

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: <http://www.journals.elsevier.com/infection-disease-and-health/>

Research paper

Twitter and Middle East respiratory syndrome, South Korea, 2015: A multi-lingual study

Isaac Chun-Hai Fung ^{a,*}, Jing Zeng ^{b,1}, Chung-Hong Chan ^c, Hai Liang ^d,
Jingjing Yin ^e, Zhaochong Liu ^f, Zion Tsz Ho Tse ^{g,2}, King-Wa Fu ^{c,2}

^a Department of Epidemiology and Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA 30460-8015, USA

^b Digital Media Research Centre, Queensland University of Technology, Brisbane, QLD, Australia

^c Journalism and Media Studies Centre, The University of Hong Kong, Hong Kong

^d School of Journalism and Communication, Chinese University of Hong Kong, Hong Kong

^e Department of Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA 30460-8015, USA

^f Department of Computer Science, The University of Georgia, Athens, GA 30602, USA

^g College of Engineering, The University of Georgia, Athens, GA 30602, USA

Received 26 June 2017; received in revised form 24 August 2017; accepted 24 August 2017

KEYWORDS

MERS;
Social media;
Language;
Culture;
Health
communication

Abstract *Background:* Different linguo-cultural communities might react to an outbreak differently. The 2015 South Korean MERS outbreak presented an opportunity for us to compare tweets responding to the same outbreak in different languages.

Methods: We obtained a 1% sample through Twitter streaming application programming interface from June 1 to 30, 2015. We identified MERS-related tweets with keywords such as 'MERS' and its translation in five different languages. We translated non-English tweets into English for statistical comparison.

Results: We retrieved MERS-related Twitter data in five languages: Korean ($N = 21,823$), English ($N = 4024$), Thai ($N = 2084$), Japanese ($N = 1334$) and Indonesian ($N = 1256$). Categories of randomly selected user profiles ($p < 0.001$) and the top 30 sources of retweets ($p < 0.001$) differed between the five language corpora. Among the randomly selected user profiles, K-pop fans ranged from 4% in the Korean corpus to 70% in the Thai corpus; media ranged from 0% (Thai) to 14% (Indonesian); political advocates ranged from 0% (Thai) to 19% (Japanese); medical professionals ranged from 0% (Thai) to 7% (English). Among the top 30 sources of retweets for each corpus (150 in total), 70 (46.7%) were media; 29 (19.3%) were

* Corresponding author. Jiann-Ping Hsu College of Public Health, Georgia Southern University, PO Box 8015, Statesboro, GA 30460-8015, USA. Fax: +1 912 478 0171.

E-mail addresses: cfung@georgiasouthern.edu (I.C.-H. Fung), zengjing310@gmail.com (J. Zeng), u3003078@connect.hku.hk (C.-H. Chan), hailiang@cuhk.edu.hk (H. Liang), jyin@georgiasouthern.edu (J. Yin), zhaochongliu@gmail.com (Z. Liu), ziontse@uga.edu (Z.T.H. Tse), kwf@hku.hk (K.-W. Fu).

¹ These two individuals serve as co-first authors.

² These two individuals serve as co-senior authors.

<http://dx.doi.org/10.1016/j.idh.2017.08.005>

2468-0451/© 2017 Australasian College for Infection Prevention and Control. Published by Elsevier B.V. All rights reserved.

Please cite this article in press as: Fung IC-H, et al., Twitter and Middle East respiratory syndrome, South Korea, 2015: A multi-lingual study, Infection, Disease & Health (2017), <http://dx.doi.org/10.1016/j.idh.2017.08.005>

K-pop fans; 7 (4.7%) were political; 9 (6%) were medical; and 35 (23.3%) were categorized as 'Others'. We performed chi-square feature selection and identified the top 20 keywords that were most unique to each corpus.

Conclusion: Different linguo-cultural communities exist on Twitter and they might react to the same outbreak differently. Understanding audiences' unique Twitter cultures will allow public health agencies to develop appropriate Twitter health communication strategies.

© 2017 Australasian College for Infection Prevention and Control. Published by Elsevier B.V. All rights reserved.

Highlights

- Tweets about the 2015 South Korean MERS outbreak in 5 languages were compared.
- User profiles and keywords used were different across the 5 data sets.
- Health communication on Twitter should be culture and language specific.

Introduction

Different linguo-cultural communities may react to an infectious disease outbreak differently, either due to varying geographical distances from the disease outbreak [1], or cultural differences that differentiate one community's perception of the outbreak from another [2]. Outbreaks may affect populations whose primary language is not English.

Twitter corpora in different languages provide us with an opportunity to analyze the perception of and reaction to disease outbreaks among different populations. Twitter is a social media platform that allows users to communicate in different languages. According to Statista, there were 328 million monthly active Twitter users worldwide in 2017 [3]; two-thirds of Twitter accounts were outside the US [4]; and Twitter supported 35 or more languages [5]. The transnational nature of Twitter as a social media platform where users of different linguo-cultural backgrounds can react to the same event online in real time presents us with an opportunity in which we can examine some of the hidden assumptions of social media public health literature. Many analyses of Twitter data in public health literature only studied tweets written in English. For example, a recent systematic review on social media studies about Ebola found that all seven studies on Twitter data about Ebola covered English tweets only [6]. In a recent paper on Twitter data about Zika, while the authors studied the trends of Zika-related tweets in English, Spanish and Portuguese, respectively, they limited their content analysis to English language tweets given their own language limitation [7]. However, health-related English language Twitter content may not be generalizable to Twitter content in other languages. The assumed 'global' nature of Twitter and the analysis of a monolingual corpus of Twitter data mask the underlying diverse nature of the different linguo-cultural groups of users who communicate on Twitter using different languages [8]. We contend that it is essentially an Anglophone-centric paradigm with a hidden language bias towards English [9].

The 2015 Middle East respiratory syndrome (MERS) outbreak in South Korea provided us with a unique

opportunity to perform a cross-cultural comparison between different linguo-cultural communities of Twitter users [10]. Because Twitter users in the various Asian countries adjacent to South Korea write in different languages, by comparing Twitter corpora in different languages that are associated with a single infectious disease outbreak, we can control for time and event, and identify the differences between the corpora.

In this study, we are going to examine the following hypotheses.

1. The types of Twitter profiles of those who tweeted about MERS differed between the corpora of five different languages.
2. Except for keywords that are apparently specific to this outbreak (i.e., 'MERS', 'Korea', 'Korean', and 'South Korea'), there are unique keywords (when translated into English) used in MERS-related tweets that distinguish the corpora of five different languages.
3. The types of the top 30 Twitter profiles that received the most retweets for their MERS-related tweets differed between the corpora of five different languages.

By examining the aforementioned hypotheses, we will test if Twitter usage pertaining to the 2015 South Korean MERS outbreak differed along linguo-cultural lines and explore the implications of our findings for global health communications.

Methods

This is a cross-sectional study. We obtained a 1% random sample of Twitter data via Twitter streaming application programming interface (API) [11]. APIs are a set of protocols that enable third-party software applications to retrieve structured data from online platforms. Twitter streaming API provides streams of live feeds, and it is the most commonly used data source for Twitter research [12]. For this study, we ran Python scripts to access a 1% random sample through the sample function of Twitter streaming API [13]. From this sample, we retrieved MERS-related

Download English Version:

<https://daneshyari.com/en/article/8583125>

Download Persian Version:

<https://daneshyari.com/article/8583125>

[Daneshyari.com](https://daneshyari.com)