The Arab Spring and Its Social Media Audiences:

English and Arabic *Twitter* Users and Their Networks

Axel Bruns, Tim Highfield, and Jean Burgess

ARC Centre of Excellence for Creative Industries and Innovation, Queensland University of

Technology

Author Note

Axel Bruns, ARC Centre of Excellence for Creative Industries and Innovation, Queensland

University of Technology; Tim Highfield, ARC Centre of Excellence for Creative Industries

and Innovation, Queensland University of Technology; Jean Burgess, ARC Centre of

Excellence for Creative Industries and Innovation, Queensland University of Technology.

Correspondence concerning this article should be addressed to Axel Bruns, ARC Centre of

Excellence for Creative Industries and Innovation, Queensland University of Technology,

Z1-515, Creative Industries Precinct, Musk Ave, Kelvin Grove 4059, Australia. Email:

a.bruns@qut.edu.au

Abstract

While popular media narratives about the role of social media in driving the events of the 2011 'Arab Spring' are likely to overstate the impact of *Facebook* and *Twitter* on these uprisings, it is nonetheless true that protests and unrest in countries from Tunisia to Syria generated a substantial amount of social media activity. On *Twitter* alone, several millions of tweets containing the hashtags #libya or #egypt were generated during 2011, both by directly affected citizens of these countries, and by onlookers from further afield. What remains unclear, though, is the extent to which there was any direct interaction between these two groups (especially considering potential language barriers between them).

Building on hashtag datasets gathered between January and November 2011, this paper compares patterns of *Twitter* usage during the popular revolution in Egypt and the civil war in Libya. Using custom-made tools for processing 'big data', we examine the volume of tweets sent by English-, Arabic-, and mixed-language *Twitter* users over time, and examine the networks of interaction (variously through @replying, retweeting, or both) between these groups as they developed and shifted over the course of these uprisings. Examining @reply and retweet traffic, we identify general patterns of information flow between the English- and Arabic-speaking sides of the Twittersphere, and highlight the roles played by users bridging both language spheres.

Keywords: Twitter, Arab Spring, language networks, Egypt, Libya

The Arab Spring and Its Social Media Audiences: English and Arabic *Twitter* Users and

Their Networks

The 'Arab Spring' uprisings in 2011 saw widespread anti-government protests, and

some régime changes, in many Middle Eastern and North African (MENA) countries, from

Libya and Tunisia to Bahrain and Syria. Social media were among the tools used by

protesters to organize themselves and to disseminate footage from rallies. These were not

only used by local activists, but also attracted comments from a worldwide media audience,

for example in Twitter hashtag conversations such as #egypt and #libya. These hashtags were

used to mediate a wide range of practices of political participation among a diverse group of

social media users – from distanced observation and information-sharing in a globalized

'ambient journalism'[1] environment through to narration of direct experience and even

coordination of on-the-ground activities. However, there is no reason to assume that these

diverse activities were really 'connected' via the hashtag, or that one geographically or

culturally distinct group of users ever encountered another, hence highlighting the question of

whether social media, in such contexts, facilitates the flow of information across social

boundaries. This paper addresses these questions via an analysis of language differences in

social media communication focused on the Arab Spring, in doing so describing new

methods for the analysis of large-scale *Twitter* data.

We focus on discussions on *Twitter* concerning the uprisings in Egypt and Libya,

tracked between January and November 2011. These two cases, showing citizen opposition to

long-serving leaders, ultimately took different forms in their pursuit of revolution. The

Egyptian uprising initially saw a short series of large protests in January and February 2011,

---

[1] Alfred Hermida, "From TV to Twitter: How Ambient News Became Ambient Journalism," *M/C Journal* 13.2

(2010), accessed July 25, 2013, http://journal.media-culture.org.au/index.php/mcjournal/article/view/220.

resulting in the ouster of President Hosni Mubarak. In Libya, anti-government protests

quickly transformed into a civil war, resulting in months of bloodshed before the capture and

death of Libyan leader Colonel Gaddafi. In both cases, developments were accompanied by

widespread discussion on *Twitter*, in both Arabic and English. Our focus in this article is on

the relative levels of activity in Arabic, English, and mixed-language tweets featuring the

#egypt and #libya hashtags, and on the interactions between these different linguistic groups.

This enables us to track the changing circumstances of these revolutionary conflicts, and to

examine the relative contributions of different language groups to their discussion.

Context, Background and Approach

The organization and coverage of public protests is one of many purposes for which

Twitter has been used; many other social, political, and educational functions have also been

identified.[2] However, the specific contribution made by the platform remains debatable. In

June 2009, *Twitter* was viewed as the medium of choice for activists, both local and

international, to dispute the Iranian election result using the #iranelection hashtag[3,4], to the

point that the Iranian protests were dubbed an (ultimately unsuccessful) '*Twitter* revolution'.

At the same time, opinions remain divided about the extent to which these and other protests

were in a narrower sense *led* by activists using social media to express their views and

---

[2] Kate Crawford, "Following You: Disciplines of Listening in Social Media," *Continuum* 23.4 (2009),

doi:10.1080/10304310903003270.

[3] Alex Burns and Ben Eltham, "Twitter Free Iran: An Evaluation of Twitter's Role in Public Diplomacy and

Information Operations in Iran's 2009 Election Crisis," *Record of the Communications Policy &*

Research Forum 2009, Sydney, November 19-20, 2009, accessed July 25, 2013,

http://www.networkinsight.org/publications/record_of_the_2009_cprf.html/group/16.

[4] Devin Gaffney, "#IranElection: Quantifying Online Activism" (paper presented at WebSci10, Raleigh, NC,

Apr. 26, 2010), accessed July 25, 2013, http://journal.webscience.org/295.

orchestrate resistance[5,6,7,8]. On the evidence available, it appears that social media were *additional* communication tools for activists, rather than drivers of the demonstrations themselves.[9] The Arab Spring uprisings have attracted similar descriptions, as social media are used to share details about protests and generate support for movements, in a highly hybridized media environment[10] in which *Twitter* has achieved increased uptake both in the population at large and among news organizations and journalists themselves. Although the Egyptian and Libyan governments attempted to block domestic Internet access during the uprisings, protestors used workarounds to post to *Twitter*.[11] Once the Egyptian blackout was lifted, mobile phone videos were uploaded directly from the demonstrations to *YouTube*[12], and shared through social media. The volume of tweets hashtagged #egypt or #libya

---

[5] Barrie Axford, "Talk about a Revolution: Social Media and the MENA Uprisings," *Globalizations* 8.5 (2011), doi:10.1080/14747731.2011.621281.

[6] Walid El Hamamsy, "BB = BlackBerry or Big Brother: Digital Media and the Egyptian Revolution," *Journal of Postcolonial Writing* 47.4 (2011), doi:10.1080/17449855.2011.590325.

[7] Malcolm Gladwell, "Does Egypt Need Twitter?" *The New Yorker: News Desk*, Feb. 2, 2011, accessed July 25, 2013, http://www.newyorker.com/online/blogs/newsdesk/2011/02/does-egypt-need-twitter.html.

[8] Evgeny Morozov, "Facebook and Twitter Are Just Places Revolutionaries Go," *The Guardian: Comment Is Free*, Mar. 7, 2011, accessed July 25, 2013, http://www.guardian.co.uk/commentisfree/2011/mar/07/facebook-twitter-revolutionaries-cyber-utopians.

[9] Malcolm Gladwell, "Twitter, Facebook, and Social Activism: Small Change – Why the Revolution Will Not Be Tweeted," *The New Yorker*, Oct. 4, 2010, accessed July 25, 2013, http://www.newyorker.com/reporting/2010/10/04/101004fa_fact_gladwell.

[10] Andrew Chadwick, "The Hybrid Media System" (paper presented at the 6th European Consortium for Political Research General Conference, Reykjavík, Iceland, August 25-26, 2011.

[11] Jillian C. York, "Egypt: A Voice in the Blackout, Thanks to Google and Twitter," *Global Voices*, Feb. 1, 2011, accessed July 25, 2013, http://globalvoicesonline.org/2011/02/01/egypt-a-voice-in-the-blackout-thanks-to-google-and-twitter/

[12] El Hamamsy, "BB=Blackberry".

highlights the attention which the uprisings received from *Twitter* users both domestic and further afield; however, there are questions about whether *Twitter* was a stable means of coordinating demonstrations on the ground, or primarily a channel for international observers to discuss the uprisings.[13]

Either way, coverage of the Arab Spring on *Twitter* provides important examples for the formation of issue publics through shared hashtags. By including '#egypt' or '#libya' in their tweets, *Twitter* users are connecting their comments to a wider discussion. Bruns and Burgess[14] argue that these conversations on common topics can create *ad hoc* issue publics, which can 'respond with great speed to emerging issues and acute events'. Such events include crises and emergencies, including civil unrest and natural disasters[15]; hashtags have been used to concentrate the flow of information from emergency authorities in such cases as the earthquakes in Christchurch, New Zealand (#eqnz), and the floods in Queensland, Australia (#qldfloods), both in 2011.[16,17] Indeed, the convention of using hashtags to mark

---

[13] Genevieve Barrons, "'Suleiman: Mubarak Decided to Step Down #egypt #jan25 OH MY GOD': Examining the Use of Social Media in the 2011 Egyptian Revolution," *Contemporary Arab Affairs* 5.1 (2012), doi:10.1080/17550912.2012.645669.

[14] Axel Bruns and Jean Burgess, "The Use of Twitter Hashtags in the Formation of *Ad Hoc* Publics" (paper presented at the 6th European Consortium for Political Research General Conference, University of Iceland, Reykjavík, August 25-27, 2011), accessed July 25, 2013, http://eprints.qut.edu.au/46515/.

[15] Kate Starbird and Jeannie Stamberger, "Tweak the Tweet: Leveraging Microblogging Proliferation with a Prescriptive Syntax to Support Citizen Reporting," *Proceedings of the 7th International ISCRAM Conference, Seattle, WA, May 2010*, accessed July 25, 2013, http://repository.cmu.edu/silicon_valley/41/.

[16] Axel Bruns and Jean Burgess, "Local and Global Responses to Disaster: #eqnz and the Christchurch Earthquake," *Proceedings of Earth: Fire and Rain - Australian & New Zealand Disaster and Emergency Management Conference, Brisbane, Australia, 16-18 April 2012*, accessed July 25, 2013, http://eprints.qut.edu.au/50739/.

topical tweets first spread (before becoming fully integrated into *Twitter* architecture)

following their use in the coverage of wildfires in San Diego in 2007.[18,19]

Hashtagged discussions emerge without being controlled by any one organization or

user. Politicians, journalists, and emergency authorities may all be contributing to the

ongoing coverage, and may indeed be central figures to these discussions, but any account is

able to use, or ignore, hashtags in their own tweets. Any *Twitter* user could include #egypt or

#libya in their tweets, regardless of their proximity to the uprisings or involvement in the

protests (the range of participants discussing #egypt is studied by Lotan *et al.*[20]). Discussion

of events in Egypt, for example, also used the #Jan25 hashtag, signifying the 'Day of Revolt'

against President Hosni Mubarak. While this hashtag was widely used, it was not studied here

(tweets containing #Jan25 as well as #egypt are present within the dataset, however). In the

present context it should be especially noted that the use of these hashtags was not limited to

English speakers, in spite of the use of the English names of these countries as hashtags. At

the time of the uprisings, Arabic speakers were forced to use hashtags in Latin characters:

although *Twitter* supports the use of non-Latin characters in tweets themselves, as of January

---

[17] Axel Bruns, Jean Burgess, Kate Crawford, and Frances Shaw, *#qldfloods and @QPSMedia: Crisis Communication on Twitter in the 2011 South East Queensland Floods* (Brisbane: ARC Centre of Excellence for Creative Industries and Innovation, 2012), accessed July 25, 2013, http://cci.edu.au/floodsreport.pdf.

[18] Chris Messina, "Twitter Hashtags for Emergency Coordination and Disaster Relief," *Factory City*, Oct. 22, 2007, accessed July 25, 2013, http://factoryjoe.com/blog/2007/10/22/twitter-hashtags-for-emergency-coordination-and-disaster-relief/.

[19] Kate Starbird and Jeannie Stamberger, "Tweak the Tweet."

[20] Gilad Lotan, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, and danah boyd, "The Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions," *International Journal of Communication* 5 (2011), accessed July 25, 2013, http://ijoc.org/ojs/index.php/ijoc/article/view/1246.

2012 it was still testing its official support for right-to-left languages, especially regarding hashtags.[21] A key reason that many Arab Spring tweets combined Arabic text with an English hashtag was that the platform could not yet support right-to-left hashtags; left-to-right hashtags, by contrast, are automatically converted on publication to links to *Twitter* searches for those tags, providing easy access to the wider discussion on the topic. Therefore, a substantial volume of tweets mainly in English (as the international *lingua franca*), using Latin characters, were united with an at times equally significant volume of tweets in Arabic (as the common language of the MENA region), using Arabic characters, under the #egypt and #libya hashtags.

An additional technological innovation, in response to local Internet restrictions, can also explain some of the crossover between Arabic tweets and the English hashtag #egypt; the *Speak2Tweet* tool provided by *Google* and *Twitter* enabled users to tweet by calling an international telephone number and leaving a voice message, which was subsequently turned into a tweet and automatically accompanied by the #egypt hashtag.[22] Here, too, comments in various languages were combined with an English hashtag, thus aggregating multilingual tweets about the Egyptian revolution, although this does not necessarily translate to greater links between linguistic groups.

The resulting heterogeneous, bi- or multilingual nature of these hashtags immediately raises questions about the structure of their participant communities. Were there two or more separate groups of commenters, writing in Arabic and English but using the same hashtag? To what extent were bilingual users acting as boundary riders, connecting different language

---

[21] *Twitter*, "Twitter Translation Center Adds Right-to-Left Languages," *Twitter Blog*, Jan. 25, 2012, accessed July 25, 2013, http://blog.twitter.com/2012/01/twitter-translation-center-adds-right.html.

[22] Ujjwal Singh, "Some Weekend Work That Will (Hopefully) Enable More Egyptians to Be Heard," *Google Blog*, Feb. 1, 2011, accessed July 25, 2013, http://googleblog.blogspot.com.au/2011/01/some-weekend-work-that-will-hopefully.html.

communities and facilitating information flows between them? Previous studies of blogging

within the MENA region have noted the presence of blogs written in English alongside sites

in Arabic, leading Zuckerman to suggest that some of these sites may act as 'bridgeblogs',

intended to inform readers 'from a different nation, religion, or culture'.[23] Similarly, a study

of Arabic language blogs found a group of sites from across the Levant acting as an 'English

bridge', writing in both English and Arabic (Etling et al., 2010).[24] While Egyptian bloggers

did not necessarily act in this way, they played 'key roles in movement politics',[25] using the

Internet to circumvent the regulation of political organisation offline.

Indeed, prior to the uprisings, Egyptian blogging was credited as having 'intensified

current trends in politics and media', following media outlets' increasingly critical coverage

of the Mubarak régime.[26] Blogs became publishers of commentary or reports that could not

be featured in the traditional media, even those opposed to Mubarak. As an active Egyptian

blogosphere developed, the bloggers involved formed activist networks, in Egypt and abroad,

and with international journalists and other foreign bloggers. These links enable the wider

spread of information, sharing reports in Egypt and with a more distributed worldwide

audience.[27]

But writing in different languages does not automatically mean that an individual is

acting as a bridge between different groups of users. Herring *et al*.'s study of language

networks on *LiveJournal* found that English, and other languages, would be featured within

---

[23] Ethan Zuckerman, "Meet the Bridgebloggers," *Public Choice*, 134.1-2 (2008): 48, doi:10.1007/s11127-007-
    9200-y.

[24] Bruce Etling, John Kelly, Robert Faris, and John Palfrey, "Mapping the Arabic Blogosphere: Politics and
    Dissent Online," *New Media & Society*, 12.8 (2010), doi:10.1177/1461444810385096.

[25] Etling *et al*., "Mapping the Arabic Blogosphere", 1240.
[26] Tom Isherwood, "A New Direction or More of the Same? Political Blogging in Egypt," *Arab Media &
    Society* 6 (2008), 13, accessed July 25, 2013, http://www.arabmediasociety.com/?article=693.

[27] Tom Isherwood, "A New Direction", 9.

journal entries "in formulaic or emblematic uses", connecting users of different linguistic backgrounds even without a thorough understanding of the languages concerned.[28] The intent – or result – of using another language on *LiveJournal* is not to reach a new international audience, like Zuckerman's bridge bloggers, but to participate extensively within *LiveJournal*'s "cosmopolitan environment". Within the MENA region, bloggers will use both Arabic and English, or Arabic and French, in their posts, so that these languages are strongly interconnected, rather than used by distinct groups of bloggers.[29]

Use of English or Arabic may also be affected by the topics discussed in posts, and by the intended audience. Jansen's study of digital activism in the Middle East found that in Syria, Arabic was employed for discussion of 'more general issues like government, unemployment, and poverty', while English was used for comments on specific activist issues, including individual cases of arrest or harassment.[30] Jansen argues that blogging in English may be aimed at drawing more, global attention to particular issues. In their analysis of #sidibouzid tweets around the Tunisian revolution, Poell and Darmoni found that the most active users would post in multiple languages, tailoring their content for different audiences and acting to connect groups of users commenting on the uprising in Arabic, English, and French.[31] Although determining the subjects of tweets written in Arabic and English during

---

[28] Susan C. Herring, John C. Paolillo, Irene Ramos-Vielba, Inna Kouper, Elijah Wright, Sharon Stoerger, Lois Ann Scheidt, and Benjamin Clark, "Language Networks on LiveJournal," *Proceedings of the 40th Hawaii International Conference on System Sciences* (Los Alamitos, CA: IEEE Press, 2007), 9, doi:10.1109/HICSS.2007.320.

[29] Bruce Etling *et al.*, "Mapping the Arabic Blogosphere", 1229.

[30] Fieke Jansen, "Digital Activism in the Middle East: Mapping Issue Networks in Egypt, Iran, Syria and Tunisia," *Knowledge Management for Development Journal* 6.1 (2010), 48, doi:10.1080/19474199.2010.493854.

[31] Thomas Poell, and Kaouthat Darmoni, "Twitter as a Multilingual Space: The Articulation of the Tunisian Revolution through #sidibouzid," *NECSUS: European Journal of Media Studies* 1.1 (2012), accessed

the Arab Spring is beyond the scope of this article, the two languages (and others) may have

been employed in different tweets by individual users for similar purposes (for a content

analysis of #egypt tweets written in English, see Papacharissi & de Fatima Oliveira[32]).

Although the different languages represented in the datasets used here do not map

onto distinct geographic regions, it is important to distinguish the patterns of social media use

around the Arab Spring originating from local and international users. Howard *et al*.,

examining tweets containing geolocational data as well as the #egypt hashtag, found that the

early discussions were led by users found outside Egypt and its neighbors. In the weeks

leading up to Mubarak's resignation, a greater proportion of tweets came from local users

(and from users who did not provide location information).[33] In addition, Freelon's analysis

of several Arab Spring hashtag datasets found that spikes in *Twitter* activity in most

discussions were led by international users, rather than those within the MENA region.[34]

While language is not in itself an accurate means of determining location, comparing the use

of English and Arabic tweets over the same period allows us to examine whether spikes in

activity are led by particular linguistic groups.

July 25, 2013, http://www.necsus-ejms.org/twitter-as-a-multilingual-space-the-articulation-of-the-

tunisian-revolution-through-sidibouzid-by-thomas-poell-and-kaouthar-darmoni/.

[32] Zizi Papacharissi, and Maria de Fatima Oliveira, "Affective News and Networked Publics: The Rhythms of

News Storytelling on #Egypt," *Journal of Communication* 62 (2012), doi:10.1111/j.1460-

2466.2012.01630.x.

[33] Philip N. Howard, Aiden Duffy, Deen Freelon, Muzammil Hussain, Will Mari, and Marwa Mazaid, *Opening

Closed Régimes: What Was the Role of Social Media During the Arab Spring?*, 16-17, accessed July 25,

2013, http://pitpi.org/?p=1051.

[34] Deen Freelon, "The MENA Protests on Twitter: Some Empirical Data," *DFreelon.org*, May 19, 2011,

accessed July 25, 2013, http://dfreelon.org/2011/05/19/the-mena-protests-on-twitter-some-empirical-

data/.

In this study, we investigated the following questions through our comparison of Latin and non-Latin tweets:

1. Do tweets containing Latin and non-Latin characters follow similar patterns in responding to the events of the Arab Spring?

Based on previous research into online communication in the region, it would be expected that the use of English (and other Latin languages) would be prominent within the #egypt and #libya hashtags, however:

2. Is this use consistent throughout the uprisings, or does the volume of tweets from different language groups follow more varied patterns of troughs and spikes in response to specific events?

And finally:

3. Are the different language groups (Latin and non-Latin) interconnected; and is there evidence of bridging between these groups of *Twitter* users?

The presence of bridges in other online contexts suggests that an examination of user interactions in the #egypt and #libya hashtags would find some users acting as bridges between Arabic and English speakers. These hashtags also provide an automatic tie between the groups, through *Twitter*'s conversion of hashtags into hyperlinks. However, this does not necessarily mean that bridging is taking place; we examine the networks of @replies and retweets within the datasets to identify connections between users tweeting in Latin and non-Latin languages. As part of this examination, this study also establishes methods for identifying, and comparing, the languages used within large datasets of tweets, which have applications for further research into multilingual social media discussions.

Methods

Our datasets were collected through the *Twitter* API. Using a modified version of the

open source tool *yourTwapperkeeper* (see Bruns[35]), we tracked #egypt and #libya from early

2011 (23 Jan. 2011 for #egypt, 16 Feb. 2011 for #libya); for the purposes of our analysis, our

data collection period terminates on 30 November 2011. Due to the vagaries of collecting

data from the *Twitter* API, we cannot expect to have gathered a fully comprehensive dataset

for either hashtag: given the long timeframe of data collection, unavoidable outages both on

Twitter's and on our side will have combined to create several brief gaps in the archives.

Further, as the API is the only sanctioned access point to *Twitter* data at scale, it is impossible

to independently verify exactly how much data may have been excluded from collection:

short of comparing datasets with other researchers tracking these hashtags over the same

period, there is no reliable method for finding gaps in the data (see also Freelon on similar

limits to his study[36]). This is a fundamental problem of all research drawing on third-party

APIs; it is an unavoidable aspect of doing 'big data' research.[37]

At the same time, the overall volume of tweets which we did capture is immense, and

sufficient as a basis for the examination of broad patterns in *Twitter* activity. A chronological

overview of the data points to obvious gaps: for #egypt, we received no tweets at all on 31

Jan., 5-7 Feb., 31 Mar., 1 and 2 Apr., 2-4 Aug., 15 Sep., 16 Oct., and 23, 26, 27, and 29 Nov.

2011; for #libya, we are missing data for 31 Mar., 1 Apr., 15 Apr., 2-4 Aug., 15 Sep., 16 and

---

[35] Axel Bruns, "Switching from *Twapperkeeper* to *yourTwapperkeeper*," *Mapping Online Publics*, June 21,

    2011, accessed July 25, 2013, http://www.mappingonlinepublics.net/2011/06/21/switching-from-

    twapperkeeper-to-yourtwapperkeeper/.

[36] Deen Freelon, "The MENA Protests".

[37] danah boyd and Kate Crawford "Critical Questions for Big Data," *Information, Communication & Society*

    15.5 (2012), doi:10.1080:1369118X.2012.678878.

21 Oct., and 26 and 29 Nov. 2011.[38] This means that for #egypt, we missed 16 days in over ten months of data collection; for #libya, we missed 11 days in nine and a half months.

yourTwapperkeeper datasets are available in simple comma- or tab-separated value formats, containing the tweets themselves and a range of additional metadata; most importantly, these metadata include the numerical *Twitter* ID and username of the sender, as well as the exact timestamp of the tweet. Further metadata can be extracted from the tweets: chiefly, this includes the usernames of any *Twitter* users mentioned (through @replies or retweets), and the – usually shortened – URLs of any links included with the tweet. Further processing also reveals the specific type of tweet: by parsing its syntax, it is possible to distinguish between simple @replies and retweets (in the form "RT @user …", "MT @user …", "via @user …" or ""@user … "" – that is, enclosing the original tweet in quotation marks), or to identify tweets as *original* tweets that neither @reply to nor retweet another user.

In the present context, it is especially important to distinguish between tweets in different languages. The *Twitter* API itself does not provide sufficient information to make immediate distinctions: while amongst the metadata returned by the API is a language code for each tweet, that code is simply inherited from the language setting made globally by the tweet sender, and does not reflect the specific language of the tweet itself. Tweets by an Egyptian user, tweeting in Arabic, who left their global *Twitter* profile setting at the English default would be marked as 'English'; tweets by a French user who set their profile to French but converses in English and Arabic would be marked as French. The specific language of tweets can only be ascertained by individually analysing each tweet itself, then.

For the purposes of the present analysis, working with datasets that largely contain tweets in English and other European languages on the one hand, and in Arabic on the other,

---

[38] All dates and times here are in Cairo time.

this analysis can be considerably simplified: a useful approach to distinguishing these two groups is to examine whether tweets are written in Latin or non-Latin characters. While the non-Latin group will also contain tweets in various other scripts (Chinese, Japanese, Korean, etc.), the presence of such languages in our present datasets is negligible in comparison to Arabic script; additionally, in an analysis of conversational networks between *Twitter* users, such third language groups should form distinct conversational networks at a distance from the dominant Arabic and English groups. Similarly, any major distinctions in the Latin group should indicate the presence of various European languages.

Since all standard Latin characters and punctuation marks have been assigned ASCII character codes below 128, a simple method for coding tweets as 'Latin' or 'non-Latin' is to count the number of characters with a code above 127 in a tweet. Should that number pass a certain threshold, the tweet is coded as 'non-Latin'. Through a trial and tuning process[39], we determined that a threshold of 10 non-Latin characters results in a reliable distinction between Latin and non-Latin tweets. This threshold value is preferable to a strict zero as it allows for the presence of several accented characters as they are common in various European languages (äöüß, áéíóú, etc.) as well as for 'fancy' punctuation marks (" instead of ", etc.), all of which have also been assigned character codes above 127.

Such automated coding of tweets was implemented using *Gawk*, a programmable command-line tool for processing CSV/TSV data files.[40] In addition to coding the tweets themselves, we can also calculate a cumulative language score for each *Twitter* user participating in these datasets, indicating what percentage of their tweets was in non-Latin scripts. This can be used to distinguish different *Twitter* user groups: those posting mainly in

---

[39] See http://mappingonlinepublics.net/2012/01/28/creating-basic-twitter-language-metrics/ for details.

[40] Axel Bruns, "Creating Basic Twitter Language Metrics," *Mapping Online Publics*, Jan. 28, 2012, accessed

    July 25, 2013, http://www.mappingonlinepublics.net/2012/01/28/creating-basic-twitter-language-

    metrics/.

Latin characters (in the present context, mainly in English); those posting mainly in non-Latin

characters (mainly in Arabic), and those using a mixture of scripts (and thus perhaps acting as

information brokers between different language communities). Similarly, we can calculate

the ratio of Latin and non-Latin tweets across all users per timeframe (e.g. per day or hour),

to show when different language communities were especially active.

Beyond this coding of language, we also extracted a range of other metrics from the

Twitter datasets (see Bruns for an extended discussion of these metrics and the methods used

to obtain them[41]): we track the number of tweets made (also broken down into tweet

categories including original tweets, @replies, retweets, and tweets containing URLs) as well

as the number of active users per timeframe; further, for each user we determine the number

of hashtagged tweets sent and received (again also broken down into the different tweet

categories).

Finally, following the 1/9/90 rule which has become an unofficial standard for

analyses of user communities where activity broadly follows a 'long tail' distribution[42,43,44],

---

[41] Axel Bruns, "Taking Twitter Metrics to a New Level," *Mapping Online Publics*, Jan. 2, 2012, accessed July

25, 2013, http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-

1/, http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-2/,

http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-3/,

http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-4/.

[42] Chris Anderson, "The Long Tail," *Wired* 12.10 (2004), accessed July 25, 2013,

http://www.wired.com/wired/archive/12.10/tail.html.

[43] Chris Anderson, *The Long Tail: Why the Future of Business Is Selling Less of More* (New York: Hyperion,

2006).

[44] Steven J.J. Tedjamulia, Douglas L. Dean, David R. Olsen, and Conan C. Albrecht, "Motivating Content

Contributions to Online Communities: Toward a More Comprehensive Theory," *Proceedings of the 38th

Annual Hawaii International Conference on System Sciences (HICSS)* (Los Alamitos, CA: IEEE Press,

2005), doi:10.1109/HICSS.2005.444.

we divide the userbase of active contributors into three groups: one group of lead users which contains the most active one percent of participants; a second group of highly engaged users which contains the next nine percent of active participants; and a third group comprising the remaining 90% of least active users. These divisions are determined by ranking users on the basis of the number of tweets they have contributed to the hashtag: the top one percent of users on this ranked list are included in the first group, the next nine percent in the second group, and the remaining userbase in the third group. Finally, a fourth group contains all those whose usernames are mentioned in @replies and retweets, but who did not themselves post to the hashtag. For each of the first three groups, we again becomes track their contribution to the hashtag over time, and determine overall patterns of activity such as their relative use of original tweets, @replies, retweets, or tweets containing URLs.

Overall Patterns

Based on this methodology, we are able to determine overall patterns for both #egypt and #libya, over the total period covered by each dataset – 23 Jan. to 30 Nov. 2011 for #egypt, 16 Feb. to 30 Nov. 2011 for #libya.

#egypt

In total, we captured some 7.48 million #egypt tweets from over 445,000 unique users between 23 Jan. and 30 Nov. 2011. *Twitter* activity for #egypt peaks at a significantly higher level during the early stages of the revolution than at any other subsequent point (Figure 1). While unfortunately, data for several days in this early period (31 Jan., 5-7 Feb.) are missing from our overall dataset (visible as gaps in the graphs which follow), the resignation of President Mubarak on 11 Feb. has the greatest resonance in the available data: we recorded more than 205,000 #egypt tweets from over 82,000 unique users during this day. During this

early stage the composition of the *Twitter* community is also markedly different from that

recorded during the majority of the overall period: throughout almost all of February, tweets

using Latin characters retain the majority; it is only on 26 Feb. that the balance first swings

towards non-Latin tweets. From then on, the situation is markedly different: from 1 Mar. to

30 Nov., an average of more than 75% of the #egypt tweets sent each day are composed in

non-Latin characters.

FIGURE 1 HERE

This demonstrates a substantial shift in attention: while during the first month, and

especially around the key days of régime change, a significant number of non-Arabic-

speaking users participate, their interest dissipates as the situation moves from outright

revolution to a more long-term reshaping of the political system; the remaining #egypt

userbase (an average of over 7,000 unique users per day, posting nearly 24,000 tweets per

day during the 1 Mar. to 30 Nov. period) is likely to be composed largely of Egyptian locals

and expatriates with a more direct interest in the continuing process of change.

One additional possible explanation for these changes is also the existence of the

alternative hashtags #Jan25, referring to the so-called "Day of Revolt" which ignited the

protest movement against the Mubarak régime. Notably, our data record only a relatively

minor spike of less than 9,500 #egypt tweets on 25 Jan., substantially less than the over

205,000 tweets on 11 Feb.; it is conceivable that the majority of early *Twitter* activity around

the Egyptian protests took place under the #Jan25 hashtag, shifting to #egypt only once the

initial aim of the protests (Mubarak's resignation) was achieved, and as the further passage of

time made the #Jan25 tag seem anachronistic. The #Jan25 tag may also have had

substantially greater resonance with directly involved local users, participating in or closely

following the 25 Jan. protests, than with onlookers further afield; it is possible, therefore, that #25Jan hashtag activities attracted a proportionally larger number of Egyptian (and generally Arabic-speaking) *Twitter* users, in turn leaving #egypt to be dominated by English speakers, and that this imbalance only changed once a greater number of Arabic speakers transitioned to #egypt.

Such shifts in the userbase can also be traced by examining the relative contributions made by each of the three user groups outlined above. Figure 2 indicates the percentage of all daily tweets contributed by the three groups, and shows significant activity by the normally less active groups especially during the first stage: until the end of February, the lead users contributed only an average of 36% of all tweets per day; from March onwards, the same group accounts for an average of 60% of all tweets each day. In other words, this early stage saw a substantially larger presence of – in the long term – less engaged users; when these users exit the hashtag conversation as the 'hot' phase of the revolution comes to an end, the two user groups who have a more long-term commitment to discussing political change in Egypt increasingly come to dominate the conversation. But it should also be noted that from July onwards, lead users are again pushed back, in favor of a greater contribution especially from the second group of users: this may point both to the growing frustration with the slow pace of changeover from the Supreme Council of the Armed Forces to a civilian administration, which began to be voiced at this time, and to the building anticipation of popular elections, which began on 28 Nov. 2011. It is interesting to note that while the balance of contributions by the three groups gradually shifts from mid-year, the total volume of #egypt tweets remains relatively stable.

FIGURE 2 HERE

For further illustration, Figure 3 specifically compares the daily contributions made by the least active group with the daily percentage of tweets in a Latin character set, and points to a strong correlation between these metrics. Especially during the early stage of the revolution, the presence of a large number of normally relatively inactive users also coincides with a large number of Latin (i.e. mainly English) tweets; this implies that Arabic-language users are especially well represented in the leading groups of most active contributors to #egypt, while less active contributors are more likely to be from non-Arabic backgrounds, and may have been attracted to the #egypt discussion largely because of the widespread media coverage of the revolution, but have limited interest in the longer-term process of political change.

FIGURE 3 HERE

These differences also become apparent from a further examination of the activity patterns for each of the three groups (Figure 4). As is to be expected, the lead users are responsible for a disproportionate percentage of all #egypt tweets; this one percent of most active users contributed nearly 56% of all tweets. Their tweets are also substantially more likely to be what original tweets (that is, neither @replies nor retweets – 64% of their tweets fall into this category); by contrast, the majority of the tweets contributed by the least active user group – 65% – are retweets. The leading user group are also most likely to share URLs: some 56% of their tweets contain hyperlinks to external resources, compared with under 40% for each of the other two groups.

A further striking difference between these three groups is evident from their language preferences. For the lead group, an average of nearly 75% of their tweets use non-Latin characters; this reduces to 63% for the highly engaged users, and drops to 43% for the

large group of least engaged users. This means that Arabic speakers are relatively

overrepresented amongst the most engaged groups, while the least engaged group of users

contains a substantially larger number of non-Arabic speakers.

FIGURE 4 HERE

#libya

Patterns in the #libya dataset are somewhat different from those for #egypt, as Figure 5

indicates. Over the course of the data collection period, we captured over 5.27 million tweets

originating from more than 476,000 users. Total usage of the hashtag spikes early on at over

320,000 tweets per day on 21 Feb. 2011, as first reports of unrest are covered by world

media, but after this relatively brief moment of heightened activity the #libya hashtag

continues to operate at a much lower volume: from the start of April, the daily average

remains at a comparatively low 10,500 tweets. As in #egypt, the number of unique users

contributing to the hashtag each day generally correlates with the number of tweets; it peaks

at over 80,000 on 22 Feb., but reaches only an average of 3,600 for the period after 1 April.

FIGURE 5 HERE

A notable difference from #egypt emerges with the percentage of Arabic (i.e. non-

Latin) tweets per day: here, #libya shows a surprisingly limited number of tweets using non-

Latin characters. From 16 Feb. to 15 Oct., the average percentage of non-Latin tweets

remains at a lowly 18%; it rises to 29% only during the last one and a half months. Contrary

to #egypt, fluctuations in language use cannot be traced back to the relative contributions

made by the different user groups: at 23%, the percentage of non-Latin tweets posted by the

top 1% group of lead users over the entire period differs little from that of the least active

group, at 27%.

Rather, an explanation for the generally comparatively low number of Arabic tweets

in the #libya dataset must be sought in the user demographics, and in the nature of the

conflict. In Egypt, where protests were centered on demonstrations in the urban setting of

Cairo, significant use of *Twitter* in covering the crisis may well have been considerably more

likely than in Libya, where régime change was achieved only after a long military campaign

unfolding across the country; additionally, differing Internet and social media take-up, and

subsequent blocking of access to such communication tools, is likely to have influenced the

respective level of domestic *Twitter* use in these countries. Media reports during the Libyan

civil war, suggesting that the Gaddafi régime attempted to block Libya's access to the global

Internet, would explain the low number of Arabic tweets in the #libya dataset; further, the

substantial rise in Arabic tweets from 20 Oct. 2011 may indicate that such restrictions were

lifted as the régime fell (Gaddafi himself was killed on that day).

Figure 6 again compares overall daily activity with the respective contributions made

by the three user groups. As before, the top 1% of most active users is generally responsible

for the vast majority of all tweets; over the entire period, they contribute some 57% of all

tweets, while the least active group only contribute 16% of all tweets. Again, however, the

contribution of the less active user groups also rises considerably when the overall number of

tweets peaks; on 23 Aug., for example, the lead user group accounts for less than 27% of all

tweets, with the other two groups driving overall hashtag activity on that day (the day rebels

overran Gaddafi's Bab al-Azizia compound in central Tripoli).

FIGURE 6 HERE

Compared #egypt, the activity patterns for these different user groups in #libya (Figure 7) show few notable trends. There is, as expected, a marked difference in the overall level of contributions made by the three groups; the lead users are also somewhat less likely to send retweets (56% of their tweets were retweets, compared to 66% of the tweets made by the least active group), and more likely to post original tweets (36% compared to 27%). There also is no clear pattern in the relative use of Latin or non-Latin scripts; differences between the groups are relatively minor.

FIGURE 7 HERE

This is remarkably different from #egypt, where lead users were substantially more likely to post original tweets (65% of their messages were neither @replies nor retweets), and to do so in Arabic (nearly 75% of their tweets used non-Latin script). What these observations strongly suggest is the relative absence – because of Internet blockages or a more limited take-up of *Twitter* – of a domestic *élite* of Libyan *Twitter* users reporting on the latest developments, as well as of an active ex-pat community to take up and disseminate their messages further. *Twitter* as a communications tool *was* used to document and discuss the unfolding events of the Libyan civil war – but more so by interested onlookers outside of the country, mainly using English to communicate, than by Libyan locals and their compatriots abroad.

Interactions between Language Groups

There are clear differences in the *Twitter* audiences for the #egypt and #libya streams, and the make-up of these groups changes substantially over the course of 2011. Of particular interest is the presence of different language groups, and the potential for interactions between them:

our interest is in determining to what extent information originating from predominantly

Arabic-speaking *Twitter* participants is able to reach English-speaking users, and vice versa.

Such interactions can be traced by examining the flow of @replies and retweets (collectively,

@mentions) between participating accounts; for both #egypt and #libya they consist largely

of retweets, since (as Figures 4 and 7 have demonstrated) less than ten percent of all tweets

are genuine @replies. For our analysis, this is useful: retweets are generally used by *Twitter*

contributors to pass along incoming information to their own networks of followers; where

we find evidence of significant connection between Arabic- and English-language users, we

may assume that information is transmitted across language boundaries.

To examine these questions, we focus on four distinct periods selected from the

overall *Twitter* feeds for #egypt and #libya. For #egypt, we examine the period of 1-28 Feb.,

which sees the major spike in *Twitter* activity, and is characterized by a relatively high

number of users (many from the less engaged groups) tweeting in Latin characters, and the

period of 15 June to 15 Sep., marked by a steady but less spectacular daily volume of tweets

and a predominance of non-Latin tweets. For #libya, we examine 16 Feb. to 15 Mar., a

comparable one-month period during the early stages of the uprising, reaching daily volumes

surpassing even those seen in #egypt but notable for the comparative absence of non-Latin

tweets, and 1 Aug. to 30 Sep., with steady levels of activity and a slightly higher incidence of

non-Latin tweets. For each of these periods, we again divide participating users into the three

groups of lead users, highly engaged users, and least active users, as well as a final group of

passive *Twitter* accounts whose usernames are mentioned in hashtagged tweets, but who do

not themselves post hashtagged tweets during the period.

We also calculate for each user the percentage of their tweets which use more than

our threshold value of ten non-Latin characters. On this basis, we divide the overall userbase

along new lines: into groups using predominantly Latin characters (less than 33% of their

tweets pass the non-Latin threshold); predominantly non-Latin characters (more than 66% of their tweets are non-Latin); and mixing both Latin and non-Latin tweets (between 33% and 66% of their tweets are using non-Latin characters). Such distinctions can only be made for active contributors to the hashtags, of course; for the group of passive accounts which are merely mentioned, we are unable to determine their position across the language divide. In the network graphs which follow, accounts with predominantly Latin (i.e. mostly English-language) tweets will be shown in blue; those with mainly non-Latin (i.e. Arabic) tweets in green; users posting a mixture of Latin and non-Latin tweets are marked in an intermediate color that reflects that mix; passive accounts, finally, are shown in grey. Connections between users are shown in the color of the originating user.

#egypt

The two periods in the overall #egypt dataset which we examine here are marked by a substantial shift in the language mix, from a substantial majority of Latin tweets to an even more significant predominance of non-Latin tweets. Figure 8a/b shows the relative presence of the three different language groups within the total community of users, as well as within the groups of more and less active users.

FIGURE 8a/b HERE

During the 1-28 Feb. period, users tweeting predominantly in Latin characters clearly dominate: more than 78% of all users fall into that category, while only 4% and 18%, respectively, belong to the 'mixed' and 'non-Latin' groups. The distribution within the least active user group largely matches this distribution. Towards the more active end, however, the distribution changes considerably: only 67% of the highly active group, and only 55% of the lead users, tweet predominantly in Latin characters, while the presence of 'non-Latin'

users grows to 22% and 33%. The most remarkable difference is for the mixed-language group, however: constituting only 3% of the least active group, they account for 10% of the highly active group, and make up nearly 13% of the lead user group. This indicates a considerable difference in commitment to the #egypt discussion: while larger numbers of English speakers may be interested enough to tweet or retweet the occasional message relating to the situation in Egypt, even at this early stage Arabic-speaking *Twitter* users are prepared to participate in significantly more depth.

Several months later, similar patterns persist, but the balance has shifted much further towards the 'non-Latin' group. They now constitute nearly 60% of the total userbase, and are again considerably overrepresented amongst the more engaged groups; over 85% of all lead users tweet predominantly in non-Latin characters. Similar to the earlier period, too, mixed-language contributors are disproportionately represented amongst the more active groups; here, however, they constitute a larger proportion of the second, highly engaged group (at nearly 10%), but only 6% of the lead user group. One explanation for this shift may be that the 'mixed' group is more likely to include native Arabic speakers who use English as a second language than native English speakers with some knowledge of Arabic; as the overall stream of the #egypt discussion shifts more towards the use of Arabic in these later months, users who were in the 'mixed' group during the earlier phase of the uprising may now be posting Arabic-language tweets so frequently that they have moved into the 'non-Latin' group as we have defined it.

Figure 9a/b compares the total network of *Twitter* exchanges between users through @replies and retweets during these periods. Connections are depicted in the color of the originating user: @replies and retweets by 'Latin' users are shown in blue; those by 'non-Latin' users in green; and those by users tweeting in a mixture of character sets in the corresponding mixed color. The balance between predominantly blue (Latin) and green (non-

Latin) regions in the network shifts substantially from the 1-28 Feb. to the 15 June to 15 Sep.

period. During the former period, in fact, some 68% of all connections through @replies and

retweets originate from the 'Latin' user group, 10% from the 'mixed' group, and 22% from

the 'non-Latin' group[45]; during the latter, the situation is reversed, and even more one-sided:

only 18% of all @mentions originate from 'Latin' users, 9% from 'mixed' users, and 73%

from 'non-Latin' participants. If the least active group of contributors is excluded from this

calculation, the situation changes slightly: for the earlier period, the 'Latin' group now

accounts for a slightly lower 64% of all @mentions; for the later period, however, the

contribution of 'non-Latin' users rises yet further, to over 78% of all @mentions. Several

outliers may be detected in these network graphs (especially amongst 'non-Latin' users in the

June-September period); it is likely that ideological, geographic, or other sociodemographic

factors are responsible for their separation from the core of the network.

        FIGURE 9a/b HERE

        The overall flows of information across the network, for which @replies and retweets

provide a proxy measure, can be examined further by visualizing aggregate flows (Figure

10a/b). These graphs show that interaction by 'Latin' and 'non-Latin' groups during both

periods is largely amongst themselves: the indicators of self-linking are considerably more

prominent in Figure 10a/b than any connections across language boundaries. During

February, some 80% of all @replies and retweets by 'Latin' users reference others in the

same group; 65% of the @mentions by 'non-Latin' users mention other 'non-Latin'

participants. For 'Latin' users, in fact, the second most prominent source of information are

---

[45] Here and throughout, these percentages refer to the relative number of connections (network edges) between users from these different language groups; we do not take into account the frequency with which such connections between any pair of participants may have been repeated during each timeframe (i.e. the specific weight of each network edge).

'passive' accounts: 10% of their tweets reference those accounts (amongst which news organizations and other sources will play an important role), often likely retweeting information while adding the '#egypt' hashtag to the original messages. Where they look beyond their own group, by contrast, 'non-Latin' users divide their attention almost equally between 'mixed' (14%) and 'Latin' (16%) sources; they draw on passive accounts only for 5% of their @mentions. The 'mixed' group, finally, act considerably differently: only 15% of their @replies and retweets are directed at other mixed-language users, but 42% reference 'Latin' accounts and 37% connect to 'non-Latin' accounts. While the overall contribution of the 'mixed' group to #egypt is relatively minor, therefore, their main role appears to be an attempt to bridge the major language groups.

FIGURE 10a/b HERE

During the period of 15 June to 15 Sep., the situation is reversed, and more: as originators of only 18% of all @mentions, 'Latin' users now play an even lesser role than 'non-Latin' users did during February. Due in part to their overall dominance, the 'non-Latin' group is similarly self-focused: over 82% of their tweets mention other 'non-Latin' users, with between four and 7 percent mentioning each of the other three groups. Conversely, as #egypt is now predominantly a non-Latin *Twitter* stream, the remaining 'Latin' users are also forced to look beyond their own group for more information: while 56% of their tweets continue to reference other 'Latin' participants, 11% draw on the 'mixed' group, and 14% contain @mentions of 'non-Latin' users. Indeed, if the 90% least active users are excluded from the analysis, the cross-language links from 'Latin' to 'non-Latin' users increase from 14% to over 19% (and from 11% to 12% for links to the 'mixed' group): those 'Latin' users who are amongst the most active overall contributors to #egypt are also significantly more

likely to seek information beyond their own group. The 'Latin' group also remains especially focused on 'passive' accounts, however: some 19% of their tweets continue to inject information from such non-participating accounts into the #egypt discussion, through retweeting. Finally, more so than during the earlier period, the 'mixed' accounts have also accepted the dominance of 'non-Latin' accounts: 53% of their @mentions reference those accounts, compared to only 21% referring to 'Latin' users. Intra-group @mentions remain characteristically low for this group: only 14% of their @mentions refer to fellow mixed-language accounts.

#libya

Dominated throughout by 'Latin' users, the situation in the #libya hashtag differs considerably from that in #egypt. During the early phase of the revolution, the overall #libya userbase presents what is nearly a mirror image of the situation in #egypt: some 82% of all participating users during this time fall into the 'Latin' category (Figure 11a/b). However, when broken down into the groups of more or less engaged users, the distribution of language groups becomes more complicated: while the second most active group again includes a larger number of 'non-Latin' and mixed-language users, that trend is reversed again for the leading user group. 'Non-Latin' users constitute 14% of the least active group, 20% of the second group of highly engaged users, but again only 14% of the lead group; by contrast, the mixed-language group accounts for only 3% of the least active group, 8% of the highly engaged group, and nearly 9% of the leading group. It appears that similar to #egypt, during this early phase the #libya hashtag attracted a substantial number of relatively random English-language commenters, a comparatively large number of fairly active Arabic-speaking users, but also a substantial number of very highly active English-language participants.

FIGURE 11a/b HERE

This pattern is even more pronounced for the August/September period. By this time, users tweeting mainly in non-Latin characters have become substantially more active in the #libya community; they now account for 29% of the total userbase, and constitute 30% of the highly engaged user group. Surprisingly, however, they have not only failed to make any inroads into the lead user group, but have indeed been pushed out of this group by an even more active English-language élite, to the point where they now constitute only 8% of that lead group. Further, the mixed-language group also appear to have been squeezed out of the overall hashtag community by this increasing language polarization: now accounting for only 2% of the total #libya userbase, they also constitute only 6% of the highly engaged group, and 4% of the lead user group.

Figure 12a/b again compares the overall network of @replies and retweets across the two periods we have chosen (16 Feb. to 15 Mar. and 1 Aug. to 30 Sep., respectively), and shows a gradual thinning of and cluster formation in the network: not only do connections between the predominantly 'Latin' and 'non-Latin' sections of the network weaken from the earlier to the latter period, but even within these sections themselves distinct, loosely connected clusters emerge (available space in this article does not permit us to examine the unifying traits of these distinct clusters). During the earlier period, nearly 80% of all connections through @mentions originated from the 'Latin' group of users, while the 'non-Latin' group accounted for just over 13%; the 'mixed' group contributed only 7% to the total number of @mentions. This distribution remains steady once the least active 90% of users are removed from the network, too. In August and September, during the final battle for control of Tripoli, the situation becomes more polarized: while at 79%, the 'Latin' dominance remains steady, the contribution of the 'mixed' group drops to only 4%, and that of 'non-

Latin' users increases to nearly 17%; if only the top 10% of most active users are considered, however, the 'Latin' group now accounts for over 85% of all @mentions, and the 'non-Latin' group drops back to just over 10%.

FIGURE 12a/b HERE

An analysis of the aggregate flow of information further supports these observations. Figure 13a/b is clearly dominated by the presence of 'Latin' users, who largely make intra-group @mentions (more than 85% of their @mentions are directed to other 'Latin' participants, in both periods); where they connect outside their own group at all, they do so mainly to 'passive' *Twitter* accounts (8% and 11% of their @mentions, respectively, are pointing to that group during the two periods, while @mentions of any of the other groups fail to account even for as little as 4% of the total @replies and retweets sent by 'Latin' users. FIGURE 13a/b HERE

Nonetheless, a small but internally active group of 'non-Latin' users does remain: respectively, during the two periods, 66% and 76% of the @mentions originating from 'non-Latin' users are directed at other members of that group. During February/March, 'Latin' users are the next most important information source for 'non-Latin' users, at 16%, followed by mixed-language users at 11% and 'passive' accounts at 6%; in August and September, however, external, 'passive' sources become more important (at 9%), while @mentions of the 'Latin' and 'mixed' groups drop to 8% and 6%, respectively. Finally, while in #libya the efforts of the 'mixed' group of users do not amount to a substantial level of activity, it is nonetheless notable that their information-sourcing processes do not reflect the balance of power which prevails within the #libya community: while during both periods, over 50% of

their @mentions refer to 'Latin' users, a similarly considerable over 30% of their mentions are directed to the significantly smaller group of 'non-Latin' users. This intermediary group of mixed-language users do continue to play a role in enabling an information flow across language boundaries, therefore, even if their more limited presence in the #libya hashtag means that direct connections between 'Latin' and 'non-Latin' users must play a greater role here, compared to #egypt.

Conclusion

Space available in this article has only allowed us to examine the broad patterns of *Twitter* usage by Arabic and English speakers in the Egyptian and Libyan uprisings, and to point to the relative presence of highly active élite users in each case; even this already highlights significant differences between the two cases. These differences are clearly aligned with sociodemographic and technological distinctions between the countries, as well as with the different course of events followed by each revolution.

We found that there is a substantially larger group of Arabic-speaking users participating in the #egypt discussion than in #libya; this observation supports research which found – albeit on the basis of geolocated tweets, which account for only a minute percentage of all messages on *Twitter* – that the Egyptian *Twitter* population is larger by an order of magnitudes than the Libyan.[46] As a consequence, discussion under the #libya hashtag is likely to consist largely of outsiders looking in, rather than – as in #egypt – of locals and expatriates discussing the unfolding political crisis in their country.

---

[46] Beatrice Karanja, "New Research Reveals How Africa Tweets," *African Arguments*, Jan. 26, 2012, accessed July 25, 2013, http://africanarguments.org/2012/01/26/new-research-reveals-how-africa-tweets-by-beatrice-karanja-portland-communications.

Even in #egypt, however, we found a substantial shift over time, from a comparative dominance of users tweeting in Latin characters to an overwhelmingly Arabic-speaking userbase. This shift may be driven in part by the early prominence of alternative hashtags – chiefly, the #Jan25 hashtag which referenced to the date of the first major demonstrations, and which subsided thereafter. But our analysis has also shown the already considerable presence of an Arabic-speaking élite amongst the top one percent of most active contributors to #egypt even at this early stage; as other users shifted from #Jan25 to #egypt proper, and as long-term interest by international participants waned, this established élite became the nucleus around which a largely Arabic-language discussion unfolded.

Our analysis of activity patterns in #egypt and #libya provides a complement especially to Lotan et al.'s analysis of the activities of a small group of highly active *Twitter* users who commented on the uprisings in Tunisia and Egypt.[47] Where that study traced patterns of dissemination for a limited number of high-profile examples, our research points to the degree to which information exchanges are able to bridge existing language divides. Though outside the scope of the present article, further work will be able to examine the relative prominence of specific news sources (as URLs cited in tweets, and/or as major *Twitter* contributors themselves) in the English- and Arabic-language networks, and the extent to which such resources are shared across the language divide, or specific to one or the other of these language communities; this will shed further light onto the relative uses of *Twitter* for disseminating both mainstream and eyewitness accounts of the uprisings to local and international followers of these hashtags.

---

[47] Gilad Lotan et al., "The Revolutions Were Tweeted."

Such analyses also enable us to move beyond simplistic arguments about whether or not the events of the Arab Spring constituted '*Twitter* revolutions' (see e.g. Sullivan[48] and Morozov[49] for examples of the opposing perspectives in this argument). The differences we have found between the Egyptian and Libyan uprisings already point to the fact that the real situation is far more complex, and not only highly dependent on national and regional specificities, but also considerably changeable over time. The substantial level of Arabic tweets in the case of #egypt certainly points to the fact that *Twitter* – and, by extension, other online media – did play a role in informing, organizing, and reporting protest activities in the country (and most likely continue to do so now, as post-election unrest persists), but this does not necessarily translate into support for the popular narrative of Egypt as a social media revolution. In Libya, the situation is notably different – here, the consistent lack of local *Twitter* activity makes it difficult to escape the conclusion that other, more conventional forms of communication were significantly more important to the successful pursuit of régime change, and that *Twitter* interest in the uprising was driven largely by onlookers from further afield. Future research will show whether – in the wake of these political transformations – *Twitter* and other online and social media will become established for the long term as tools for political communication in both countries.

---

[48] Andrew Sullivan, "Could Tunisia Be the Next Twitter Revolution?", *The Atlantic: The Daily Dish*, Jan. 13, 2011, accessed July 25, 2013, http://www.theatlantic.com/daily-dish/archive/2011/01/could-tunisia-be-the-next-twitter-revolution/177302/.

[49] Evgeny Morozov, "Facebook and Twitter".

Bibliography

Anderson, Chris. "The Long Tail." *Wired* 12.10 (2004). Accessed July 25, 2013.

  http://www.wired.com/wired/archive/12.10/tail.html.

———. *The Long Tail: Why the Future of Business Is Selling Less of More*. New York: Hyperion, 2006.

Axford, Barrie. "Talk about a Revolution: Social Media and the MENA Uprisings." *Globalizations* 8.5 (2011):

  681-686. doi:10.1080/14747731.2011.621281.

Barrons, Genevieve. "'Suleiman: Mubarak Decided to Step Down #egypt #jan25 OH MY GOD': Examining the

  Use of Social Media in the 2011 Egyptian Revolution." *Contemporary Arab Affairs* 5.1 (2012): 54-67.

  doi:10.1080/17550912.2012.645669.

boyd, danah, and Kate Crawford. "Critical Questions for Big Data." *Information, Communication & Society*

  15.5 (2012): 662-679. doi:10.1080:1369118X.2012.678878.

Bruns, Axel. "Creating Basic Twitter Language Metrics." *Mapping Online Publics*, Jan. 28, 2012. Accessed

  July 25, 2013. http://www.mappingonlinepublics.net/2012/01/28/creating-basic-twitter-language-

  metrics/.

———. "Switching from *Twapperkeeper* to *yourTwapperkeeper*." *Mapping Online Publics*, June 21, 2011.

  Accessed July 25, 2013. http://www.mappingonlinepublics.net/2011/06/21/switching-from-

  twapperkeeper-to-yourtwapperkeeper/.

———. "Taking Twitter Metrics to a New Level." *Mapping Online Publics*, Jan. 2, 2012. Accessed July 25,

  2013. http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-1/,

  http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-2/,

  http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-3/,

  http://www.mappingonlinepublics.net/2012/01/02/taking-twitter-metrics-to-a-new-level-part-4/.

———, and Jean Burgess. "Local and Global Responses to Disaster: #eqnz and the Christchurch Earthquake."

  Proceedings of Earth: Fire and Rain - Australian & New Zealand Disaster and Emergency Management

  Conference, Brisbane, Australia, 16-18 April 2012. Accessed July 25, 2013.

  http://eprints.qut.edu.au/50739/.

———, and Jean Burgess. "The Use of Twitter Hashtags in the Formation of *Ad Hoc* Publics." Paper presented

  at the 6th European Consortium for Political Research General Conference, University of Iceland,

  Reykjavík, August 25-27, 2011. Accessed July 25, 2013. http://eprints.qut.edu.au/46515/.

———, Jean Burgess, Kate Crawford, and Frances Shaw. *#qldfloods and @QPSMedia: Crisis Communication*

on Twitter in the 2011 South East Queensland Floods. Brisbane: ARC Centre of Excellence for Creative

Industries and Innovation, 2012. Accessed July 25, 2013. http://cci.edu.au/floodsreport.pdf.

Burns, Alex, and Ben Eltham. "Twitter Free Iran: An Evaluation of Twitter's Role in Public Diplomacy and

Information Operations in Iran's 2009 Election Crisis." *Record of the Communications Policy &*

Research Forum 2009. Sydney, November 19-20, 2009. Accessed July 25, 2013.

http://www.networkinsight.org/publications/record_of_the_2009_cprf.html/group/16.

Chadwick, Andrew. "The Hybrid Media System." Paper presented at the 6th European Consortium for Political

Research General Conference, Reykjavík, Iceland, August 25-26, 2011.

Crawford, Kate. "Following You: Disciplines of Listening in Social Media." *Continuum* 23.4 (2009): 525-535.

doi:10.1080/10304310903003270.

El Hamamsy, Walid. "BB = BlackBerry or Big Brother: Digital Media and the Egyptian Revolution." *Journal of*

Postcolonial Writing 47.4 (2011): 454-466. doi:10.1080/17449855.2011.590325.

Etling, Bruce, John Kelly, Robert Faris, and John Palfrey. "Mapping the Arabic Blogosphere: Politics and

Dissent Online." *New Media & Society* 12.8 (2010): 1225-1243. doi:10.1177/1461444810385096.

Freelon, Deen. "The MENA Protests on Twitter: Some Empirical Data." *DFreelon.org*, May 19, 2011.

Accessed July 25, 2013. http://dfreelon.org/2011/05/19/the-mena-protests-on-twitter-some-empirical-

data/.

Gaffney, Devin. "#IranElection: Quantifying Online Activism." Paper presented at WebSci10, Raleigh, NC,

Apr. 26, 2010. Accessed July 25, 2013. http://journal.webscience.org/295.

Gladwell, Malcolm. "Does Egypt Need Twitter?" *The New Yorker: News Desk*, Feb. 2, 2011. Accessed July 25,

2013. http://www.newyorker.com/online/blogs/newsdesk/2011/02/does-egypt-need-twitter.html.

———. "Twitter, Facebook, and Social Activism: Small Change – Why the Revolution Will Not Be Tweeted."

The New Yorker, Oct. 4, 2010. Accessed July 25, 2013.

http://www.newyorker.com/reporting/2010/10/04/101004fa_fact_gladwell.

Hermida, Alfred. "From TV to Twitter: How Ambient News Became Ambient Journalism." *M/C Journal* 13.2

(2010). Accessed July 25, 2013. http://journal.media-

culture.org.au/index.php/mcjournal/article/view/220.

Herring, Susan C., John C. Paolillo, Irene Ramos-Vielba, Inna Kouper, Elijah Wright, Sharon Stoerger, Lois

Ann Scheidt, and Benjamin Clark. "Language Networks on LiveJournal." *Proceedings of the 40th*

Hawaii International Conference on System Sciences. Los Alamitos, CA: IEEE Press, 2007. doi:10.1109/HICSS.2007.320.

Howard, Philip N., Aiden Duffy, Deen Freelon, Muzammil Hussain, Will Mari, and Marwa Mazaid. *Opening Closed Régimes: What Was the Role of Social Media During the Arab Spring?*. 16-17. Accessed July 25, 2013. http://pitpi.org/?p=1051.

Isherwood, Tom. "A New Direction or More of the Same? Political Blogging in Egypt." *Arab Media & Society* 6 (2008): 1-17. Accessed July 25, 2013. http://www.arabmediasociety.com/?article=693.

Jansen, Fieke. "Digital Activism in the Middle East: Mapping Issue Networks in Egypt, Iran, Syria and Tunisia." *Knowledge Management for Development Journal* 6.1 (2010): 37-52. doi:10.1080/19474199.2010.493854.

Karanja, Beatrice. "New Research Reveals How Africa Tweets." *African Arguments*, Jan. 26, 2012. Accessed July 25, 2013. http://africanarguments.org/2012/01/26/new-research-reveals-how-africa-tweets-by-beatrice-karanja-portland-communications.

Lotan, Gilad, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, and danah boyd. "The Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions." *International Journal of Communication* 5 (2011): 1375-1405. Accessed July 25, 2013. http://ijoc.org/ojs/index.php/ijoc/article/view/1246.

Messina, Chris. "Twitter Hashtags for Emergency Coordination and Disaster Relief." *Factory City*, Oct. 22, 2007. Accessed July 25, 2013. http://factoryjoe.com/blog/2007/10/22/twitter-hashtags-for-emergency-coordination-and-disaster-relief/.

Morozov, Evgeny. "Facebook and Twitter Are Just Places Revolutionaries Go." *The Guardian: Comment Is Free*, Mar. 7, 2011. Accessed July 25, 2013. http://www.guardian.co.uk/commentisfree/2011/mar/07/facebook-twitter-revolutionaries-cyber-utopians.

Papacharissi, Zizi, and Maria de Fatima Oliveira. "Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt." *Journal of Communication* 62 (2012): 266-282. doi:10.1111/j.1460-2466.2012.01630.x.

Poell, Thomas, and Kaouthat Darmoni. "Twitter as a Multilingual Space: The Articulation of the Tunisian Revolution through #sidibouzid." *NECSUS: European Journal of Media Studies* 1.1 (2012). Accessed July 25, 2013. http://www.necsus-ejms.org/twitter-as-a-multilingual-space-the-articulation-of-the-tunisian-revolution-through-sidibouzid-by-thomas-poell-and-kaouthar-darmoni/.

Singh, Ujjwal. "Some Weekend Work That Will (Hopefully) Enable More Egyptians to Be Heard." *Google Blog*, Feb. 1, 2011. Accessed July 25, 2013. http://googleblog.blogspot.com.au/2011/01/some-weekend-work-that-will-hopefully.html.

Starbird, Kate, and Jeannie Stamberger. "Tweak the Tweet: Leveraging Microblogging Proliferation with a Prescriptive Syntax to Support Citizen Reporting." *Proceedings of the 7th International ISCRAM Conference, Seattle, WA, May 2010*. Accessed July 25, 2013. http://repository.cmu.edu/silicon_valley/41/.

Sullivan, Andrew. "Could Tunisia Be the Next Twitter Revolution?" *The Atlantic: The Daily Dish*, Jan. 13, 2011. Accessed July 25, 2013. http://www.theatlantic.com/daily-dish/archive/2011/01/could-tunisia-be-the-next-twitter-revolution/177302/.

Tedjamulia, Steven J.J., Douglas L. Dean, David R. Olsen, and Conan C. Albrecht. "Motivating Content Contributions to Online Communities: Toward a More Comprehensive Theory." *Proceedings of the 38th Annual Hawaii International Conference on System Sciences (HICSS)*. Los Alamitos, CA: IEEE Press, 2005. doi:10.1109/HICSS.2005.444.

Twitter. "Twitter Translation Center Adds Right-to-Left Languages." *Twitter Blog*, Jan. 25, 2012. Accessed July 25, 2013. http://blog.twitter.com/2012/01/twitter-translation-center-adds-right.html.

York, Jillian C. "Egypt: A Voice in the Blackout, Thanks to Google and Twitter." *Global Voices*, Feb. 1, 2011. Accessed July 25, 2013. http://globalvoicesonline.org/2011/02/01/egypt-a-voice-in-the-blackout-thanks-to-google-and-twitter/.

Zuckerman, Ethan. "Meet the Bridgebloggers." *Public Choice* 134.1-2 (2008): 47-65. doi:10.1007/s11127-007-9200-y.

Figures

Figure 1. #egypt tweets and unique users per day, compared with daily percentage of non-Latin tweets.
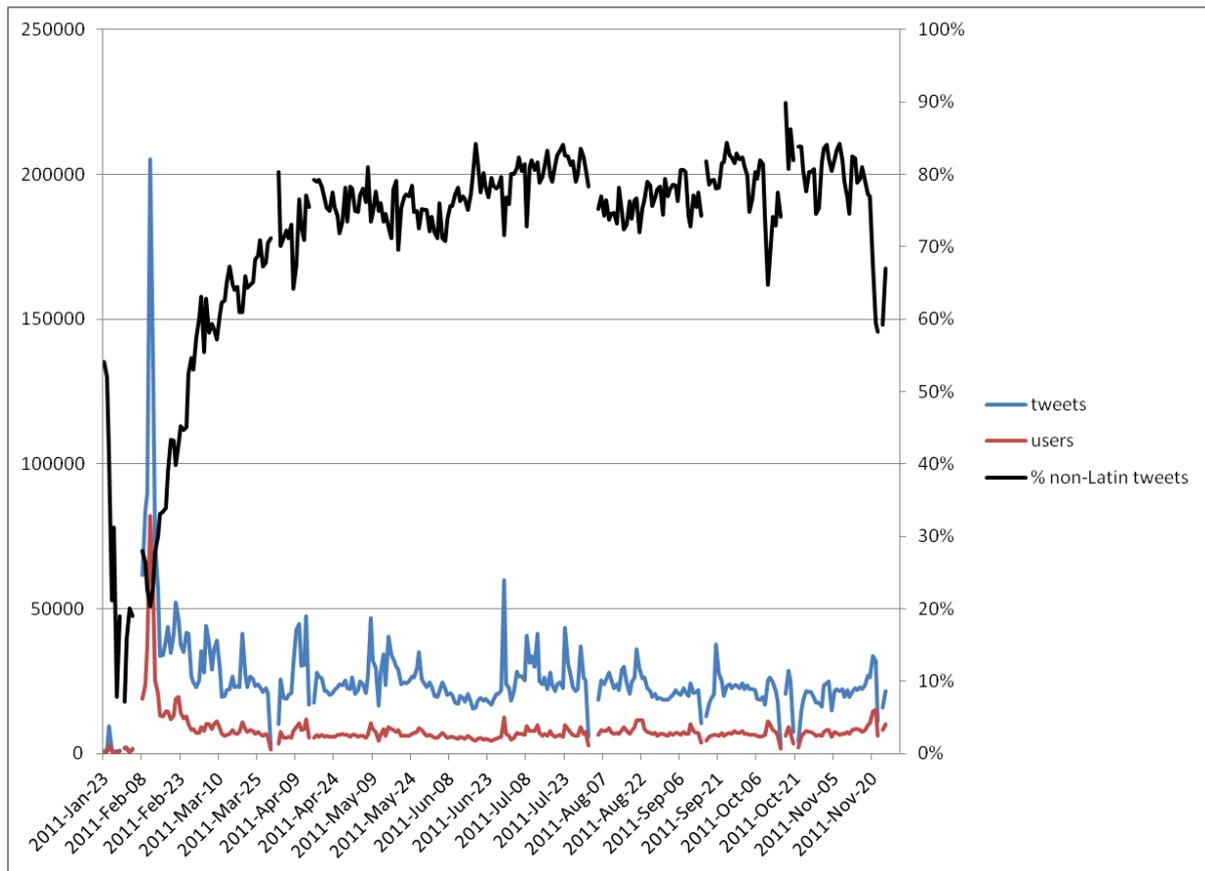
Figure 2. #egypt tweets and unique users per day, compared with daily contributions by different user groups.

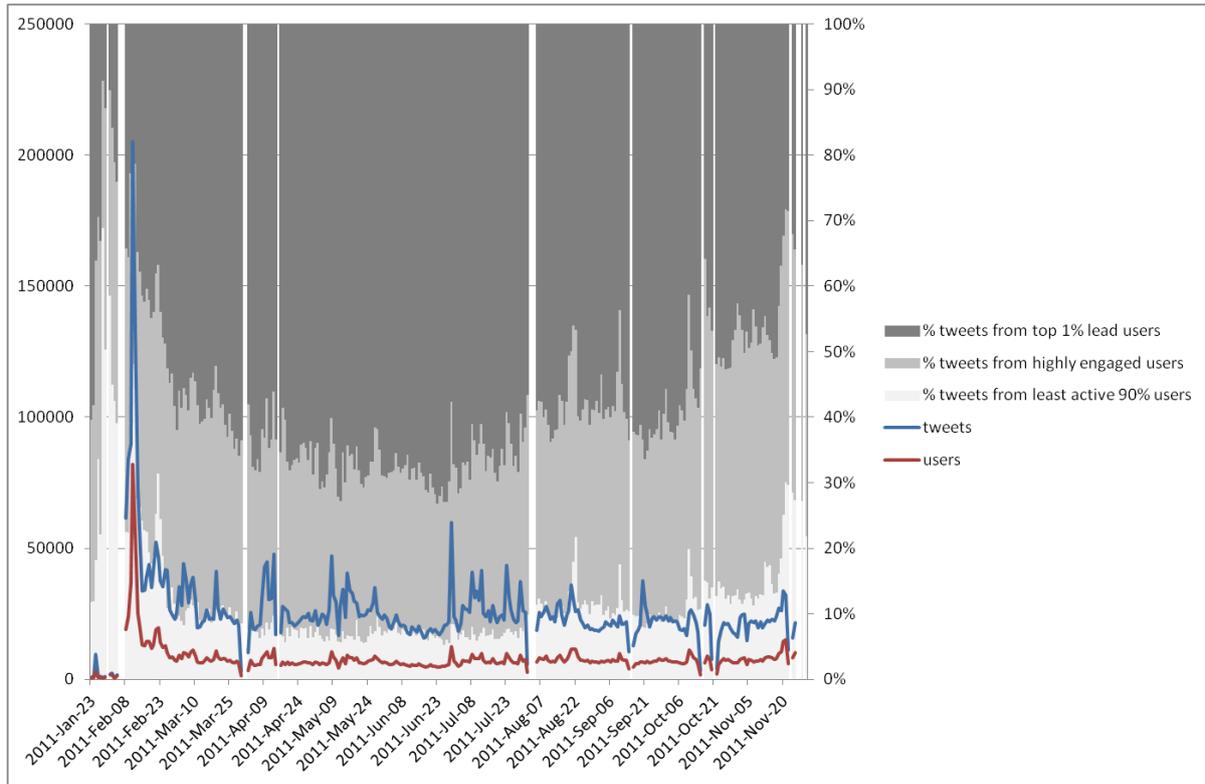*Figure 3.* Daily percentage of Latin tweets, compared with percentage of tweets from least active   users.
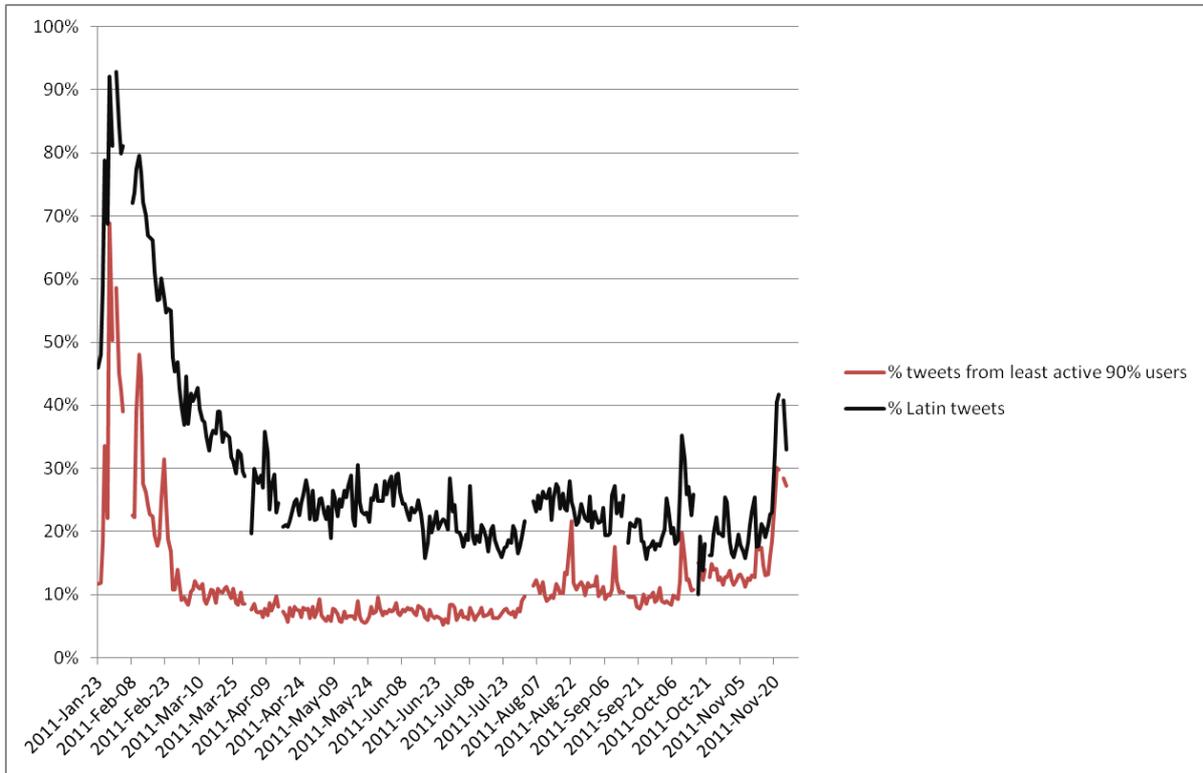
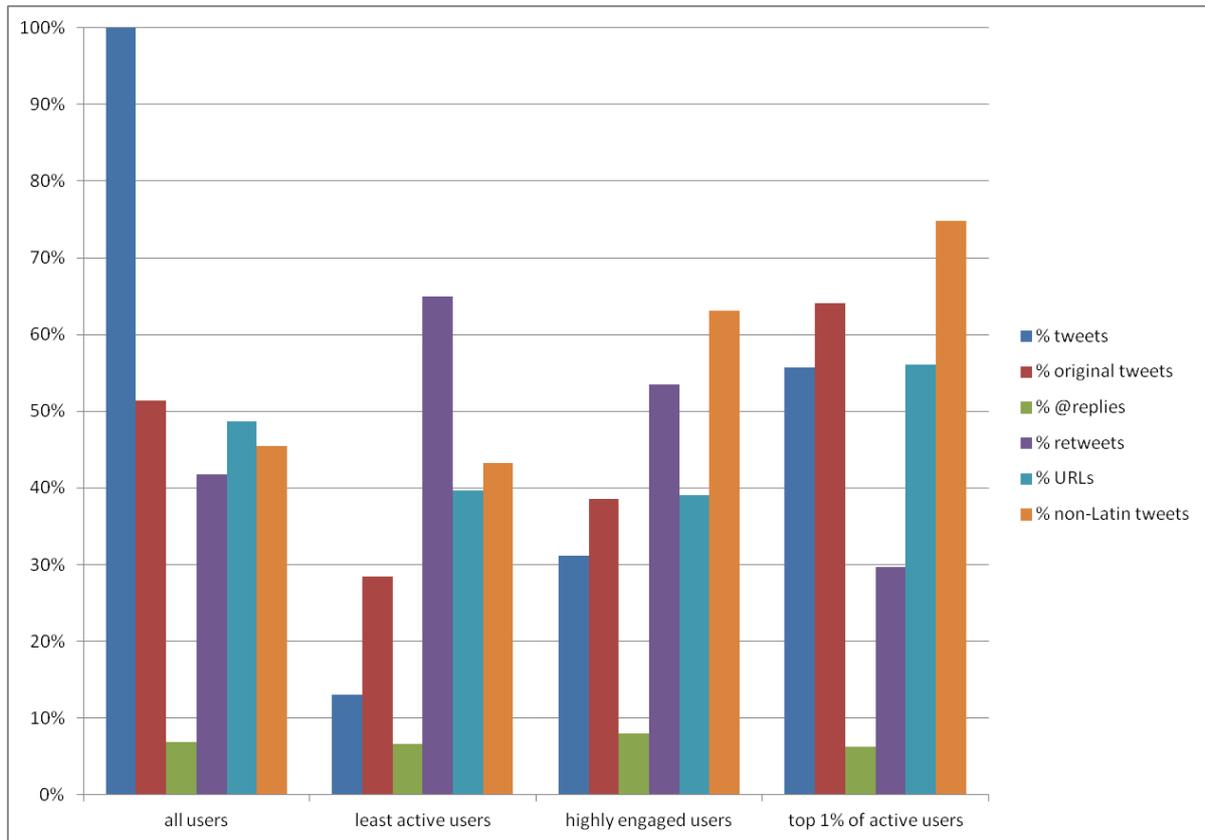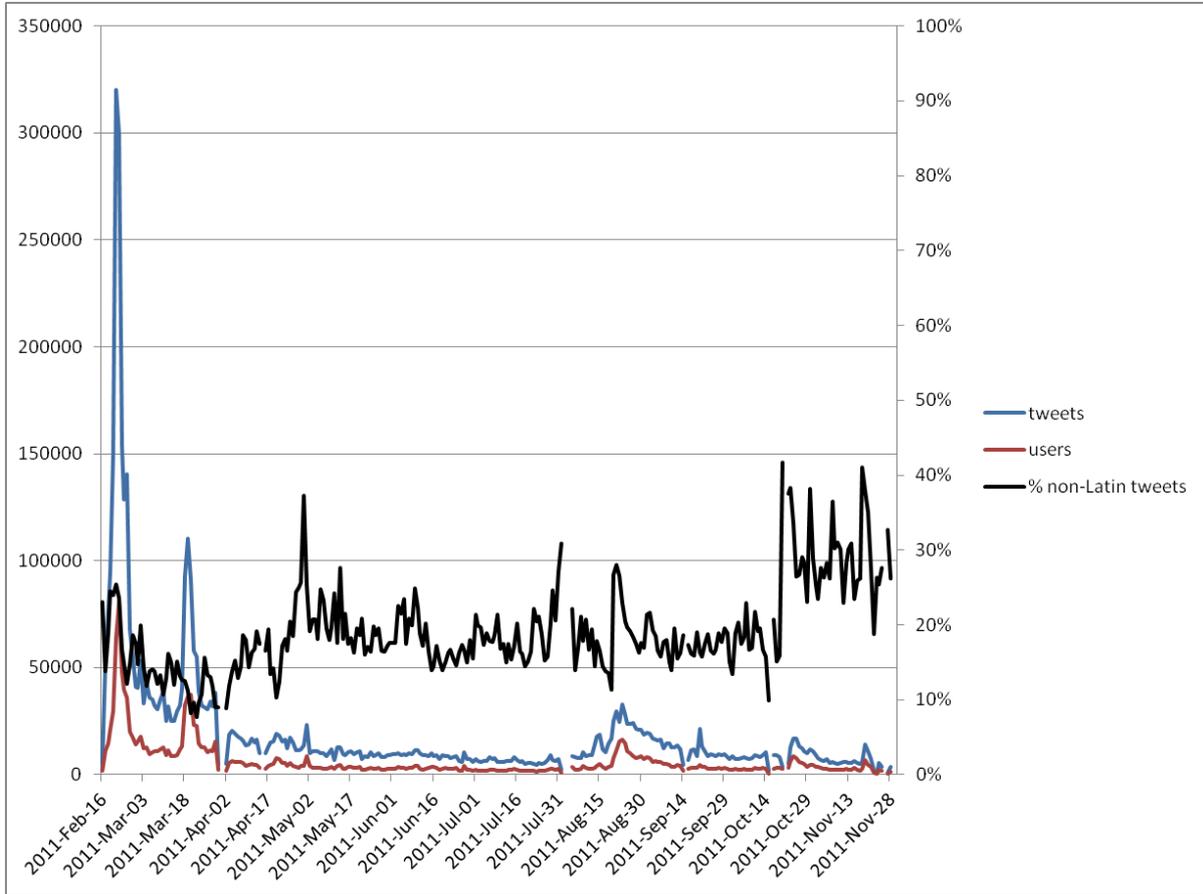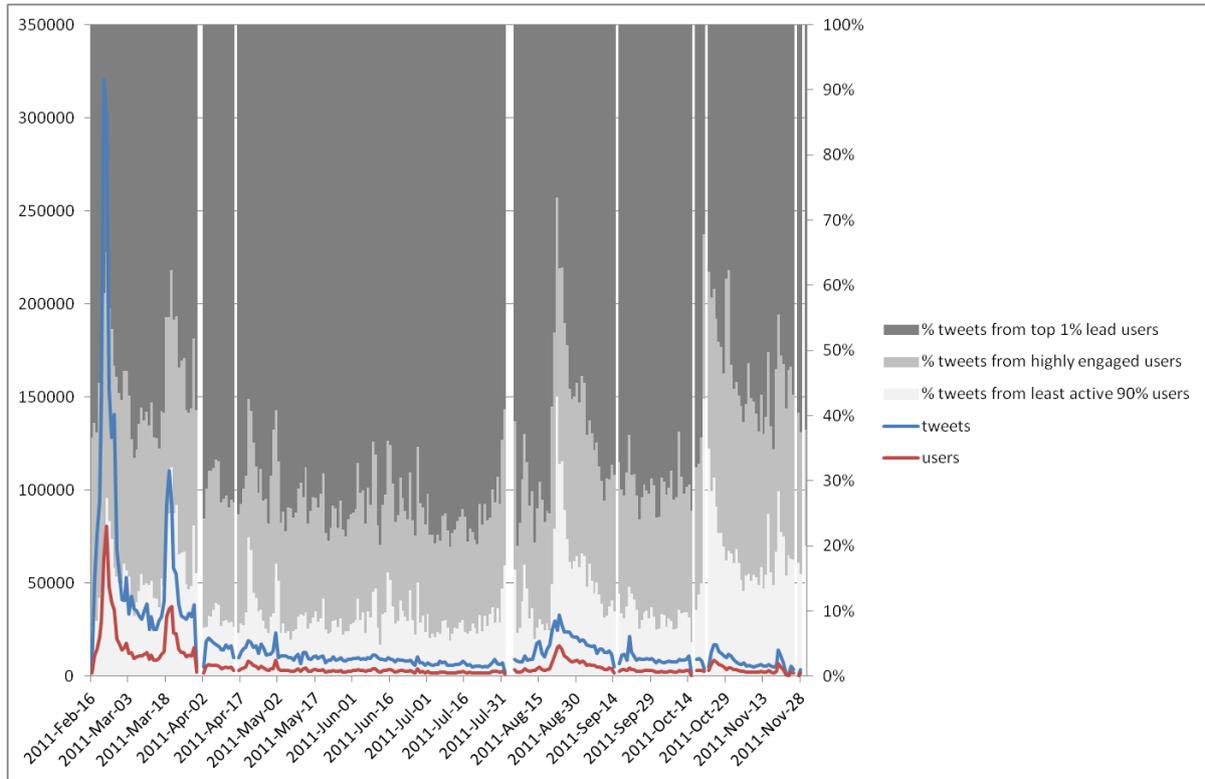*Figure 4.* #egypt contribution patterns across the different user groups.

Figure 5. #libya tweets and unique users per day, compared with daily percentage of non-Latin tweets.

Figure 6. #libya tweets and unique users per day, compared with daily contributions by different user groups.

Figure 7. #libya contribution patterns across the different user groups.
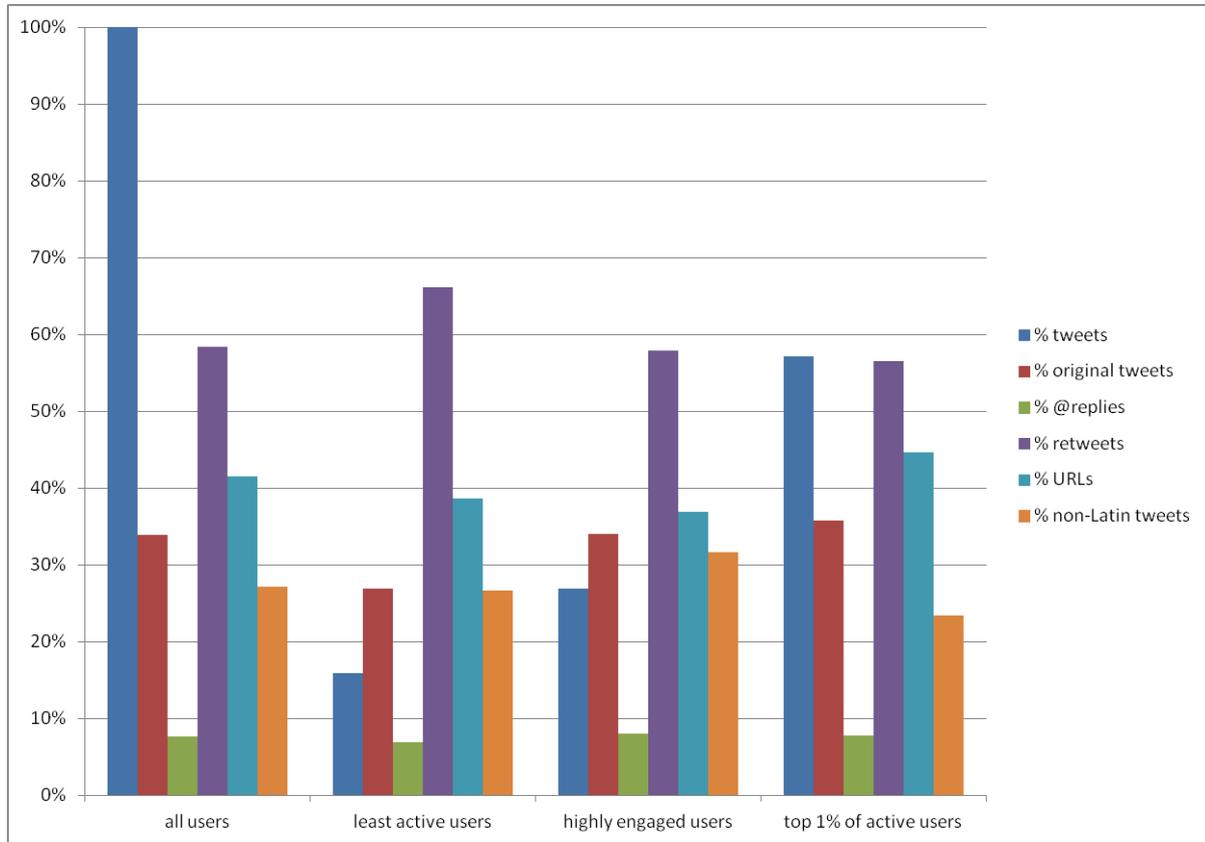
Figure 8a/b. #egypt language groups as percentage of total userbase, 1-28 Feb. and 15 June
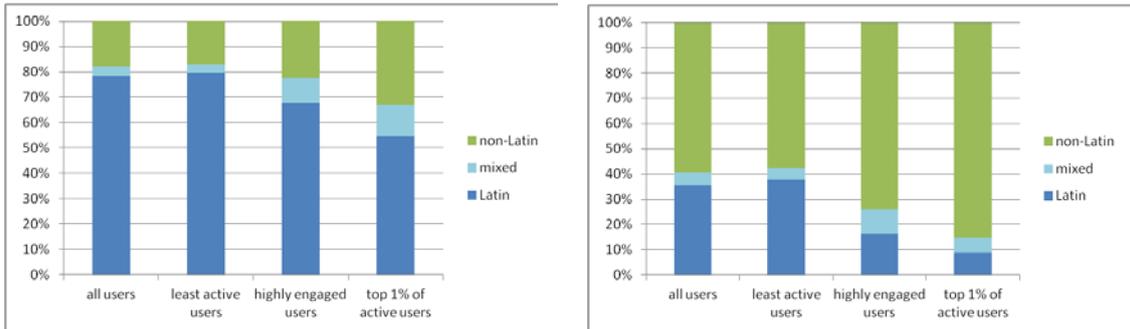
to 15 Sep.

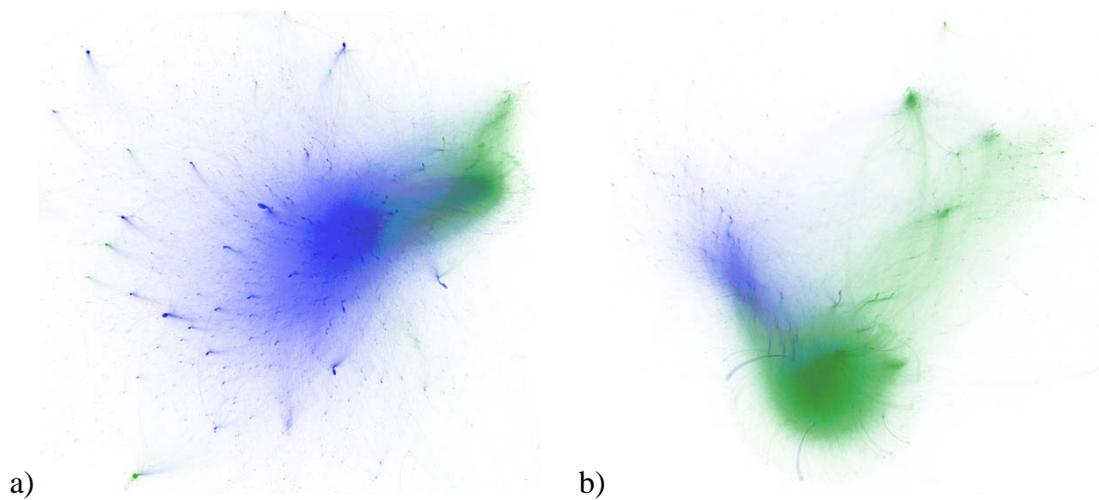*Figure 9a/b.* #egypt @reply/retweet networks, 1-28 Feb. and 15 June to 15 Sep.

a)   b)

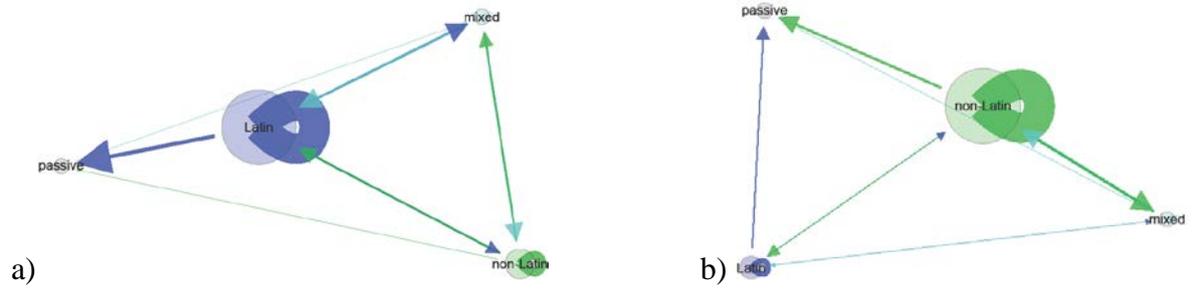*Figure 10a/b.* Aggregate #egypt @reply/retweet networks, 1-28 Feb. and 15 June to 15 Sep.

Figure 11a/b. #libya language groups as percentage of total userbase, 16 Feb. to 15 Mar. and
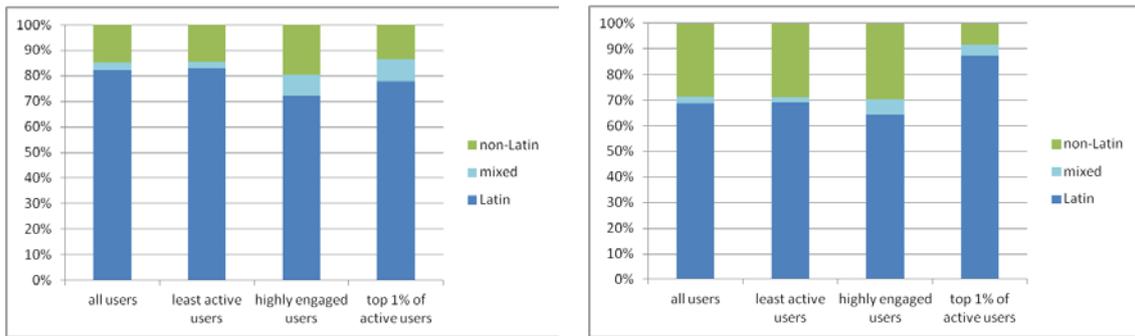
1 Aug. to 30 Sep.

Figure 12a/b. #libya @reply/retweet networks, 16 Feb. to 15 Mar. and 1 Aug. to 30 Sep.

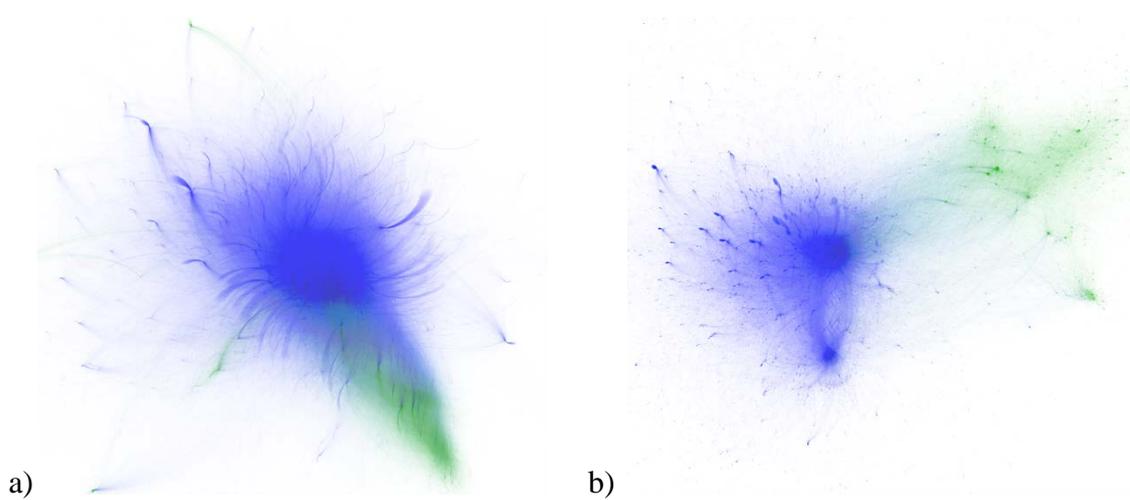a)   b)

Figure 13a/b. Aggregate #libya @reply/retweet networks, 16 Feb. to 15 Mar. and 1 Aug. to 30 Sep.