

PATRICIA RONAN
Tweeting with Trump

1. Introduction

While it is known that spoken and written language show clearly identifiable differences (e.g. Biber 1988), which are largely attributable to ease of processing, claiming that spoken and written language are absolutely and clearly distinct is an oversimplification. Instead, we find features that are considered typical of both written and oral language in different genres of communication. Koch and Oesterreicher (1985; 1985/2012) and Koch (1999) have shown that beyond a simple division between written and spoken, other factors determine what linguistic features are found in any text, and that these factors can be divided into such expressing – and creating – distance and into such creating – and expressing – proximity and immediacy. Partly, these features coincide with the written/spoken dichotomy, but other features like familiarity of the participants, co-location of the participants of the speech event and public or private setting play a large role.

One type of language which has repeatedly been said to comprise both features of written and spoken language is new media language (e.g. Crystal 2009; Jonsson 2015). Amongst these media, Twitter, and especially the political use of language on Twitter, has very much come to the fore of the public's consciousness. This is due not least to the use of this medium by the 45th President of the United States, Donald Trump. The tweets from his account receive world-wide attention both because of their content and their style. Though the tweets from Trump's Twitter account, @realDonaldTrump, may in fact be multi-authored (Grieve 2017), distinct stylistic features emerge. For any political message, the use of language is particularly consequential as language is the main element of political persuasion (Partington and Taylor 2019). When considering the language of distance and the language of immediacy, investigating the language used on social media is particularly rewarding. Here, the investigation of the political language of, arguably, the most powerful man in the world, the President of the United States of America, is instructive. As Lockhart argues, "[t]he Commander-in-Chief's use of Language on Twitter was a rhetoric event that was urgent" (2019, 2). Not only do the tweets receive world-wide attention, in many cases they also constitute speech acts that introduce new realities (Searle 1976), such as firing and hiring staff publicly.

The current study pursues two research questions: first, how Donald Trump's tweets situate within the dimensions of language of distance or immediacy. By this, we determine just how aloof, or not, the Commander-in-Chief of the United States is on Twitter. Second, we want to find out whether the linguistic features used in the tweets are typical of the features that have been identified to be significantly different in Computer-Mediated Communication (CMC) from written and spoken genres. In order to answer these research questions, 200 tweets, the tweets of one week in the summer of 2019, are taken from the Twitter account @realDonaldTrump via the Trump Twitter Archive, <www.trumptwitterarchive.com>. These tweets are analyzed manually for

those features of Biber's (1988) Multidimensional Analysis (MDA) which Jonsson (2015) has identified as significantly different from written and spoken genres in CMC. While the data base of my current study is still small, results gleaned from it already allow us some interesting insights into Trump's Twitter use.

After this short introduction, the typical features of written and spoken language, and the language of distance and immediacy, are presented. Then data and methodology of the current study are explained before it is shown how these features are used in the corpus of tweets.

2. Written or Oral Language Use and Twitter

Studies on the language of CMC, or on messaging language respectively, have repeatedly pointed out that such modes of communication can contain both features that are commonly associated with spoken language or informal language, and other features that are associated with written or more formal language (e.g. Crystal 2009; Jonsson 2015; Tagliamonte and Denis 2008). In the following, let us first trace the study of written and spoken language and of the language of immediacy versus distance. Subsequently, we will see how these can be applied to communication patterns on social media in general and on Twitter in particular.

In an early and influential large scale study on genre features in English language corpora, Biber (1988), in his MDA framework, determined typical grammatical features which tend to coincide in a given text, and which correlate with written versus spoken language use. Due to the fact that speech more typically than not is produced and comprehended under time constraints, spoken language is neither highly complex nor highly integrated. However, depending on the formality of the situation, the style used in spoken contexts can be more involved, using e.g. discourse particles, 1st person pronouns or contractions. Yet, features of highly integrated spoken registers include attributive adjectives, prepositional phrases, phrasal coordination and careful word choice (Biber 1988, 43). For a brief overview of the model, see also Buschfeld and Leuckert (this volume).

On the basis of the German language, a comparable model was developed by Koch and Oesterreicher (1985). Koch and Oesterreicher created a three-dimensional model including graphic or phonic representation of registers, thus allowing for the fact that something that has been spoken can be written down, like an interview, or something that has been written can be read off, like a sermon (Koch and Oesterreicher 1985, 18; 1985/2012, 444). Medium, as well as involvement of the participants, typically correlates with specific patterns in turn taking, spontaneity, familiarity between participants, situation embedding or publicness of the discourse. On the basis of these, we can then classify language as showing features of immediacy or of distance.

Typically, this would lead us to find the following features in the language of immediacy, and find the opposite of these features in language of distance (Koch and Oesterreicher 1985, 27; 1985/2012, 454). In terms of morphosyntactic features, in the language of immediacy we find additions, anacolutha, mistakes in congruence, holophrastic utterances, phenomena of segmentation, theme-rheme sequences, and

little hypotaxis. As specific lexical features in the language of immediacy, we find *passe-partout* words, lexical poverty, low type-token ratios, expressive language (hyperbole and cursing), and lexical richness is low and found in specific contexts only. The typical textual-pragmatic features of the language of immediacy are backchannels, gap-fillers, corrections, markers of structure, reductive particles, narrative present, direct speech in quotations, and different expectations as to textual coherence.

Koch (1999) carries forward the Koch and Oesterreicher (1985) model and applies it to texts that are less typically dialogic, namely transcripts of historical court records on the one hand and late medieval Italian cartoons on the other hand. For this he creates models of typical dialogicity. As parameters he suggests physical immediacy versus distance, familiarity of partners, emotionality, context embeddedness, deictic immediacy or distance, dialogicity or unidirectionality, communicative cooperation or not, spontaneity or not, and free versus focused topic development.

In a move to apply Biber's MDA framework to new media genres, Jonsson (2015, 29) investigates both asynchronous and synchronous CMC language. She determines written, spoken and conversational writing modes, which are expressed through the media of writing (professional letters, academic prose and so on) and asynchronous CMC, such as Twitter, Facebook posts, blog-comments, emails or others. Spoken language uses the medium of speech in face-to-face communication, telephone conversations, interviews, and also less spontaneous genres like prepared speeches and broadcasts. Conversational writing shows an even higher degree of orality features than spoken language and makes use of synchronous CMC, such as instant messaging, web chats etc., and super-synchronous communication uses split window communication and Unix Talk (Jonsson 2015, 23-29).

In her analysis, Jonsson (2015, 111-129) finds that in CMC, the degree of abstraction, and the involvement with the addressee varies. Generally, 1st person pronouns are more frequent in all types of CMC than in spoken, and much more than in written modes of interaction. Addressing one communication partner in particular is positively correlated with synchronous CMC. By contrast, synchronous CMC does not devote much attention to other people and to concepts, which leads to a low number of 3rd person pronouns. Concerning word length, Jonsson (2015, 130-149) finds that while short words are one of the most prevalent features of spoken language, longer words convey more specific, and often specialized meanings, they express rarer concepts, and are more difficult to retrieve under the time pressure of spoken dialogue. While word length in synchronous CMC even drops below the level of spoken interaction, asynchronous conversation may reach the word length of written texts. Also the Type-Token ratio (TTR), the amount of different lexical words per text, seems high in CMC, but Jonsson shows this to be largely misleading. She finds that a high TTR can be due not only to rich vocabulary of the author, but also to the use of many non-standard spellings, which lead to an inflated count of presumed word types due to both typos and flouted spelling conventions.

In the comparison with written and spoken genres, Jonsson (2015, 150) identifies the use of the following linguistic categories as most significantly different (differing by more than two standard deviations) in conversational writing: 1st person pronouns, 2nd person pronouns, direct *wh*-questions, analytic negations, demonstrative pronouns,

indefinite pronouns, present tense verbs, predicative adjectives, contractions and prepositional phrases. In order to determine in how far Twitter language, in this case the political use of language from Donald Trump's Twitter account, compares to the characteristics of CMC language, these distinctive linguistic categories are traced in the data set of Trump tweets.

Twitter, as a means of CMC, is subsumed under Jonsson's heading of asynchronous CMC. Arguably, it is a unique category of writing which, due to its restricted token count per tweet, requires brevity. Thus, Twitter language leaves no room for lengthy explanations. Further it can also be argued to be ephemeral: while spoken language is ephemeral in the sense that it cannot be retrieved after having been uttered (Koch and Oesterreicher 1985/2012, 451-452), old tweets are superseded by responses and new tweets and are thus out of sight and out of consciousness (McIntosh 2020, 7). Similarly, Stolee and Caton (2018, 153) observe that contradictory comments on, or reactions to, the same stimulus at different points in time are not an issue for the Twitter community due to the temporal sequencing of the Twitter messages; a tweet represents an instant reply to an event, which is then superseded by new events and new Twitter feeds. These characteristics of Twitter, Stolee and Caton (2018, 157-160) argue, are perfect for Donald Trump's message, in which a tweet addresses the latest crisis of a moment. That Trump speaks his mind seemingly uninhibitedly in such moments of (perceived) crisis is seen by his supporters as a sign of his candidness and genuineness. Typing or language errors in the tweets add to his perceived authenticity and unfilteredness. Factuality of a tweet is less of an issue to a community that has, by the time the facts are checked, already moved on to a new issue. This gives possibility to escape from long-term consequences of the words (2018, 160). Voters also appreciate that they feel able to enter a dialogue with their president by interacting on Twitter, especially if they watch the same television programmes as him (2018, 161). By these standards, on Twitter we can indeed see the President unplugged.

The stylistic features of his language use on Twitter are investigated by Clarke and Grieve (2019). Using an automated approach to the analysis of Trump's tweets, the authors classify more than 21,000 tweets produced between 2009 and early 2018 on the basis of the stylistic features identified in Biber's (1988) MDA framework. Clarke and Grieve's large-scale analysis finds that the stylistic features of Trump's tweets largely pattern according to five dimensions of the tweets. Clarke and Grieve discard their first dimension, text length, from scrutiny as it predetermines the possible number of features that can be found in a text (2019, 10-12).

The other dimensions according to which the tweets pattern are conversational style, campaigning style, engaged style, and advisory style (Clarke and Grieve 2019). One clear patterning of tweets is found along a dimension of conversational style, with grammatical features that belong to informal, spoken style language. Often, there is interaction with other Twitter users, such as initial mentions, especially the use of 2nd person pronouns, contractions, question words and question marks, analytic negation and predicative adjectives (2019, 12-15), as in "@RaydelMusic That's great – you will love what we are doing!" (30 January 2013, D2: 0.653; cf. Clarke and Grieve 2019, 12). Non-conversational, literary style levels by contrast showcase many nouns of

different types, many noun-premodifiers and prepositions. Clarke and Grieve observe that after declaring his candidacy for the presidency, Trump's Twitter language became more informal, then rose in formality after he was elected, to grow more informal again during his time in office when reaching out to his voter base (2019, 13-14).

A further dimension along which the tweets differ are whether they promote Trump's campaign or not. Those tweets that do, mainly found in the run-up to the election, use many 1st person pronouns, possessive determiners/pronouns, modals of predication and time adverbs and imperatives, as well as capitalisations and exclamation marks. Non-campaign tweets are more declarative and less self-focused: they offer more opinions and descriptions and thus more 3rd person references, predicative adjectives and superlatives (Clarke and Grieve 2019, 15-16). The fourth dimension identified is engagedness of style. Engaged style uses interactive features such as question words, question marks, 2nd person pronouns, analytic negations or modals of possibility. Negative engagement is marked by features showing fixed opinions such as attributive and predicative adjectives, copular verbs, amplifiers, superlatives or exclamation marks. An example of highly engaged style is "I still don't know who I'm going to choose. @GeraldoRivera or @LeezaGibbons? Who do you like? @ApprenticeNBC" (13 February 2015, D4: 0.798; Clarke and Grieve 2019, 16). Clarke and Grieve (2019, 17-18) observe that since Donald Trump has come to power, his tweeting style has become less engaged than it was before he was elected.

The last dimension observed in the study is that of advisory style. Where this is present in the tweets, we find advice, with features including imperatives, personal pronouns, especially *it* and 2nd person pronouns, predicative adjectives, copular *be*, superlatives or perception verbs. Tweets that are low on the advisory domain use pronouns, particularly in the 3rd person, narrative features like past tense, perfect and progressive aspect and devices creating an "'us versus them' dichotomy" (Clarke and Grieve 2019, 19), such as modals of necessity, passives and rhetorical questions. The different stylistic dimensions introduced above are used on the Trump Twitter account to reflect political goals and to promote Trump and his campaign. Thus, the authors conclude, whether individual tweets are written on impulse or not, the overall Twitter campaign shows a clear political strategy.

3. Data and Method of the Study

The current study carries out a smaller-scale, manual, qualitative and quantitative analysis of Twitter data from a set period of time. The data was collected from the tweets @realDonaldTrump, taken from the Trump Twitter Archive at <www.trumptwitterarchive.com/>. Like Grieve (2017), I acknowledge that it is more than likely that not all the tweets were written by the account owner only. However, as the Twitter account is perceived in the public eye as representing the Trump presidency, I believe that the content should be analysed as a whole, without trying to subdivide tweets according to potential authors.

The period from which the tweets are investigated is that from 29 August to 5 September 2019. This time span represents a non-campaign time. According to Clarke

and Grieve's (2019) dimensions, this time span is not likely to coincide with a period of the especially highly engaged pre-election style. The period did offer multiple reasons for potential high engagement, however: the President is under pressure for his handling of Hurricane Dorian and the communication strategy surrounding his approach. He is furthermore suffering from the fallout of his dismissal of erstwhile FBI director James Comey, and is under criticism for his trade war with China. Thus, this arguably represents a time where heightened activity from the Twitter account could be expected.

In this period, 200 tweets were sent. Of these tweets, those that are mere retweets of other Twitter users' content were not considered for linguistic analysis. Tweets that are retweets of Trump's own tweets, by contrast, are considered. This leaves 137 tweets with 4,531 words for the detailed textual analysis of this paper. In future research, the scope of this analysis will be extended to a larger dataset to obtain more representative results, however.

Data analysis took place in two steps. First, it was determined according to the criteria given in Koch (1999) in how far this data set from Donald Trump's Twitter account fits the criteria of language of distance versus language of immediacy. Second, it was investigated how the data set can be situated in the broad context of Biber's (1988) MDA framework. For this, Jonsson's (2015) application of Biber's framework to CMC was used, in order to situate Trump's Twitter language in terms of typical features of CMC language. Those features that are identified by Jonsson (2015) as being distinctive for the language of CMC were searched for and evaluated manually in the data in hand. This approach has the advantage that the researcher is able to get to know the data set well, without running the danger of overlooking potentially relevant features. In order to determine quantitative results – which need to be considered with a large degree of caution as the data set is too small to be truly representative – results have been normalized to 1,000 words.

An exception had to be made for the TTR, which Jonsson determines on the basis of 400 word-segments, which she then normalized to a ratio of types per 100 words (2015, 91). Due to the extreme shortness of tweets (no more than 280 characters per tweet) tweets typically have insufficient context to determine TTRs (cf. Clarke and Grieve 2019). A measure of lexical density that generally deals well with short texts is Measure of Textual Lexical Density (MTLD), but even for this a text length of at least 100 tokens is advocated so that any meaningful repetition could be assumed (Koizumi 2012). Thus, in order to compare lexical richness in tweets, it would be most profitable to compare the lexical richness of Donald Trump's tweets against tweets of other Twitter users, which is not the purpose of this particular paper but is planned for a separate analysis at a later time. For notational purposes only, a Mean Segmental Type-Token Ratio (MSTTR) is calculated on 400-word-slots and the figures are normalized to a TTR per 100 words, as done by Jonsson (2015).

4. Results

When analyzing the presence of features of distance versus immediacy in Twitter language, and in the tweets from @realDonaldTrump in particular, we find elements of

both distance and immediacy according to Koch (1999). This is not unduly surprising as Jonsson (2015) identifies asynchronous CMC, of which Twitter language forms a part, as comprising features typical of both written and spoken genres of language. Results from the investigation of the Trump twitter data along these two models are discussed below.

4.1 Language of Immediacy and Distance in Trump's Twitter Data

We have raised the question in how far features of Trump's tweets can be counted as part of Koch's (1999) immediacy and distance paradigms. Table 1 gives an overview of how features of the tweets can be situated in these paradigms.

Language of immediacy	Applicable to tweets?	Language of distance	Applicable to tweets?
Physical (spatial, temporal) immediacy	-	Physical (spatial, temporal) distance	+
Privacy	-	Non-privacy	+
Familiarity of partners	-/+	Lack of familiarity	+/-
Highly emotional	?	Lack of emotion	?
Context embedded	-	Context dissociation	+
Deictic immediacy	-/+	Deictic distance	+/-
Dialogue	-/+	Unidirectionality	+/-
Communicative cooperation	-	No/little cooperation	+
Free topic development	+	Focused topic development	-
Spontaneity	-?	Planned	+?

Table 1: Applicability of dialogic features to Tweeting following Koch (1999, 402-403)

Table 1 indicates that Twitter language does not need – nor provide – physical immediacy. Temporal immediacy may of course exist if users check their Twitter feed immediately, but probably temporal distance is the more typical feature. Yet some of Trump's voter-base are temporally close to him and can decode references made to television shows in his tweets. This could be seen in an incident where Trump referred to what happened "last night in Sweden," which referred to a news programme aired on TV the previous evening (Stolee and Caton 2018, 162). The parameters of context embeddedness and deictic immediacy go together with physical immediacy. While typically these do not apply to tweets, both may be created if followers read tweets immediately and are engaged in the same activities as the tweeter, e.g. watching the same television shows. Privacy, an item that is crucial to spoken discourse, can be achieved on Twitter by choosing settings that only allow accepted followers to see your tweets, but this obviously is not the case for Donald Trump's Twitter account. As a result, the conversation partners are not necessarily familiar – though Trump's followers

of course know him and he knows some of the other Twitter users he interacts with. Indeed, interacting with and evaluating other people is an important characteristic of Trump's Twitter use, as argued for instance by Clarke and Grieve (2019, 15-16), who discuss his engaged style. Depending on whether or not tweets rank high on conversation style or engaged style domains or low on advisory domains (Clarke and Grieve 2019), the tweets may or may not display features of emotionality.

While typical written language is unidirectional, this is not the case for tweets. Clarke and Grieve (2019) have shown this in particular in their discussion of Trump's engaged style, which interacts with his followers such as by asking questions, or directly addressing other Twitter users via their Twitter handles. A high frequency of retweets of other users' content also belongs to the category of engaged tweets. Of the 200 tweets in the corpus in hand, 63 (31.5%) were retweets. According to Stolee and Caton, retweets and embedded tweets constitute dialogues as they "creat[e] a dialogical narrative curated by one user" (2018, 153). In addition, other Twitter users' handles are directly addressed in 29 tweets (14.5%). In about half of these cases, the tweets address multiple journalists or agencies (examples 1, 2). Single users are addressed 15 times. With one exception, these single users are either journalists or Republican politicians (example 3). A further 17 tweets (8.5%) contain embedded tweets. Thus more than half of the tweets in the corpus, 54.5%, use overt dialogic, interactional features with other Twitter users.

1. I am monitoring Hurricane Dorian and receiving frequent briefings and updates. It is important to heed the directions of your State and Local Officials. This is an extremely dangerous storm, please prepare and be safe! @FEMA @NWS @NOAA @Readygov @FLSERT <https://t.co/Fee9EB1wQk> (30 August 2019 11:13:39 PM)
2. The IG found that James Comey leaked Classified Documents to his attorneys, which can be (is) a violation of the Espionage Act. This is consistent with all the lies that Comey has been spreading. @GreggJarrett @ByronYork @LouDobbs (31 August 2019 07:11:16 AM)
3. Thanks @RepDougCollins – TRUE! <https://t.co/o49tPKdWct> (30 August 2019 01:05:03 PM)

But it is not only Trump who is interactive, Seoane Pérez et al. (2019, 27) note that reader interaction with Trump's tweets is high: in sample sets of Trump and Clinton tweets from between July and November 2016, Trump's were generally more likely to be liked, retweeted or receive replies than Clinton's.

Regardless of this potentially dialogic nature, we generally find little communicative cooperation in tweets, which put this parameter into the "distance" domain, instead we have a completely free topic development, which clearly links to the "immediacy" domain. An interesting case is constituted by the location along the spontaneous versus planned axis of the model. While language of distance, and thus typical written language, would be planned ahead, typically language of immediacy is spontaneous. While it seems expectable that public utterances of a person holding a high political office should be highly planned, just how much forethought has gone into the tweets in hand is sometimes questionable.

4. The incompetent Mayor of London, Sadique Kahn, was bothered that I played a very fast round of golf yesterday. Many Pols exercise for hours, or travel for weeks. Me, I run through one of my courses (very inexpensive). President Obama would fly to Hawaii [...]. (3 September 2019 09:20:10 AM)

Example 4 in fact contains two misspellings of the Mayor of London's name, Sadiq Khan, which might point to spontaneous, unchecked tweeting, possibly prompted by criticism of Trump. Stolee and Caton (2018, 159-160) indeed argue that Trump's tweeting is neither self-edited, nor self-censored, which, for his supporters, adds to his authenticity. Nearly the same tweet is sent again 28 minutes after the first one, but the spelling of Khan's name is corrected in the second tweet. Even though Khan has a Twitter account too, there is no dialogicity here: no direct Twitter communication takes place with the opponent.

Combining these observations, we can see that, in spite of the physical distance of the tweeter and other Twitter users, various features of the language of distance are found: physical distance, lack of privacy, context dissociation, lack of cooperation. Whether the discourse is planned in all cases has been questioned above. We also find lack of familiarity, unidirectionality and deictic distance in most cases, but exceptions exist. Yet, very strong features of immediacy are found: dialogicity is built in many cases with the help of retweets and direct addresses of other Twitter users. Deictic distance can be, and is, overcome by the use of deictic markers, which followers can decode if they watch the same television shows or follow the same news reports (Stolee and Caton 2018): if they understand the context of the tweet, they are "in-the-know" and can feel part of an in-group (example 5).

5. Such a phony hurricane report by lightweight reporter @jonathancarl of @ABCWorldNews [...]. (2 September 2019 06:12:29 PM)

If an in-group feeling can be created successfully throughout the Twitter community, the fact that in most cases there is no reciprocal familiarity of Trump and his Twitter followers may be overcome and the feeling of familiarity may be created. That using forms of CMC can indeed move language further towards language of immediacy is observed by authors who compare CMC media with non-CMC media: Landert and Jucker (2011) show this for letters to the editor in print and online news media. Online letters can be situated more closely towards the language of immediacy pole than their print media equivalents. To account for this, Landert and Jucker (2011) introduce public accessibility and privacy dimensions into Koch and Oesterreicher's model. However, as cautioned by Dürscheid (2016, 379-381), when investigating communication through new media we need to bear in mind that a certain medium of communication does not mean all genres or registers of communication are identical on it.

4.2 Comparing Trump's Tweets to other Types of CMC

After having determined how Trump's tweets fit into the paradigms of language of immediacy and of distance, I would now like to continue with the question of how linguistic features of Trump's tweets relate to typical features of CMC. For this, I am

comparing the use of those features that have been identified by Jonsson (2015, 150) as differing most significantly from Biber's (1988) Multidimensional Analysis of spoken and written registers. These features are: the use of 1st and 2nd person pronouns, direct *wh*-questions, use of analytic negation, demonstrative and indefinite pronouns, present tense verbs, predicative adjectives, contractions, and prepositional phrases. Jonsson (2015) uses normalized frequencies per 1,000 words, thus, for purposes of comparison, the same normalized frequencies are given here.

Let us consider the use of 1st and 2nd person pronouns first (example 6).

6. [...] I am getting the North Carolina Emergency Declaration completed and signed tonight. Hope **you** won't need it! (3 September 2019 07:43:10 PM)

This tweet represents a nice example of an engaged style: next to the self-promoting 1st person pronouns typical of Trump's campaign style (Clarke and Grieve 2019), we find second person addressees being cared for, analytic negation and a modal auxiliary "won't," which are typical of highly engaged style. Overall, we find that in the corpus of tweets, both 1st and 2nd person pronouns are used less frequently than in average asynchronous CMC, of which Twitter forms part (Jonsson 2015), namely at 23.24 and 4.0 per 1,000 words respectively (for a tabular overview, see Table 2 below). A further feature of engaged tweets (Clarke and Grieve 2019) is the use of *wh*-questions as in example (7).

7. Comey very vulnerable. But **where** is the Supreme Court. **Where** is Justice Roberts? (1 September 2019 07:58:44 AM)

Compared to the use of *wh*-question words in other genres of communication (Jonsson 2015, 150), at 1.1 examples per 1,000 words the feature is frequent compared to written and spoken interaction, but low in comparison to other asynchronous CMC. Similarly, the use of analytic negations such as *not*, *no*, *neither*, or *nor* (example 8), which is also identified as frequent in engaged styles by Clarke and Grieve (2019, 17), is comparatively low at 10.85 per 1,000 words in comparison to typical asynchronous CMC observed by Jonsson (2015, 150), but more frequent than in written, and less frequent than in spoken genres.

8. I have **not forgotten** that when it was announced that I was going to do The Apprentice [...] @DebraMessing [...] profusely thanked me, even calling me "Sir." (1 September 2019 08:26:03 AM)

A very interesting feature is constituted by the use of demonstrative pronouns. Demonstrative pronouns, as deictic markers, are more typical of the language of immediacy than the language of distance. Correspondingly, Jonsson (2015) finds demonstratives to be considerably more frequent in spoken language than in written genres, and also more frequent than in synchronous and asynchronous CMC genres. However, in the data in hand, demonstratives are found with a higher frequency (example 9).

9. Has **anyone** noticed that the top shows on @foxnews and cable ratings are **those** that are Fair (or great) to your favorite President, me! (31 August 2019 07:15:06 AM)

In relation to Jonsson's (2015) typical asynchronous CMC data, at 11.51 per 1,000 words, the number of demonstratives here is comparatively high. This suggests that conceptual space is indeed being shared by Trump and his Twitter followers: they know what he is talking about. By contrast, the use of indefinite pronouns, such as *something*, *anything* or *anyone* (example 9) is at the lower end of the range of typical asynchronous CMC data at 3.98 per 1,000 words.

Jonsson's next feature that is distinct in CMC is the use of the present tense verbs (example 10).

10. The U.S. now **leads** the world in energy production... BUT [...] (4 September 2019 04:08:58 PM)

In asynchronous CMC, this is typically at a similar level to written genres. Here in the tweet corpus it is lower than in any of Jonsson's genres (2015, 150), 46.71 per 1,000 words. Clarke and Grieve (2019, 15) find that the typical present tense feature, 3rd person singular verbal *-s*, is particularly indicative of the negative dimension of campaign language: Trump is externally focused. We find that instead of the present tense verbs, there is a considerable amount of reference to past events, or the outcome of past events, which are coded in the perfect, such as for example "has anyone noticed" in (9) above. The use of these past and perfects, according to Clarke and Grieve (2019), is most typical of Trump's conversational style.

Particularly instructive is the use of adjectives in the corpus of tweets. Following Biber (1988), Jonsson (2015, 161-162) argues that, while the use of attributive adjectives is typical of nominal types of written discourse, the use of predicative adjectives is typical of conversational writing. She points out that predicative adjectives are typically used as a stance-marker and that they often are evaluative and emotive. These traits can be seen clearly in Donald Trump's tweet data (example 11).

11. Read my FULL FEMA statement. What I said was **accurate!** (5 September 2019 06:48:56 AM)

In comparison with other genres, the use of predicative adjectives is very high here, at 12.4 per 1,000 words. Unfortunately, Jonsson (2015) has no data for predicative adjectives in asynchronous CMC, the category that we would have been most interested in, but the figures in the Trump tweets are considerably higher than in all other genres except for super-synchronous CMC. This very high incidence of predicative adjectives confirms that the language here is very evaluative.

The last two categories which Jonsson (2015) found to be significantly different in CMC from written and spoken genres are the amount of contractions and of prepositions. It is well-known that contractions are rare in writing and frequent in spoken language. Not unexpectedly, in typical asynchronous CMC, Jonsson (2015, 150) finds them to be between written and spoken frequencies, while the frequencies are very high in super-synchronous CMC. While we might have assumed that, given the character limit in tweets, contractions in tweets should be even above average asynchronous CMC levels, in fact, Trump's tweets show contraction levels which, at 8.8 per 1,000 words, are more similar to the figures found for written genres (4.8/1,000)

than spoken (36.1/1,000) or CMC genres (16.6/1,000). In terms of contractions, Trump is a conservative speller.

A further feature that is found comparatively rarely in the corpus is that of prepositions at 74.6 per 1,000 words (example 12).

12. [...] Absolutely worth it, we **don't** want to be servants **to** the Chinese! [...] (1 September 2019 07:25:28 AM)

That prepositional phrases should be rarer in the tweets than in written genres is not surprising as they are a typical means to transmit densely packed information, most frequent in official and academic discourse (Biber 1988, 237). As a result, prepositions are also considerably rarer in spoken conversation. In Jonsson's (2015, 150) figures, prepositional phrases are at a comparable level for both written and asynchronous CMC genres, and notably lower for spoken genres. Jonsson (2015, 166) notes, however, that prepositional phrases are infrequent in conversational writing, particularly in synchronous and super-synchronous CMC. Where they are used, they mainly serve to lower complexity rather than to increase it, for example the preposition might govern just a pronoun. Thus, the comparatively low frequency of prepositions in the tweets is not unexpected. Whether it can be considered particularly low would need to be determined in comparison with specific tweet corpora, which is outside the scope of this contribution.

The final issue to be considered here is that of lexical parameters. First, let us address the question how the average word length in the tweets compares to that in other genres. I have found no information on whether hashtags and Twitter handles are included in Jonsson's (2015) statistics and have thus decided to retain such items, e.g. "#FakeNewsCNN" or "@ABCWorldNews," in the data. While Jonsson computes an average word length of 4.5 characters for asynchronous CMC, the word length here is slightly higher at 4.61 and thus matches that of Jonsson's average for written genres (2015, 131). It notably exceeds the average word length in speech and synchronous and super-synchronous CMC (cf. Table 2). In this context it is worth noting again that there are comparatively few contractions in the tweets, which also increases average word length.

The second lexical parameter is that of TTR. As pointed out in section 3 above, typically TTRs are determined on the basis of lengthy textual chunks, such as per 1,000 words. The shorter the investigated texts become, the less meaningful the TTR will be: a new utterance on a new topic typically means that entirely different contexts are touched and new words will be used. Correspondingly, short texts should have more different words, and thus a higher TTR, than long, coherent text. In order to obtain any meaningful results from the very-short-text category of tweets, the only approach that would provide any meaningful results is a comparison of the lexical diversity of one Twitter user compared to a corpus of other Twitter users' data. This would allow us to determine how fast the pace in the Trump tweets is, how often new topics are introduced and how diverse the vocabulary is that is used in their discussion. Using this approach, we might combine tweets into 400-word-chunks and then determine the mean TTR of these chunks, which would give us a Mean Segmental TTR, such as has also been used to determine TTR by Jonsson (2015). By this method we find that the

Mean Segmental TTR of the Trump tweets is 59 types per 100 tokens (i.e. words). At first sight this seems to correspond to averages that are also determined for written and synchronous CMC (see Table 2) by Jonsson (2015). However, as stated, topic coherence for tweets is not comparable to these genres and thus this similarity could be spurious.

Category	Writing	ACMC	Speech	SCMC	SSCMC	Trump-tweets
1 ps. pronouns	17	57.8	52.8	56.9	88.9	23.02
2 ps. pronouns	5	17.6	23	50.4	45	4
<i>Wh</i> -questions	0.1	2.9	0.8	3.5	3.9	1.1
Analytic neg.	6.4	15.1	13.9	13.1	29.7	10.85
Dem. pronouns	2.3	6.5	10.6	6.6	16.4	11.51
Indef. pronouns	0.9	4.6	3.1	11.7	6	3.98
Pres. tense verbs	64.6	67.6	112.3	147.2	168.5	46.71
Predicative adj.	4.8	n.a.	4.9	8.4	15.3	12.4
Contractions	4.6	16.6	36.1	30.8	55	8.8
Prep. phrases	117.3	116.9	91.1	47	42	74.6
Av. word length	4.6	4.5	4.2	4	3.7	4.61
TTR	52.8	56.8	46.8	54.9	52	59 ¹

Table 2: Results integrated into Jonsson's (2015) features

Table 2 illustrates various key features of the Trump tweets in comparison with distinctive features of CMC writing. Concerning pronoun use, we find a very low count of 1st and 2nd person pronouns compared with other asynchronous CMC. Clarke and Grieve (2019) find these two features to be highest in Trump's engaged and campaign centred styles. To determine in how far the pronouns are used considerably differently from other Twitter users, however, we would need to compare to a more diverse tweet corpus. On the basis of the figures here, both direct addresses of the readers and also talking in the first person seem low compared to other genres. By contrast, we have

1 Note the methodological difficulties and resulting basic incomparability when comparing the Mean Segmental TTR of tweets to other genres discussed above.

very high counts of demonstrative pronouns. This suggests that the discourse here, for a genre that is essentially not one of immediacy, is comparatively deictic and referential. Here, the suggestion that online media are more deictic than offline media (Landert and Jucker 2011) must be considered. But we can also assume that the Trump account creates a common mental space with an in-group of followers (cf. Stolee and Caton 2018, 162). By contrast, the use of indefinite pronouns like *someone/anywhere* etc. is broadly in line with other types of asynchronous CMC and references are no more vague than elsewhere.

With regard to syntax, we find a low count of the syntactic features *wh*-questions and analytic negations compared to other asynchronous CMC. Both of these features were identified as typical of Trump's engaged tweeting style by Clarke and Grieve (2019), which does not seem to be very prominent here. Also, given the low use of overtly marked present tense verbs in comparison to other asynchronous CMC, the tweets show an external focus. Instead of present tense forms, an analysis would find more past tense and present perfect verb forms, which are indicative of Clarke and Grieve's (2019) conversational styles. However, as the use of these parameters does not differ significantly in CMC in Jonsson's (2015) model on which the current study is based, the investigation of this feature must be left to future work. Further, Table 2 clearly shows the very high counts of predicative adjectives, comparable only to synchronous and super-synchronous CMC, and their very high incidence mirrors the very evaluative language use (Biber 1988; Jonsson 2015).

The count of prepositional phrases in the tweets is low in comparison to asynchronous CMC. It is much rarer in speech than writing (Biber 1988) and most rare in (super-)synchronous CMC (Jonsson 2015, 150). The low count in the tweets indicates a low syntactic complexity. Very interesting is the low number of contractions compared to other CMC genres. This may point to a conservative spelling style.

For the lexical feature of average word length, Table 2 shows us that the average length in the tweets corresponds to average word length in written language. I have suggested above that this may be due to a conservative spelling style, but also due to my decision not to remove hashtags and Twitter-handles from the data base. Concerning the TTR, it must be borne in mind that tweets should not be compared to other genres and should only be compared to TTRs of other tweet corpora. Only then can we determine if the 59 tokens per 100 types can be considered a high or a low rate.

In answer to our research questions we expectedly find both features of the language of immediacy and distance (Koch, 1999; Koch and Oesterreicher 1985) in the tweets. The level of immediacy is generally high, which is typical of a CMC medium (Dürscheid 2016; Landert and Jucker 2011). This can be shown in particular on the basis of high levels of dialogicity, in particular through retweets and directly addressing other Twitter users. Given the medium of Twitter and the physical distance between the participants, the use of fewer deictic markers such as demonstrative pronouns would have been expected. I argue, however, that using deictic markers may help to create an in-group feeling amongst those of Trump's followers who are able to decode the references. This holds in particular if followers check their Twitter feed regularly and there is no temporal distance between sending and receiving the tweet. The tweets also

appear spontaneous – and thus arguably authentic – due to both "unfiltered" evaluative language and misspellings.

Concerning the question of how the typical features of CMC are used, the most notable findings are that compared to asynchronous CMC, the language is considerably more evaluative, as indicated by the high use of predicative adjectives. In spite of strong dialogicity, use of both 1st and 2nd person pronouns is low. As these pronouns have been identified as relating to specific styles of Trump's tweeting (Clarke and Grieve 2019), the period that is represented in the small data set here may have influenced this outcome. The data set does not stem from a campaign period and there is a strong focus of the tweets on external factors like media criticism and Hurricane Dorian. To gain confirmation of this, a larger data set would need to be investigated. Also in spite of dialogicity, there is a low use of the typical high engagement feature *wh*-questions (Clarke and Grieve 2019). Overall, a picture thus emerges of a highly interactive tweeter, with strong personal opinions and a, for the medium, conservative spelling style, notwithstanding spelling mistakes deriving from spontaneous tweeting.

5. Conclusion

The results of this study show that Koch and Oesterreicher's model provides a very useful framework for the analysis of Twitter language. Combining their framework with Biber's or Jonsson's proves very fruitful as the MDA framework provides a good base for the selection of features on which to base the analysis.

In answer to the research question how tweets sent from Donald Trump's Twitter account situate within the dimensions of language of immediacy and distance, we have found that the tweets are situated between the two poles, but, like other genres in online communication, positioned closer to language of immediacy. In the investigated tweets, we find more features of the language of immediacy. In particular, these are signs of spontaneity, such as misspellings, as well as dialogic features, which derive from interaction with other Twitter users and are particularly noteworthy in the examined tweets. To what extent these features deviate from practices by other Twitter users is, however, a question that should be pursued in further research.

In answer to the question how typical linguistic features of language of CMC are used in the tweets, we find that features of evaluative language are strongly represented. In spite of dialogic interaction with other Twitter users, key interactional features identified by Jonsson (2015), especially the use of *wh*-questions, remain rare.

However, the study so far is based on a small data set. In order to obtain a more robust picture of Trump's Twitter use, a larger corpus with more time depth should be investigated and the results gleaned from this will need to be compared with a tweet corpus for parameters such as word length and particularly Type-Token Ratio to be compared meaningfully.

Works Cited

Biber, Douglas. *Variation across Speech and Writing*. Cambridge: Cambridge University Press, 1988.

- Clarke, Isobelle, and Jack Grieve. "Stylistic Variation on the Donald Trump Twitter Account: A Linguistic Analysis of Tweets Posted between 2009 and 2018." *PLoS ONE* 14.9 (2019). DOI: 10.1371/journal.pone.0222062.
- Crystal, David. *Txtng: The Gr8 Db8*. Oxford: Oxford University Press, 2009.
- Dürscheid, Christa. "Nähe, Distanz und neue Medien." *Zur Karriere von 'Nähe und Distanz.'* *Rezeption und Diskussion des Koch-Oesterreicher-Modells*. Eds. Helmuth Feilke and Mathilde Hennig. Berlin: De Gruyter, 2016. 357-385. DOI: 10.1515/9783110464061-013.
- Grieve, Jack. "Trump Twitter Analysis." *RStudio* 2017. <www.rpubs.com/jwgrieve/340342> [accessed 31 January 2019].
- Jonsson, Ewa. *Conversational Writing: A Multidimensional Study of Synchronous and Supersynchronous Computer-Mediated Communication*. Frankfurt a.M.: Peter Lang, 2015.
- Koch, Peter. "Court Records and Cartoons: Reflections of Spontaneous Dialogue in Early Romance Texts." *Historical Dialogue Analysis*. Eds. Andreas H. Jucker, Gerd Fritz, and Franz Lebsanft. Amsterdam: John Benjamins, 1999. 399-429. DOI: 10.1075/pbns.66.16koc.
- Koch, Peter, and Wulf Oesterreicher. "Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte." *Romanistisches Jahrbuch* 36 (1985): 15-43. DOI: 10.1515/9783110244922.15.
- Koch, Peter, and Wulf Oesterreicher. "Language of Immediacy – Language of Distance: Orality and Literacy from the Perspective of Language Theory and Linguistic History." *Communicative Spaces. Variation, Contact, and Change: Papers in Honour of Ursula Schaefer*. Eds. Claudia Lange, Beatrix Weber, and Göran Wolf. Frankfurt a.M.: Peter Lang, 1985/2012. 441-473. DOI: 10.15496/publikation-20415.
- Koizumi, Rie. "Relationships Between Text Length and Lexical Diversity Measures: Can We Use Short Texts of Less than 100 Tokens?" *Vocabulary Learning and Instruction* 1.1 (2012): 60-69. DOI: 10.7820/vli.v01.1.koizumi.
- Landert, Daniela, and Andreas H. Jucker. "Private and Public in Mass Media Communication: From Letters to the Editor to Online Commentaries." *Journal of Pragmatics* 43.5 (2011): 1422-1434. DOI: 10.1016/j.pragma.2010.10.016.
- Lockhart, Michele, ed. *President Donald Trump and his Political Discourse. Ramifications of Rhetoric via Twitter*. London: Routledge, 2019.
- McIntosh, Janet. "Introduction: The Trump Era as a Linguistic Emergency." *Language in the Trump Era. Scandals and Emergencies*. Eds. Janet McIntosh and Norma Mendoza-Denton. Cambridge: Cambridge University Press, 2020. 1-44. DOI: 10.1017/9781108887410.001.
- Partington, Alan, and Charlotte Taylor. *The Language of Persuasion in Politics*. London: Routledge, 2019.
- Seoane Pérez, Francisco, Irene Asiaín Román, and Javier Lorenzo Rodríguez. "Seizing the Populist Rhetorical Toolkit: a Comparative Analysis of Trump and Clinton's Discourse on Twitter during the 2016 U.S. Presidential Campaign." *President*

- Donald Trump and his Political Discourse. Ramifications of Rhetoric via Twitter.* Ed. Michele Lockhart. London: Routledge, 2019. 13-32. DOI: 10.4324/9781351038782-2.
- Searle, John R. "A Classification of Illocutionary Acts." *Language in Society* 5.1 (1976): 1-23. DOI: 10.1017/S0047404500006837.
- Stolee, Galen, and Steve Caton. "Twitter, Trump, and the Base: A Shift to a New Form of Presidential Talk?" *Signs and Society* 6.1 (2018): 147-165. DOI: 10.1086/694755.
- Tagliamonte, Sali A., and Derek Denis. "Linguistic Ruin? LOL! Instant Messaging and Teen Language." *American Speech* 83.1 (2008): 3-34. DOI: 10.1215/00031283-2008-001.
- Trump Twitter Archive. www.trumptwitterarchive.com. [accessed 26 October 2020].²

2 Migrated to: <https://www.thetrumparchive.com/> in 2021 [accessed 14 June 2021].

