

## Corpus Approaches to Discourse Analysis: Persuasion Strategies in a Large Corpus of #Covid19NZ Tweets

Social media platforms afford a window into the fast-paced and ever-changing world of public opinion. The language used in such outlets allows us to interrogate strategies used for constructing different discourse perspectives. In this paper, we investigate links between grammatical constructions (directives) and politeness strategies to express speaker stance in a large corpus of Twitter posts related to Covid19 in New Zealand. The overall aims of the project are to document an under-studied construction (directives) in a sizeable naturally occurring dataset, and to probe the extent to which we see patterns matching linguistic usage with discursive goals.

Twitter provides the opportunity to collect diachronic data directly from its API by either searching for keywords, hashtags of interest, or usernames of Tweeters. For this project, we collected Tweets posted between 22 February 2020 and 10 November 2020 containing the hashtag #Covid19NZ and variants of it. Excluded from the data were any Tweets which only consisted of a retweeted message, hashtag, link or other multi-modal feature without additional text. In total, the Twitter API contained 40,243 Tweets meeting our criteria. Our analysis centres on three key linguistic properties (among others), namely presence/absence of directive, politeness strategy and Tweeter stance. Here we report on the part of the data of most interest to us, namely Tweets that contained directives. We manually coded 1,001 Tweets from our corpus of 40,243 Tweets and within this body of data, we identified 759 distinct directive clauses, which occurred in 499 Tweets (some of these clauses occurred in the same Tweet).

One immediate problem we encountered was that identifying whether a clause contains a directive or not is not a straightforward matter because some directives are less “direct” than others. While English grammars, such as Huddleston & Pullum (2002) provide descriptions of directive constructions from a grammatical perspective, these descriptions are not sufficiently detailed to reliably identify (more) nuanced constructions which nevertheless encompass a directive discourse function, as illustrated below from our Twitter data:

- (1) And with that... The gates of Middle Earth are now officially closed. *If you're not a hobbit you can fuck right off* #Covid19NZ
- (2) *Would Jacinda care to explain this screenshot that is currently circulating NZ right at this moment? NZ moving to a Level 4 would be critical information wouldn't you agree?* #nzlockdown #COVID19nz #WhatsUpDoc?
- (3) *It's 8.35am on Friday 3 April 2020, and New Zealand's Minister of Health is David Clark* #nzpol #COVID19nz

The Tweet in (1) contains a strong expletive verb (*fuck off*) accompanied by the modal *can*, which adds an element of ambiguity between permission (e.g. *you can go now*) and directive (e.g. *can you pass the salt?*). Example (2) is also ambiguous because of the modal interrogative *wouldn't*. Both examples are coded as directives in our data; (1) is a declarative directive and (2) an interrogative directive. Examples such as the one in (3) are more difficult to classify because they are considerably more ambiguous. Though external context – mass outrage and calls for David Clark's resignation following his breach of government mandated Covid-19 measures – leads to the interpretation of this Tweet as a demand for David Clark's resignation or firing, this could also be understood as a purely declarative statement. Examples of this nature were therefore excluded from our analysis. For the 759 directive constructions identified, we coded two separate parameters: general clause type (imperative directive, declarative directive, interrogative directive) and specific grammatical construction (cohortive *let's*, closed *let*, auxiliary *do*, order/demand, plea/entreaty, advice/recommendation, warning), following Huddleston & Pullum (2002, pp. 929-945).

The second factor of interest was Tweeter stance. We discovered that Tweets fall into three delineated stance groups: “in agreement with current government measures”, “against current government measures”, “in general agreement with measures but calling for stronger measures”. Additionally, some Tweets containing directives were neutral in stance and some few Tweets were impossible to classify for stance.

The third key factor was politeness strategy. Because directives are by their very nature threatening the face of the implied or explicit addressee, they can often involve mitigation. We coded our body of Tweets

for various politeness strategies, following Huang (2007, p. 117), namely, bald on record, positive politeness, negative politeness, and off record.

Additionally, where present, we indicated the use of vocative noun phrases, subject type of the directive clause (explicit or implicit), intended addressee (individuated or non-individuated), imperative hashtags, discourse markers and Māori loanwords.

Our results show that Tweeters who either strongly agree or strongly disagree with government measures tend to flaunt politeness and go bald on record, much more than those who would like to see stronger measures in place. In particular, those who uphold the status quo (i.e. who are in favour of government measures) are the most forceful in their directives. Tweeters who want to see stronger measures implemented are the most tentative, going off record and sometimes opting for negative politeness strategies. As regards directive types used, we find that there is a strong preference for all groups to use imperative or declarative constructions, and this tendency is most pronounced for the group favouring government measures. *Let's* imperatives are disfavoured by most groups. There is also a high preference for imperative hashtags among Tweeters who uphold the status quo. The use of Māori loans is prevalent across all three groups, with those in favour of government measures leading the trend.

In sum, our data suggests that the language of persuasion on Twitter is often laden with various specific features not commonly found elsewhere (such as vocative noun phrases, imperative hashtags) and that directives are typically given in a strongly commanding and emotionally charged manner, with little mitigation. The emotional contours found in this type of language allow a unique opportunity to analyse large authentic datasets whose discourse function is to induce action – data which has previously been impossible to obtain.