aultilingual MANCHi€S)EA

Report

2013



The University of Manchester

The contents of this report are the intellectual property of the authors. No part of this report may be circulated or reproduced without explicit permission from the authors, or from the School of Languages, Linguistics and Cultures at the University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom.

What Can Twitter tell us about the language diversity of Greater Manchester?

George Bailey

Joseph Goggins

Thomas Ingham

1 Introduction

1.1 Overview

In this paper we investigate the language diversity of Greater Manchester via the medium of social-networking site Twitter. With the number of Tweets sent daily reaching the hundreds of millions (Twitter 2012), this relatively new system of communication is no longer a novelty but, for many, a part of day-to-day life. The content of Tweets is fully in the public domain, including the exact GPS co-ordinates of where they were sent from, which means as linguists we can utilise Twitter almost as a corpus. The study should prove to be interesting given the lack of any previous literature on the subject of online-language use in Manchester. A similar project has already been carried out in London by academics Ed Manley and James Cheshire from UCL (Spatial Analysis 2012), so another potential point of interest is whether these two huge cities resemble or differ from each other in terms of linguistic diversity.

1.2 Methodology

Our analysis will be conducted on a database of over 28,000 Tweets, which have already been run through language-detection software to determine the language of their content. Each Tweet has been recorded with the username of who sent it, the date, time and location (in latitude and longitude) of its transmission, and what language Google's translation service believes it to be sent in. The word 'believes' is important, since this is a computer-automated process with no human intervention; the results are not necessarily correct, which is something we address in section 1.3.

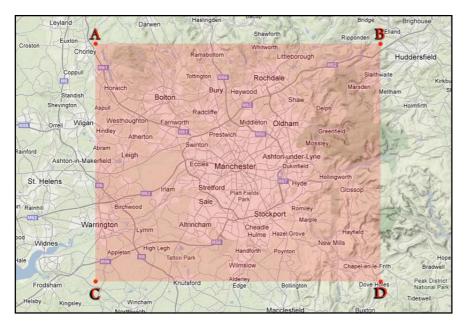
In section 2, we plan to display our results quantitatively with the aid of tables and graphs, but our findings will mainly be presented in the form of maps showing the geographical distribution of Tweets. This will be done using Google's Fusion Tables service, which places colour-coded points on a map using provided latitude and longitude values.

In addition to this general analysis of language trends across Manchester, section 2.2 sees us focus on particular areas on specific dates to see how language diversity changes when large-scale events take place in the city centre. This will be done using filters to narrow down our database to relevant co-ordinate ranges, but its effectiveness is dependent on how much data we have.

1.3 Data Clean-up

The sheer number of Tweets in our database (28,115) makes it impossible to manually go through each entry to ensure the language detection is accurate. We can, however, ensure that they were all in fact sent from Greater Manchester. To do this, we settled upon four co-ordinates that would form a 'catchment area' around the Greater Manchester area, with any tweets sent from outside these co-ordinates disregarded from our study. The co-ordinates are provided below, alongside Map (1) illustrating the area they cover.

Map (1): Greater Manchester Area



Point	Latitude	Longitude
Α		
	53.665575	-2.605716
В		
	53.665575	-1.875812
С		
	53.304806	-2.605716
D		
	53.304806	-1.875812

We have been fairly generous in deciding what comes under 'Greater Manchester' to ensure the number of Tweets we delete is kept to a minimum; this resulted in dismissing only 264 Tweets, one of which was actually sent from the English Channel, so this was clearly a necessary process.

In terms of language detection, as mentioned earlier there are too many Tweets to check the database thoroughly, but we did calculate an accuracy value. With the help of Google Translate as well as friends and fellow students who are bilingual, we took a random sample of 100 Tweets and decided one-by-one which had been correctly identified; this would give us an indication of the reliability of our database.

The language detection software was 48% accurate for this sample, but what became apparent was that many of the discrepancies came from Tweets classed as German and Tagalog, a language spoken in the Philippines. In fact, these two languages were reported as the two most frequent in our study, which came as somewhat of a surprise. Ignoring Tagalog Tweets, the overall accuracy rises to 62%; this is an easy decision to make, since all 23 'Tagalog' Tweets were incorrectly classified. We looked into German Tweets specifically and found only 2 of 50 German Tweets (4%) were actually German, so we also discounted these, increasing the accuracy further to 74%. This was a difficult decision to make since we are losing *some* correct data, but we felt it was in the best interests of the investigation.

The Tagalog issue was expected, since a similar problem was found in the London study mentioned earlier. We believe it is down to the language detection software misinterpreting strings of text like 'hahaha' as a reduplicative morpheme, a productive aspect of Tagalog's grammar. Indeed, almost all of these 'Tagalog' Tweets included strings like this. It isn't apparent why there is such inaccuracy with German, though. Lithuanian is another language with rampant misclassification, all 225 'Lithuanian' Tweets were actually English. Looking into this did shed some light on the language-detection software, since 69 of the 225 Tweets had the word 'hilarious' in them, and 111 had 'christmas'. It seems a single word can mislead the software, indicating that its predictions are based solely on lexis (why these words are deemed Lithuanian in the first place is unclear). This is further supported by other erroneous tweets, where a Tweet containing "Mourinho" (Portuguese football manager) led the software to deem the whole string Portuguese, and rather comically, many Tweets that contained "spaghetti" were classed as Italian.

1.4 Changes to Planned Research Schedule

When we came to investigate large-scale events in Manchester and their effect on language diversity, we encountered one major problem: a lack of data. We overestimated the Twitter activity at these events, so when we narrowed the data set down to, for example, the area of Manchester City's Etihad Stadium on the date of the Manchester football derby (8th of December), we were left with less than 20 Tweets to analyse. This is clearly insufficient to provide any meaningful findings, and the same problem hindered our research into other events, like the X-Factor final and music concerts at the 20,000-capacity Manchester Arena.

As a result of this, we decided to focus this aspect of our investigation solely on the christmas markets, which are said to bring in large numbers of visitors from around the UK and the rest of Europe (Christmas Markets 2012); our results are provided in section 2.3.

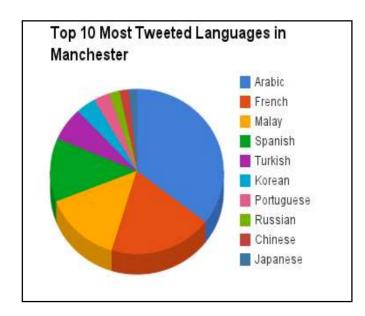
Another problem that became apparent during this process was the actual plotting of points on a map. In the fieldwork plan, we mentioned our interest in conducting a questionnaire to see if people's responses corresponded with our findings on Twitter, with the caveat that it may not be possible with our time limitations. This became reality as a result of the technical issues with the maps, which we had to dedicate a lot of time to when searching for software that allowed us to plot almost 20,000 points on a single map.

2 Investigation

2.1 General Analysis

Before we begin our analysis of language patterns and what determines concentrations of certain languages, we will offer a broad view of the general distribution in Manchester. The Tweets making up our database were sent in a two-month period from the dates of 15th November 2012 - 14th January 2013. Graph (1) below illustrates the 10 most Tweeted languages in Greater Manchester (excluding languages with less than 50% accuracy from a tested sample size).

Graph (1):



Language	No. of Tweets	Percent	Accuracy
Arabic	3470	31.43%	100%
French	1858	16.83%	100%
Malay	1397	12.66%	90%
Spanish	1218	11.03%	90%
Turkish	634	5.74%	60%
Korean	367	3.32%	100%
Portuguese	292	2.65%	90%
Russian	190	1.72%	100%
Chinese	170	1.54%	100%
Japanese	155	1.4%	85%

These results came as somewhat of a surprise, since only three of the ten languages are official languages of European countries, despite claims that Manchester attracts large numbers of European citizens (Guardian 2012). All ten, however, are from Eurasia. The fact that Arabic is the most Tweeted language (after English of course) comes as no surprise, given the high Asian population in Greater Manchester. Table (1) contains statistics from 2009 illustrating the ethnic populations by Greater Manchester region.

Table (1): Ethnic Population by Greater Manchester Area

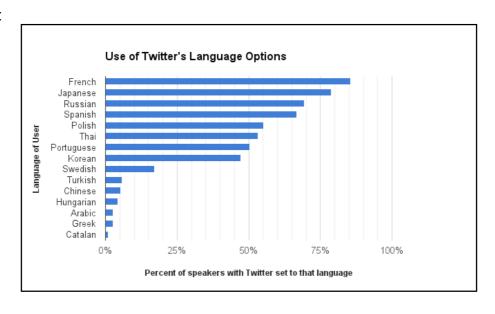
Local Authority	White British	Mixed	Asian	Black	Chinese
Oldham	80.48%	1.6%	13.57%	1.23%	0.32%
Manchester	69.93%	3.27%	11.2%	4.84%	1.76%
Rochdale	82.8%	1.42%	10.94%	0.98%	0.59%
Bolton	84.42%	1.43%	9.58%	1.24%	0.53%
Greater Manchester	83.57%	1.8%	7.14%	1.93%	0.77%
Tameside	88.86%	1.25%	5.76%	0.93%	0.42%
Trafford	82.63%	2.18%	5.43%	2.55%	1.07%
Bury	87.19%	1.64%	5.09%	1.2%	0.44%
Salford	86.45%	1.6%	3.29%	1.73%	0.62%
Stockport	89.07%	1.59%	3.21%	1.16%	0.6%
Wigan	94.68%	0.82%	1.31%	0.65%	0.29%

Reproduced from Greater Manchester Centre for Voluntary Organisation (2009)

When investigating the distribution of Arabic Tweets, we will refer back to this table to see if the likes of Oldham and Rochdale contain a higher proportion of Asian speakers than other Greater Manchester boroughs, as this table suggests.

One final topic we want to look at before our investigation into language diversity, is the languages that Twitter actually offers its service in. This is separate to the language of each Tweet's content, but one would expect the two to correlate. Since we can see what language each Tweet-er has their account set to, we investigated how many foreign-language speakers actually make use of Twitter's diverse language options. Graph (2) shows the percentage of speakers who have Twitter set to the same language they Tweet in.

Graph (2):



Language	No. of Tweets in Language	Percent of Tweets with language correspondence	
French	1,858	85.63%	
Japanese	155	78.71%	
Russian	190	69.47%	
Spanish	1,218	66.83%	
Polish	114	55.26%	
Thai	45	53.33%	
Portuguese	292	50.34%	
Korean	367	47.14%	
Swedish	117	17.09%	
Turkish	634	5.68%	
Chinese	170	5.29%	
Hungarian	23	4.35%	
Arabic	3,470 2.71%		
Greek	112 2.68%		
Catalan	105	0.95%	

It's interesting to note that of the top 3 most Tweeted languages, French and Spanish rank highly in terms of users utilising Twitter's language options, but Arabic is at the complete other end of the spectrum with a mere 2.71% of the 3,470 Arabic Tweets being sent from users with Arabic settings (hereafter referred to as 'correspondence'). The other 98% of Arabic Tweets are largely from users with English as their setting, which seems to indicate speakers of Arabic are confident in their bilingualism and possibly prefer Twitter's standard of left-to-right displays. Another point of interest is the stark contrast between Japanese with 78% correspondence and Chinese with just 5.29%, despite both languages having non-Latinate scripts (Korean places somewhere in the middle with 47.14%). Catalan sits at the bottom of the table with a 0.95% correlation; upon further inspection of these Tweets, a large majority are unsurprisingly from users browsing Twitter in Spanish, an official language in Catalonia that is actually used more often than Catalan itself in everyday speech (Statistical Institute of Catalonia 2008).

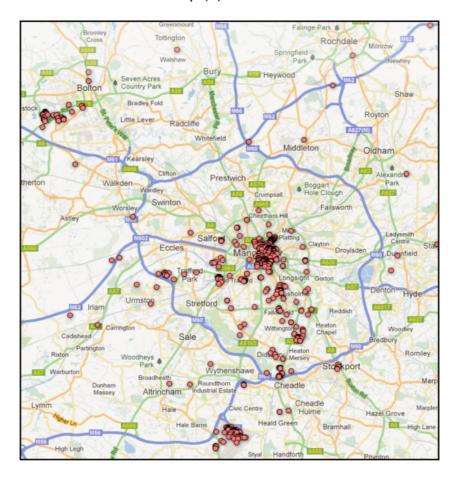
2.2 Language Trends

In this section of our investigation, we look into the geographical distribution of the top 10 most Tweeted languages in the Greater Manchester area.

2.2.1 Arabic

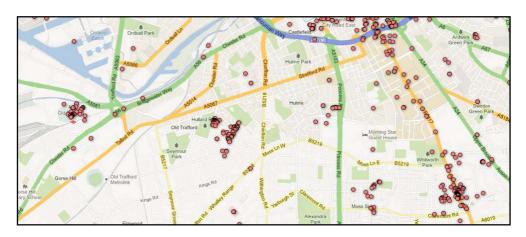
The most Tweeted language in Manchester is Arabic, and it is by no small margin with almost double the number of Tweets of the second most popular language, French. Map (2) below illustrates the regional distribution of these Arabic Tweets in Manchester.

Map (2): Arabic Tweets



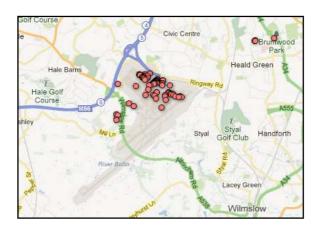
First of all we see a large grouping of Arabic Tweets in Bolton, which is to be expected since previous surveys (presented earlier in section 2.1) have found it to have the fourth largest Asian population in this region. What is surprising, given the figures previously presented in Table (1), is the lack of Arabic Tweets in Oldham and Rochdale; the Asian population of these towns either favour another language (although this isn't evident from our research), or more likely, they simply aren't Twitter users. Map (3) shows southern Manchester in more detail.

Map (3): Arabic Tweets



Here we can see a cluster on the far left, situated in Old Trafford (Manchester United's football stadium). There seem to be more Arabic Tweets in this particular area than any other foreign language, which is an interesting finding since it backs up reports of Manchester United's huge Asian following (Irrawaddy 2012). We also see a large group of Arabic Tweets around Wilmslow Road, which is perhaps the most predictable finding since this is where the 'Curry Mile' is located. Another point of interest is even further south, at Manchester Airport, as Map (4) shows.

Map (4): Arabic Tweets



Interestingly, most Tweets sent from the airport were Arabic, with only a few other languages represented. Whether this suggests a greater tendency for tourism or not is unfortunately beyond the scope of this investigation.

2.2.2 French

French is the second most Tweeted language from our data set with 1,858 Tweets sent from 227 users. Map (5) below shows the full distribution of all the French language Tweets.



Map (5): French Tweets

There seems to be a fairly widespread use of French, as we can see from the concentrations in Bolton and Oldham, as opposed to other languages which only feature prominently in the centre of Manchester. However, upon closer inspection it seems the Tweets from Oldham are largely from a single user. 53% of all French Tweets were sent by one user, *izii_priska* who accounted for a total of 987 Tweets. Map (6) below shows the location from which these were sent.



Map (6): izii_priska's French Tweets

We can see from the large concentration of these tweets sent from Oldham (throughout the entire two-month period our database covers), that this user is unlikely to be a visitor or tourist to Manchester.

2.2.3 Malay

In terms of the inner-city, Map (7) below shows dense usage of Malay in Tweets over three key areas; the city centre, Salford and in the vicinity of Rusholme.





There are enough to suggest that there may be a Malay-speaking community in the Rusholme area, and the relatively high number of tweets in the Oxford Road area, as shown by Map (8), could back that up. This suggests that there are a significant number of Malay-speaking students in Manchester.

Map (8): Malay Tweets



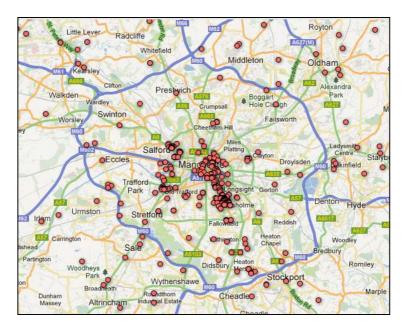
Map (9) below shows that numbers are reasonably high within the city centre and in Salford, but likely not high enough to conclude that there actual speech communities there, rather than just, for example, visitors. A significant number are concentrated in the area surrounding the University of Salford.

Map (9): Malay Tweets



Further afield, the distribution of Malay tweets is fairly regular, with evidence of sporadic Malay tweets in both the north and south of Greater Manchester, as illustrated by Map (10).

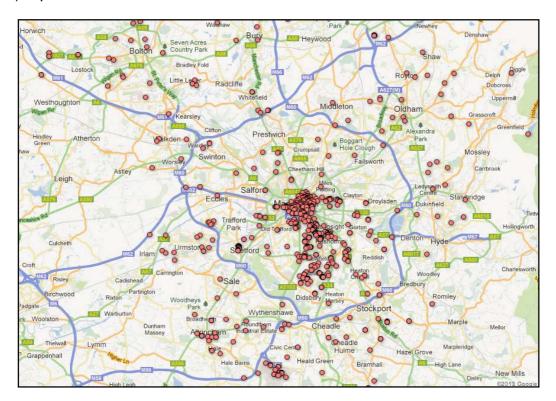
Map (10): Malay Tweets



2.2.4 Spanish

Spanish is the second most-Tweeted European language in our study, after French, and Map (11) below shows how these Tweets are distributed throughout Greater Manchester.

Map (11): Spanish Tweets



What this map makes apparent is just how wide-spread the use of Spanish is in Manchester and its surrounding regions. Although there is the expected concentration of speakers in the densely-populated city centre, there is a fairly equal spread throughout the likes of Bolton, Oldham, Prestwich, Stretford, Altrincham, and Didsbury. There also seem to be numerous Spanish speakers around Oxford Road, which we know is a hub for students both studying at university and living in its accommodation, indicating that a fairly high proportion of these Spanish speakers are possibly students.

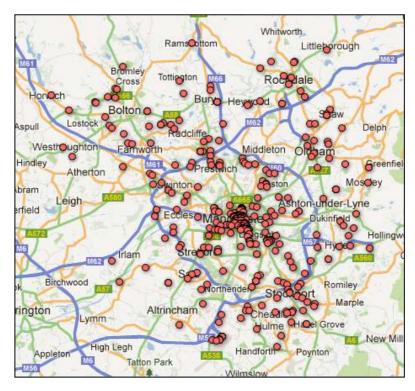
Of all our languages, Spanish Tweets are possibly the most representative of individual speaker distribution; a majority of speakers only contribute a couple of Tweets to the database, with the most prolific Tweeter, *iKiely*, only accounting for 119 of the 1,218 Spanish Tweets. Even these Tweets are all concentrated in a single area, Stretford, as Map (12) shows; this lends even more credibility to our overall Spanish map as a direct measure of the Spanish-speaking community.



Map (12): iKiely's Spanish Tweets

2.2.5 Turkish

Turkish is the fifth most Tweeted language from our data set, comprising of 634 Tweets, sent by a total of 233 different Tweeters. Map (13) shows the full distribution of Turkish Tweets throughout Greater Manchester.



Map (13): Turkish Tweets

It goes without saying that some people Tweet more often than others. Twitter user *BKorachi* sent a total of 89 Tweets during the period we investigated, the most by any Turkish Tweeter; these are shown in Map (14) below.

Map (14): BKorachi's Turkish Tweets



The second most productive tweeter was *ENBUYUKGSARAY*, with 87 Tweets. Map (15) shows the location of all 87 of his/her Tweets.

Ancoats Clayton Droylsder Manchestor Old Rd A63 A57(M) Abbey Hey Hulme Longsight Gorton Debdale Park wood Platt Fields Park cum-Hardy A6010 Reddish Fallowfield Reddish Technolo n-cum-Hardy Withington olf Club Didsbury Heaton Didsbury Northenden Stockport enshawe

Map (15): ENBUYUKGSARAY's Turkish Tweets

Other than these, most Tweeters in Turkish weren't particularly prolific. The two maps of the individual Tweeters show that their Tweets are largely form the same area and aren't very spread out, which means the distribution of points as shown in Map (13) above is fairly reflective of the actual distribution of Turkish speakers themselves, not just the Tweets they send.

2.2.6 Korean

The following map illustrates the distribution of Korean Tweets from Manchester.





What is striking about this display is how concentrated Korean speakers are, with all but a handful of Korean Tweets located in the very centre of the city.

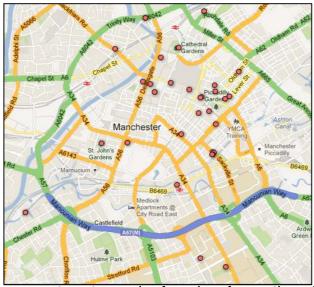
Furthermore, the total of 367 Korean Tweets in our database are actually only sent from 10 unique users, with two in particular, *Able_25* and *marielee88* contributing 274 between them. As such, the results are fairly misleading and not quite representative of the actual number of Korean speakers in Manchester. If we discount the Tweets of these two users, as we have done for Map (17) below, it's clear that there is no discernible Korean-speaking community in Manchester.



Map (17): Korean Tweets excluding those of Able_25 and marielee88

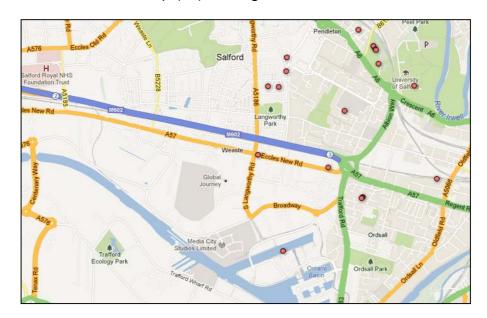
2.2.7 Portuguese

Incidences of tweets in Portuguese were found across the Greater Manchester area, with the only real noteworthy concentration found within the city centre, as shown below in Map (18).



Map (18): Portuguese Tweets

Whilst a handful of Portuguese tweets can be found as far north as Bury and as far south as Cheadle, there is no evidence from the map to suggest that there are any Portuguese communities of note within the Greater Manchester area. It is probably that the majority of Tweets generated are from visitors to the city rather than inhabitants, given that the only major concentration is based in the non-residential city centre. However, the relative dearth of Tweets in the Salford area (see Map (19) below) - which contains a number of Manchester's major tourist attractions - means Tweet-by-Tweet analysis may be necessary to ascertain whether the Tweets have been generated by visitors or Portuguese-speaking individuals living in the area.

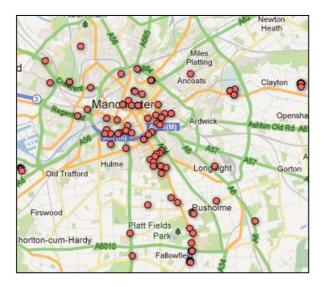


Map (19): Portuguese Tweets

2.2.8 Russian

The concentration of Russian tweets on the map is overwhelmingly focused within Manchester city centre, with a handful further south in the Rusholme and Fallowfield areas. Whilst those latter instances, found in areas heavily populated by students, could be from Russian speakers in Manchester for university, we can probably assume that the city centre tweets are likelier to be from either professional individuals or visitors, rather than actual speech communities.

Map (20): Russian Tweets



The almost total absence of Russian tweets in the wider, suburban areas of Greater Manchester is shown in Map (21); this suggests a lack of speakers further afield than the city centre, and certainly a lack of Russian-speaking communities, although a handful of tweets were detected in the Bolton area.

Bolton Country Park

Bradley Fold

Little Lever RadCliffe

Whitefield

Wardley

Clifton Prestwich

Wardley

Worsley

Swinton

Astley

Worsley

Swinton

Astley

Trafford

Park

Did Trafford

Denton

Woodheys

Partington

Woodheys

Park

Woodheys

Park

Staccoot

Stac

Map (21): Russian Tweets

2.2.9 Chinese

Chinese Tweets are highly concentrated in the city centre; interestingly, this mirrors our findings for Korean (and Japan, below). In total there are 170 Chinese Tweets in our database, but these are at least spread across a range of users, 44 in total. Map (22) highlights how these Tweets are geographically distributed.



Map (22): Chinese Tweets

There are, like with Arabic, a number of Tweets sent from Old Trafford, once again explained by Manchester United's large Asian fan base as mentioned earlier. However, we expected there to be a much greater number of Chinese Tweets around Portland Street where Manchester's China Town is located, an area with strong Chinese cultural identity and the third largest of its kind in Europe (Christiansen 2003). As Map (23) illustrates, there are only seven Chinese Tweets in this area of Manchester.





2.2.10 Japanese

Japanese was the tenth most Tweeted language from our database with 155 Tweets sent from 38 different users. Map (24) below shows the full distribution of these Japanese Tweets.



Map (24): Japanese Tweets

Although one must take into account the fact that it's a much smaller community of speakers, the majority of Japanese Tweets are sent from areas close to the city centre, with very few occurring north of Manchester.

Map (25) below shows the Tweets from user *Yuidon26*, the most prolific Japanese Tweeter sending a total of 33 Tweets.

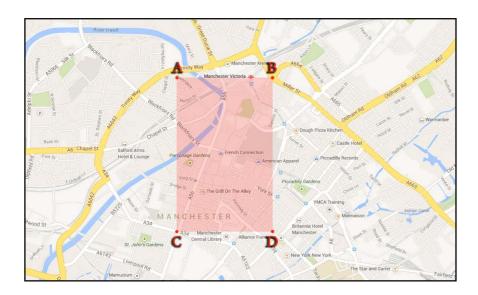


Map (25): Yuidon26's Japanese Tweets

Yuisdon26's Tweets are predominantly from the Altrincham area, not the city centre, which means that if we discount this user, the distribution of Japanese Tweets is even more concentrated in the centre of Manchester, a stark contrast with many other languages.

2.3 Manchester Christmas Markets

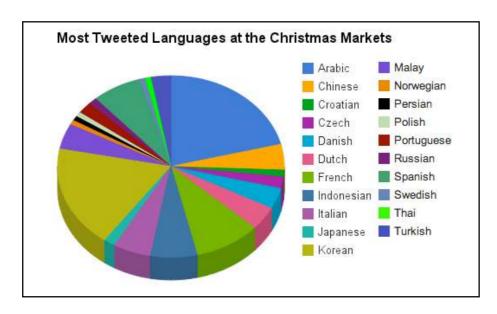
The second part of our investigation deals with the annual christmas markets in the centre of Manchester, a regular occurrence since 1998 that never fails to attract large numbers of visitors. Since it falls conveniently in the date range of our database (17th November - 23rd December 2012), it provides an interesting opportunity to see just how cross-cultural these kinds of events are. Using spreadsheet software that offers the functionality to sort and filter our data, we are left with 258 Tweets sent from the afore-mentioned date-range, and in the co-ordinate range illustrated by Map (26) below. This area covers all parts of the christmas markets: St. Ann's Square, Exchange Square, Albert Square, King Street, and others.



Point	Latitude	Longitude		
Α				
	53.487835	-2.249611		
В				
	53.487835	-2.239950		
С				
	53.478346	-2.249611		
D				
	53.478346	-2.239950		

Map (26): Christmas Market Area

Graph (3) below provides our findings from this investigation; it shows the most Tweeted languages from the christmas markets in Manchester.

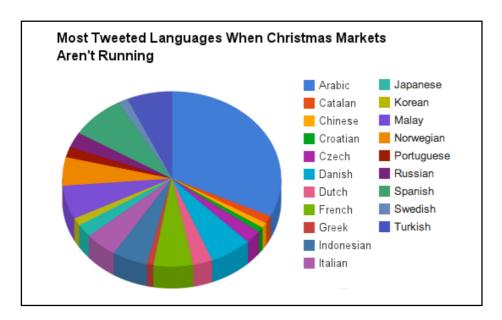


Graph (3):

Contrast this with the results shown in Graph (4), which illustrates the most Tweeted languages from the same area as shown in Map (2) above, but during the times when the christmas markets *weren't* running.

Graph (4):

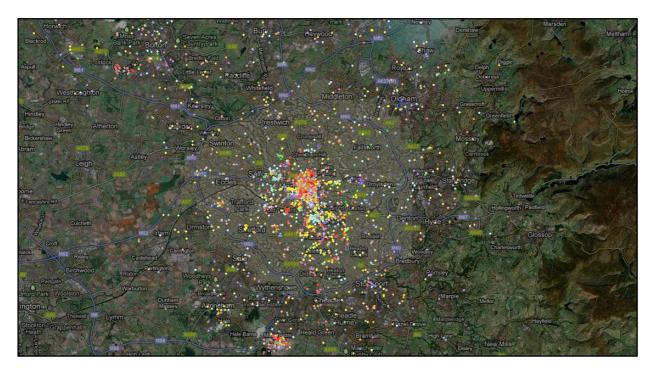
The two graphs seem to show little difference. There is a noticeably large majority of Arabic Tweets when the christmas markets aren't running, this is reduced during the christmas markets where we see a slightly more equal distribution between the languages, with a large increase in Korean and French. Other than this, there is little evidence to support our hypothesis that the christmas markets increase linguistic diversity. One possible explanation is that since Manchester is such a culturally-diverse city anyway, what should be multicultural events like the christmas markets actually do little to change the language distribution found here.



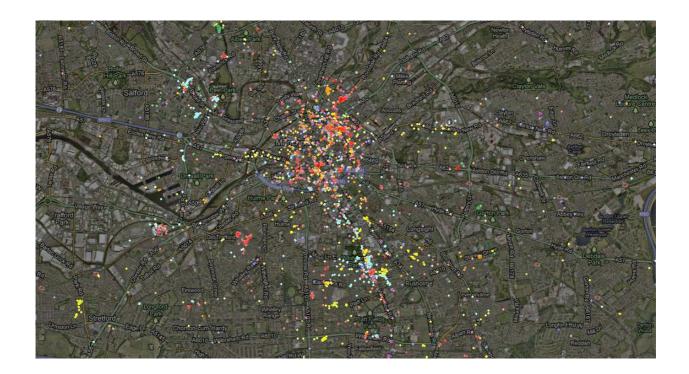
3 Comparison with London Study

As well as the individual language maps produced in Section 2, we managed to generate a single map of the top 10 most-Tweeted languages. This is presented below as Map (27).

Colour Key				
Arabic	French	Malay	Spanish	Turkish
Korean	Portuguese	Russian	Chinese	Japanese

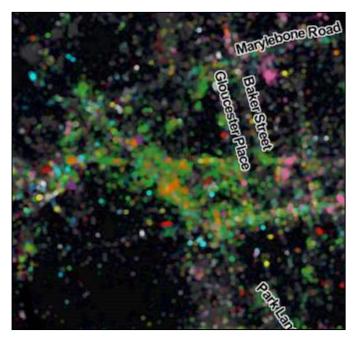


Map (27): The Top 10 Most-Tweeted Languages in Greater Manchester Map (28): The Top 10 Most-Tweeted Languages in Central Manchester



When comparing our paper with the London study that directly inspired it, we should take into consideration that the mapping software we have used is slightly less sophisticated. In hindsight, we realise that with more time, we might have been able to harness that software ourselves. The major difference between the two studies is that London is not only a much larger metropolitan area than Greater Manchester, but it is clearly far more cosmopolitan in terms of linguistic diversity, too. It's far easier to pick out clusters that indicate language communities, such as the green area indicating Arabic use as displayed in Map (29).

Map (29): Arabic Tweets in London (Spatial Analysis 2012)



In our study, as the analysis of each individual language shows in section 2, it was rare to see that kind of density and that kind of focus of one language in a particular area.

Of the nine top non-English languages looked at in the London study, seven were also present in our own analysis - Arabic, French, Malay, Spanish, Turkish, Russian and

Portuguese - suggesting that these minority languages are prevalent in significant numbers across the UK. The other two popular languages in London were German and Italian. As we discussed in section 1.3, German was seemingly rampant in our database, until we discovered the extreme inaccuracies it contained. It is unknown whether this was taken into account for the London study.

Also of note is the fact that the London study was able to take up to thirty thousand tweets into account for their results, whereas software inaccuracies that we encountered prevented us from doing likewise.

4 Conclusion

Our study into language use on Twitter has shed light on the multilingual status of Manchester, and is the first of its kind for this region. We have found a wide variety of languages used to different extents throughout the city, the likes of Arabic and French with thousands of Tweets over a two month period, and in contrast, rarer languages like Basque and Icelandic. In total there were 43 different languages in use throughout Manchester, revealing the true diversity of its community.

There *are* limitations to the use of Twitter as an index of language diversity, especially in terms of data size. Since our investigation relied upon widespread Twitter activity, there were certain aspects that proved problematic where we simply did not have enough data to work with. However, given a longer timeframe to work within, and a larger database size, this new area of linguistic research can prove to be even more fruitful and insightful.

5 Bibliography

Christiansen, Flemming (2003). *Chinatown, Europe: an exploration of overseas Chinese identity in the 1990s.* Routledge.

Christmas Markets. 2012. Available at: http://www.christmasmarkets.com/UK/manchester-christmas-market.html (Last accessed 22nd May 2013)

Greater Manchester Centre for Voluntary Organisation. 2009. Available at: http://www.gmcvo.org.uk/ethnic-population-greater-manchester-districts (Last accessed 24th April 2013)

Guardian. 2012. Available at: http://www.guardian.co.uk/uk/2012/dec/16/manchester-linguistic-ethnic-diversity (Last accessed 20th May 2013)

Irrawaddy. 2012. Available at: http://www.irrawaddy.org/archives/5324> (Last accessed 2nd May 2013)

Spatial Analysis. 2012. Available at: http://spatialanalysis.co.uk/2012/10/londons-twitter-languages/> (Last accessed 17th May 2013)

Statistical Institute of Catalonia. 2008. Available at: http://www.idescat.cat/dequavi/?TC=444&V0=15&V1=2 (Last accessed 30th April 2013)

Twitter. 2012. Available at: http://blog.twitter.com/2012/03/twitter-turns-six.html (Last accessed 17th April 2013)

6 Appendix

All maps created for this study are available to view here:

Map of Top 10 Languages: http://www.gpsvisualizer.com/display/data/1369388964-26649-82.28.120.88.html

Arabic Map:

https://www.google.com/fusiontables/DataSource?docid=1gaES5BpSSh28BFTYMwBiBg-JefKAE8TZddzJipc#map:id=3

French Map:

https://www.google.com/fusiontables/data?docid=15gklubEDCYQkLWiFn 8L6RGeHkAk5t vObUlcQHo#map:id=3

Malay Map:

https://www.google.com/fusiontables/DataSource?docid=1OOFzNbRc7VagEV7GLGvrVDiaDkzS_2EVrKl86Cc#map:id=3

Spanish Map:

https://www.google.com/fusiontables/DataSource?docid=12wbHLuv6tUv8AmbYSZT1KF9uKWRUs5HuPGOQgKg#map:id=3

Turkish Map:

https://www.google.com/fusiontables/DataSource?docid=1TiAAz2H5rEk6ChximDt RB5xR 8R7 U8sEcjMELg#map:id=3

Korean Map: https://www.google.com/fusiontables/DataSource?docid=159F10-DcBVXrE7NGuimcT4tvaxHaO4Eyol2Es04#map:id=3

Portuguese Map:

https://www.google.com/fusiontables/DataSource?docid=1aTf2l_LGQ3quURWim8nv96ibLyDlYKgnG2R60AU#map:id=3

Russian Map:

https://www.google.com/fusiontables/DataSource?docid=1kJblQFJML7FprOx-PWM5pAmM5k2zJpRDGN6QdUQ#map:id=3

Chinese Map:

https://www.google.com/fusiontables/DataSource?docid=1ALYorhHJHAm4Gt28kP4J1xvawwSjpGmGj1PmOqo#map:id=3

Japanese Map: https://www.google.com/fusiontables/DataSource?docid=1Si520UaNis-UY_UGXoMJHyxadpujpM3JpF94wuo#map:id=3