**ACADEMIC PROJECT SUBMISSION DETAILS:**

<table>
<thead>
<tr>
<th>Supervisor/s</th>
<th>Andreea Calude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Stance in Social Media - the language of taking sides during a pandemic</td>
</tr>
<tr>
<td>Field:</td>
<td>linguistics, sociolinguistics, (social) media, discourse analysis</td>
</tr>
<tr>
<td>Division/School:</td>
<td>ALPSS - Division of Arts, Law, Psychology and Social Sciences</td>
</tr>
</tbody>
</table>

**EXPECTED OUTCOMES:**

1. **INCREASE UNDERSTANDING OF STANCE.** It is anticipated that following the manual coding, we will have a working set of tweets from around the English-speaking world (at least) coded for stance (for or against current measures taken) and a diverse list of expressions of stance (ADJ, N, etc).
2. **NEW METHOD FOR IDENTIFYING STANCE MARKERS.** I hope to adapt the method of semantic vector space modelling to stance markers and to propose a method for identifying key markers of stance in a corpus.
3. **SOME RESULTS. A JOURNAL ARTICLE.** Following the manual coding of the tweets and implementing the proposed method for identifying stance markers, it is anticipated that the results will be summarised in a written journal article to be submitted for publication in a peer-reviewed journal.

**STUDENT TASKS:**

2. **CORPUS TAGGING** An agreed portion of the Tweets collected which make up the corpus will be manually coded for a number of features: stance (for or against lockdown or government measures), key words signalling stance of the overall tweet, other relevant features to the expression of stance.
3. **CAPTURING STANCE** Each key word identified is added to a database of what we will term bag of stance words. We thus end up with a bag of stance words that are for and a bag of stance words that are against (in this case) lockdown measures. (under guidance)
4. **TRENDS OVER TIME** Separately from Task 3, numbers of cases and deaths will be extracted from online sources. Ratios of the numbers of for and against tweets will be calculated from the manually-coded tweets, and matched against online stats. (under guidance)
5. **PAPER** Time permitting, we will draft up a report of the findings which would serve as a starting point for an article in a peer-reviewed journal (under guidance).
REQUIRED SKILLS:

1. excellent language skills and English language ability, ideally be native or near-native
2. willingness to work independently in a completely new field of inquiry, and ability to contribute own ideas if called upon
3. good computer literacy and willingness to learn how to use new software, ideally not complete adversity to numbers and statistics (interest in these would be a plus)
4. background (or at least some interest) in linguistics and ideally, discourse analysis, basic knowledge of French or German a bonus (but not required)
5. ability to pay attention to detail and maintain good accuracy and focus in coding data

PROJECT ABSTRACT:

The past year has seen the world battle one of the most aggressive pandemics of modern time, Covid19. There are two important characteristics which set the context of this particular pandemic apart from all others before it: (1) unprecedented access to daily information of what is going on almost anywhere in the world (from statistics relating the number of total infections and deaths to political decisions), and (2) the public’s ability to freely voice opinion (stance) toward the situation. In particular, millions of Tweets have been posted during the months that Covid19 took over the world news, as countries went in and out of lockdown, as politicians made good and bad decisions, and as people got tested, became infected or recovered. The context created by these events provides a fruitful and rich opportunity for the study of the linguistic means which Tweeters use to express polarised views on the events unfolding around them. Scanning just two examples of such Tweets provides insight into the diverse stylistic and grammatical ways employed to support or oppose measures and policies introduced by local governments and politicians or the acts of fellow community members.

Example Tweet 1 (4/30/2020 11:47:00, stance: opposing lockdown) The road to hell is paved with aspirations of perfect safety. #wedemandsafety #SafetyandHealth #SafetyFirst #SaferAtHome

Example Tweet 2 (4/29/2020 19:55:00, stance: supporting lockdown) You’re so vain you prolly think this virus won't get you!!!! #staytheFUCKhome

In example 1, the tone is subtle and the correct identification of stance requires a holistic reading of the entire sentence. The main clue comes from the word hell (a noun) and the exaggeration expressed by the noun phrase perfect safety. The hashtags are not telling of the wider message and in this tweet, they serve the purpose of community-affiliation aimed at making the tweet discoverable by its topic. In contrast, the tweet in 2 figures the adjective vain and the hashtag staytheFUCKhome which forcefully relay the writer’s supportive stance of the lockdown.

The linguistic expression of stance has implications for how meaning is construed and represented, and it gives insight into the language used to convey public opinion and belief. Stance is notoriously complex to capture because of the diverse means by which it can be expressed. This project seeks to categorize ways in which stance is signalled in Tweets related to Covid19 and to track public opinion over time, relating its changing pulse to statistics of total number of cases/deaths.

Over the past months, we have collected Tweets just like the ones above, on the topic of Covid19 with the help of computational tools by following various relevant hashtags in a number of countries around the world (#CovidNZ, #CovidUK, #CovidUS, #CovidDE, #CovidRO, #CovidFR). The tweets span the entire period of visibility of Covid19 on social media (dating back to November 2019) and the process of collection is ongoing. The final corpus will consist in at least 300,000 Tweets. The project will require manual tagging (of at least some) of the data, with the help of corpus linguistics software. It is anticipated that the ensuing analysis will inform our understanding of stance and how it is signalled.