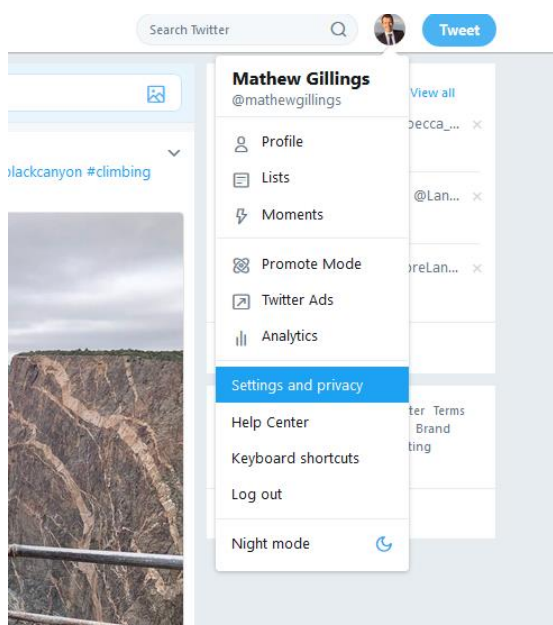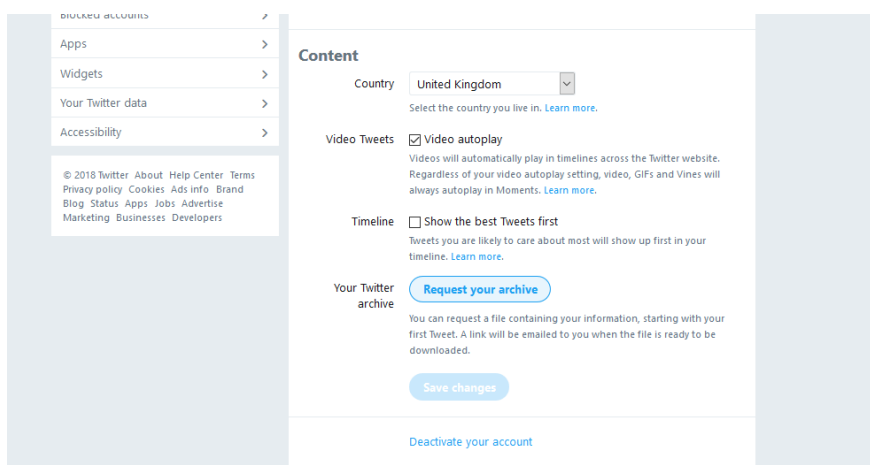# BNC 2014 – Tweet collection

#BNC2014

To collect tweets for the Written BNC2014 we rely on the generosity of the British public to give us access to their published tweets to incorporate them into the corpus. We are asking the public to provide us with their Twitter archives containing tweets sent between 2014 and 2018. Tweets should be unedited, but you are welcome to anonymise them. We will carry out our own anonymization procedures, but if you would like to remove any personal information yourself, please replace it with [anon] before submission.

Note: Tweets should be **written by British native speakers**, people who grew up in the UK and use English as their first language, **and unedited.**

Step 1: Log into your Twitter account and visit the "Settings and Privacy page".

Step 2: Click on "Request your archive". You will receive an email containing a link where you can download an archive of all your tweets since you set up your account.
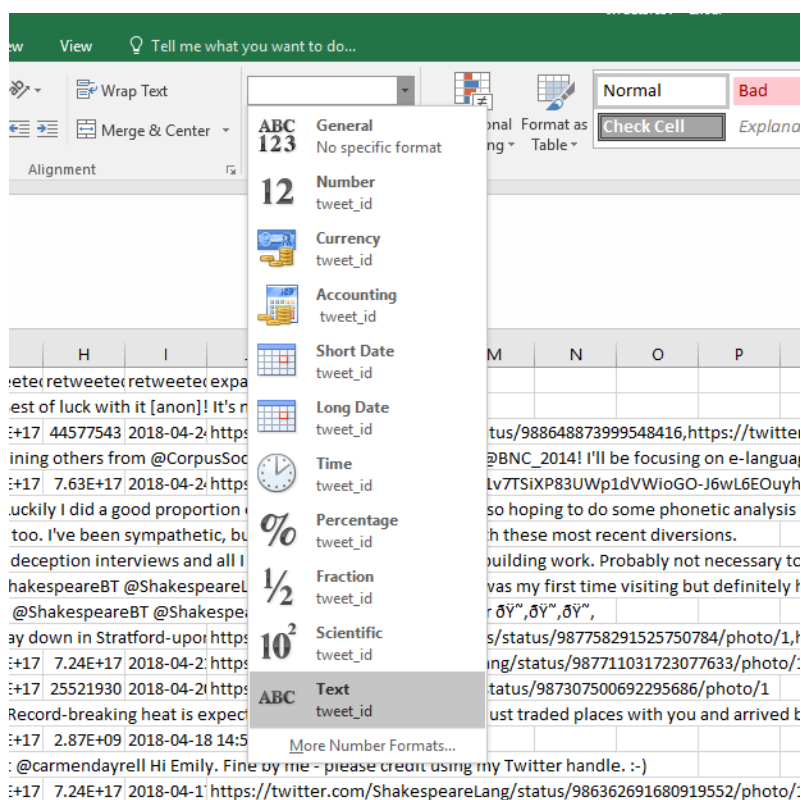

Step 3: When the download is complete, open the file entitled "tweets.csv". Tweets should be unedited, but you are welcome to anonymise them. We will carry out our own anonymization procedures, but if you would

like to remove any personal information yourself, please replace it with [anon] before submission. You can edit the spreadsheet by highlighting all the text > clicking the dropdown box on the toolbar which reads "General" > selecting "text". We understand that this may be time consuming, especially if you have a lot of tweets, so we are able to automatically remove all Twitter handles, URLs, and retweets of other users on your behalf. However, we would strongly encourage you to sift through your tweets yourself to ensure you are happy for everything you submit to be included within the corpus.

Step 4: Visit the "SMS collection" page through the CASS website and upload your .csv file through the Google form: http://cass.lancs.ac.uk/?page_id=2462