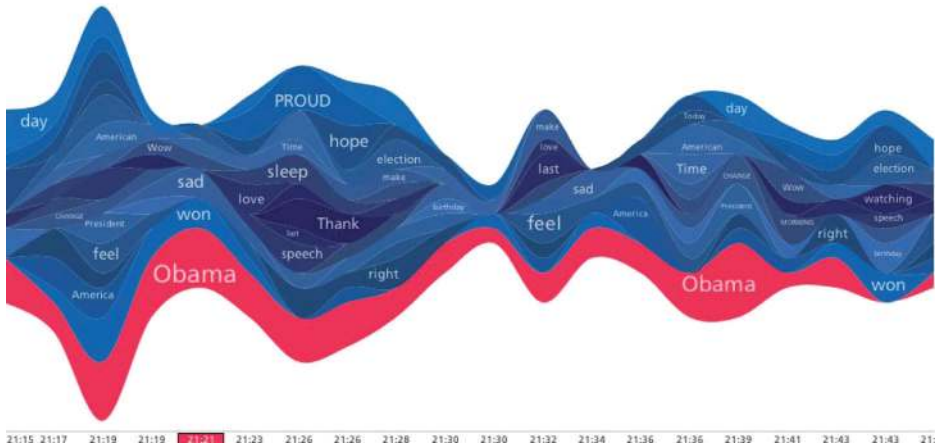
Figure 7. StreamGraph of the query 'happy' after the Obama election win Produced using Clark's (2008) system, Twitter StreamGraphs.

tag is potential target of evaluation (Figure 6). Of course in Figure 6 we have factored out time, perhaps the most important dimension in Twitter discourse. This omission means that the tweet starts to look like an artefact rather than a text unfolding dynamically over the twitter stream.

We can, however, gain some insight into what a more time-based representation might be if we consider Figure 7, a StreamGraph (Byron and Wattenberg, 2008; Clark, 2008; Havre et al., 2002) showing the unfolding of tweets that contain the string 'happy'. This graph was produced using Clark's (2008) system Twitter StreamGraphs. StreamGraphs are an example of a text visualization technique that does not efface logogenesis, the unfolding of text over time (Zappavigna, forthcoming). While an area graph usually shows a single data series, StreamGraphs are a form of stacked area graph that represent multiple data series by stacking one on top of the other. In a stacked area graph the height of the curve at a given point represents the total frequency of all features at that point and thus each data series should be read as starting at zero rather than as their accumulative height. This makes the graphing technique most useful to a linguist interested in the general trend of a data series, or in other words, the qualitative ebb and flow of the annotated units over the time series. It is also a useful technique for appreciating the relationships between the data series as they unfold by the overall impression of the relative amount of colour.

Examples of the tweets represented in the StreamGraph include:

Wow, I'm so **happy**, bring it home **Obama**

Obama:) I'm so **happy** that I'm crying! Hahhaa. Yes. He got his daughters a puppy.:) He is an amazing person and has well deserved this.

The stream along the bottom of the graph shows the coupling of 'happy (POSITIVE AFFECT) and 'Obama' unfolding over time (represented as intervals along the *X*-axis). The other

streams present in the graph are lexical couplings in temporally related tweets. There is a fairly consistent relationship between 'happy' and 'Obama' in the time window graphed. Other evaluative items in the environment of 'Obama' and 'happy' were: 'proud', 'sad', 'feel', 'hope', 'right', 'won', 'love', etc.

Given that tweets unfold over time along private and public streams, considering them from this dynamic perspective is crucial when making claims about the discourse. The argument that has been made in this article about the construction of ambient community assumes this dynamic perspective: these communities shift as hashtags shift, and different couplings of ideational and interpersonal meaning are established depending on what people are talking about at a given time. Theoretically, some topics might remain fairly constant such as, for example, the 'microgenre' of complaining about something, usually technological, that is not working for a user, marked by the #fail tag. For example, the following is typical of complaints regarding software problems:

> 'Vista spent 45 minutes installing updates.. only to say after rebooting that the update has failed and all changes are rolled back. **#fail.'**

Other tags may be more likely to shift with social and political concerns such as '#Iran' which at the time of writing was a tag facilitating political protest on Twitter in order to bypass mainstream media censorship: 'hopes that people around the world will help civilians rally for independence against unjust religious and political oppression in #Iran'. This functioning of the hashtag as inline metadata is a novel and emergent form of punctuation usage. The use of these markers appears to be related to a larger movement that Knox identifies in the evolution of what he refers to as interpersonal punctuation:

> The trajectory for interpersonal punctuation … begins with boundary marking, moves to punctuating speech function, and then to punctuating attitude and identity. At the same time, the prosody of punctuation spans (potentially) longer stretches of text, with the punctuation of attitude and identity through emoticons now able to spread over entire messages. (Knox, 2009)

For example, commas may be used to mark clause boundaries, question marks to indicate speech function and a smiley emoticon to mark positive affect. To Knox's trajectory we would add the punctuation of ambient affiliation via hashtags. This is the beginning of searchable talk where people mark the ideational targets of appraisal that act as boundaries defining ambient communities. They mark these targets both to indicate a stance and to make their talk findable.

Conclusion: microblogging and affiliation

This article has presented a case study showing the role of hashtags and evaluative language in affiliation on Twitter. The study involved a corpus-based discourse analysis of tweets about the victor of the 2008 US Presidential election. It has argued that the outpouring of emotional and evaluative language seen in this corpus is an example of the way that Twitter seems oriented toward the expression of interpersonal meaning. Criticism of Twitter as a service facilitating inane and frequent status updates about users' activities seems to have missed the social point of twittering. Twitter offers a medium for expressing personal evaluation to a large body of listeners with which one can affiliate ambiently.

This article has argued that by presenting couplings of evaluative and ideational meanings, a tweet invites the follower or searcher to share in the values presented.

The expansion of typographic meaning potential seen in hashtag usage on Twitter is the beginning of 'searchable talk'. Hashtags are used to mark potential targets of evaluation and to render these as metadata that may be found by other users. Hashtag usage on Twitter is an example of leveraging one the essential affordances of New Media: the affordance of the database to render information searchable and to make visible relationships that would not otherwise be recognizable. As Manovich (2002: 228) notes in his argument that databases have become a key form of cultural logic, 'regardless of whether new media objects present themselves as linear narratives, interactive narratives, databases or something else, underneath on the level of material organization, they are all databases'. Hashtag usage in Microblogging is an example of deploying the semiotic potential of this material organization.

Searchable talk is online conversation where people actively render their talk more findable. Talk using typographic strategies such as hashtagging becomes louder and more bondable. It appears the linguistic corollary of Morville's (2006) concept of 'ambient findability', where information can be found in any location. Now, instead, information can be found anywhere you have access to online talk and the wealth of online discourse renders Twitter an interpersonal search engine. In other words, Twitter is the place you go when you want to find out what people are saying about a topic right now and in order to involve yourself in communities of shared value that interest you in this given moment.

While this study has focused on Twitter, broader research into how we use Microblogging to affiliate ambiently is clearly important. A critical issue for any program of research into Microblogging is to consider it as a semiotic activity, that is, as an activity of making meaning with language. If we are to understand the relationships enacted by such meaning-making we require systematic ways of accounting for the role of language in creating social bonds. Affiliation is about more than connecting; it is about negotiating meanings within genres of language use. Hence, while studies of who connects with whom at which level of frequency will be useful, we need to understand the nature of what is being negotiated with language within particular patterns of social processes. The approach taken in this article has been informed by two complementary traditions in linguistics: Systemic Functional Linguistics and Corpus Linguistics, both of which offer useful, replicable strategies for analysing how language works in its social contexts. As our online talk becomes increasingly searchable and as microblogging changes our patterns of interaction, the data available to analysts for pursuing the question of ambient affiliation will grow. It is hoped that this article has been a first step into showing how we can study the inevitable linguistic complexity that arises as people commune online.

Notes

1. For an overview of research into Social Networking Sites see boyd and Ellison (2007).
2. American usage refers to the # character as a pound sign while Australian and UK usage refers to it as a hash. I will adopt the latter convention.
3. Technical terms of the Appraisal framework are distinguished from their everyday counterparts via small caps font.
4. n-grams are used in corpus linguistics to predict the sequence of items. In this instance, a 3-gram shows clusters of three words.

References

Anthony L (2009) AntConc3.2.0. Computer software. Tokyo.

Bakhtin M (1981) *The Dialogic Imagination*. Austin: University of Texas Press.

Baron N (2008) *Always On: Language in an Online and Mobile World*. Oxford and New York: Oxford University Press.

boyd dm and Ellison NB (2007) Social network sites: Definition, history, and scholarship. *Journal of Computer-mediated Communication* 13(1). Available at: http://jcmc.indiana.edu/vol13/issue11/boyd.ellison.html

Burnett G (2000) Information exchange in virtual communities: A typology. *Information Research* 5(4). Available at: http://jcmc.indiana.edu/vol9/issue2/burnett.html

Byron L and Wattenberg M (2008) Stacked graphs – geometry & aesthetics. *IEEE Transactions on Visualization and Computer Graphics* 14(6):1245–1252.

Clark J (2008) Twitter topic stream. Available at: http://www.neoformix.com/Projects/TwitterStreamGraphs/view.php

Crystal D (2006) *Language and the Internet* (2nd edn). Cambridge: Cambridge University Press.

Firth JR (1957) *Papers in Linguistics 1934–1951*. London: Oxford University Press.

Goffman E (1963) *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York: The Free Press.

Hagel J and Armstrong A (1997) *Net gain: Expanding Markets Through Virtual Communities*. Boston, MA: Harvard Business School Press.

Halliday MAK (1994) *An Introduction to Functional Grammar* (2nd edn). London: Arnold.

Halliday MAK and Matthiessen CMIM (1999) *Construing Experience Through Meaning: A Language-based Approach to Cognition*. London: Cassell.

Halliday MAK and Matthiessen CMIM (2004) *An Introduction to Functional Grammar* (3rd edn). London: Arnold.

Havre S Hetzler E, Whitney P and Nowell L (2002) ThemeRiver: Visualizing thematic changes in large document collections. *IEEE Transactions on Visualisation and Computer Graphics* 8(1): 9–20.

Herring SC (1996) *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives*. Amsterdam and Philadelphia: Benjamins.

Herring SC (2004) Computer-mediated discourse analysis: An approach to researching online communities. In: Barab SA, Kling R and Gray JH (eds) *Designing for Virtual Communities in the Service of Learning*. Cambridge and New York: Cambridge University Press, 338–376.

Herring SC (2008) Virtual community. In: Given LM (ed.) *Encyclopedia of Qualitative Research Methods*. Los Angeles, CA: Sage, 920–921.

Jakobson R (1971) *Selected writings. Second expanded edition*. The Hauge: Mouton.

Jones Q (1997) Virtual-communities, virtual settlements & cyber-archaeology: A theoretical outline. *Journal of Computer-mediated Communication* 3(3). Available at: http://www.ascusc.org/jcmc/vol3/issue3/jones.html

Knight N (2008) 'Still cool…and American too!': An SFL analysis of deferred bonds in internet messaging humour. In: Nørgaard N (ed.) *Systemic Functional Linguistics in Use, Odense Working Papers in Language and Communication* 29: 481–502.

Knight N (forthcoming) Naming culture in convivial conversational humour. In: Mahboob A and Knight NK (eds) *Directions in Appliable Linguistics*. London: Continuum.

Knox J (2009) Punctuating the home page: Image as language in an online newspaper. *Discourse and Communication* 3(2): 19–53.

Kress GR and Van Leeuwen T (2006) *Reading Images: The Grammar of Visual Design* (2nd edn). London and New York: Routledge.

Lemke JL (1984) *Semiotics and Education*. Toronto: Victoria University.

Lemke JL (1998) Multiplying meaning: Visual and verbal semiotics in scientific text. In: Martin JR and Veel R (eds) *Reading Science: Critical and Functional Perspectives on Discourses of Science*. London; New York: Routledge, 87–114.

Manovich L (2002) *The Language of New Media*. Cambridge, MA and London: MIT Press.

Martin JR (2000) Beyond exchange: Appraisal systems in English. In: Hunston S and Thompson G (eds) *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 142–175.

Martin JR and White PRR (2005) *The Language of Evaluation: Appraisal in English*. New York: Palgrave Macmillan.

McCracken H (2009) What i know about twitter. Available at: http://technologizer.com/2009/03/15/what-i-know-about-twitter/

Mischaud E (2007) *Twitter: Expressions of the Whole Self. An Investigation into User Appropriation of a Web-based Communications Platform* Master's thesis, Department of Media and Communications. London: London School of Economics and Political Science.

Morville P (2006) *Ambient Findability: What We Find Changes Who We Become*. Cambridge: O'Reilly.

O'Donnell M (2008) *UAM Corpus Tool 2.0*. Computer Application. Madrid.

Rheingold H (1993) *The Virtual Community*. Cambridge, Massachusetts: MIT Press.

Rocketboom (2009) The Twitter global mind. Available at: http://www.rocketboom.com/rb_09_mar_09_twitter_global_mind/

Vander Wal T (2007) Folksonomy coinage and definition. Available at: http://www.vanderwal.net/folksonomy.html

Wellman B (2001) Physical place and CyberPlace: The rise of personalized networking. *International Journal of Urban and Regional Research* 25: 227–252.

Zappavigna (forthcoming) Visualizing logogenesis: Preserving the dynamics of meaning. In: Dreyfus S, Stenglin M and Hood S (eds) *Semiotic Margins*. London: Continuum.

Zappavigna M, Dwyer P and Martin JR (2008) Syndromes of meaning: Exploring patterned coupling in a NSW Youth Justice Conference. In: Mahboob A, and Knight K (eds) *Questioning Linguistics*. Newcastle: Cambridge Scholars Publishing, 103–117.

Zhao S (2010) Intersemiotic relations as logogenetic patterns: Towards the restoration of the time dimension in hypertext description. In: Bednarek M and Martin JR (eds) *New Discourse on Language: Functional Perspectives on Multimodality, Identity, and Affiliation*. London: Continuum, 195–218.

Michele Zappavigna is a Research Fellow in Linguistics at the University of Sydney. Her major research interests are computer-mediated communication and social media. Currently she is working on a corpus-based study of the language of microblogging using a 100 million word Twitter corpus. Her book *The Discourse of Twitter and Social Media* will be published by Continuum in 2011. Michele also has an ongoing interest in text visualization as a tool to aid discourse analysts.