



## The impact of language and Systemic factors on tweeted countries of the world

H. Denis Wu 

Department of Mass Communication, College of Communication, Boston University, Boston, MA, USA




### ABSTRACT

This study is intended to unveil the difference of social mediated world via major languages and investigates the volume of tweets individual countries received during 2015–2016 in nine languages –Arabic, Chinese, English, French, German, Japanese, Portuguese, Russian, and Spanish. Shared language, country attributes, economic power, and communication resources were used in predicting country mention. The salient countries on Twitter overall are vastly diverse and vary from language to language. Based on cluster analysis, English and Japanese tweets distinguish themselves from other languages; yet the result from rank-order correlation test shows Arabic and French tweets treat countries differently from the rest. Core nations are still covered more in English- and French-language tweets. Shared language factor is found to predict well for tweets in Chinese, Arabic, Spanish, French, and German but not in English and Portuguese.

### KEYWORDS

International communication; twitter; language; world systems; information flow

Information about the world matters to people. Whether and the extent to which a country is covered in the world's media – traditional or social – is vitally significant on several grounds. In the twentieth century, communication scholars and policymakers were concerned about valence and volume of news about different parts of the world; owing to the ominous situation in the 1970s, their debates resulted in the declaration of New World Information Order (MacBride 1980). Mediated country representation is not only pivotal to how countries are understood but also contributes to external recognition and self-identity of a country's people (Masmoudi 1979). Also relevant are symbolic power of countries, perceived contest of geopolitical blocs, and the corollary of public diplomacy for both emerging and established countries. Since the 1970s scholars have confirmed that not every country can be equally covered – many structural factors would prohibit this from happening (Chang 1998; Golan 2010); some sort of selection, sifting, and distortion about the world appears inevitable.

**CONTACT** H. Denis Wu  [hdw@bu.edu](mailto:hdw@bu.edu)  Department of Mass Communication, College of Communication, Boston University, 640 Commonwealth Avenue, Boston, MA 02215, USA  [@hdeniswu](https://twitter.com/hdeniswu)

© 2020 Journal of International Communication

The development and subsequent diffusion of the internet fuelled hopes not only to conscientious news professionals but also concerned citizens worldwide. Compared to print and broadcast media, internet-based communication was considered more democratic, especially because of its unlimited space and ease to post, interact, and share all sorts of information among netizens (White 1997). Although the internet remains far from an idealised agora or Habermasian public sphere, social media is poised to transcend the dominance of traditional gatekeepers, wholesale information brokers – such as Thomson-Reuters – and, other government-sponsored news agencies.

While many control mechanisms and monitoring programmes implemented by authoritarian regimes are in place to block the creation, sharing, or retrieval of ‘sensitive information’ (Taubman 1998), some online platforms have demonstrated an ability to circumvent such controls, even if only briefly and non-continuously (Groshek 2010). Furthermore, online social media – Facebook, YouTube, and Twitter in particular – have become a major source of news for many people (Matsa and Shearer 2018), especially during such far-reaching, transnational crises as the Arab Springs (Lotan et al. 2011). This trend is pronounced among the young who do not consume international news on a regular basis from traditional media (Mitchell and Page 2015).

Relying exclusively on social media for news can lead to many issues, one of which, for keeping up with the world, is language. Languages used in the internet may not necessarily be the same as the languages used in the real world. Various sources (e.g., Accredited Language Services and Statista<sup>1</sup>) found that the ten most used languages on the internet as of 2017 are English, Chinese, Spanish, Arabic, Portuguese, Malay, French, Russian, Japanese, and German. It would be epistemologically interesting to compare the worlds discoursed on social media according to different languages; furthermore, identifying the pattern and focus of countries in Twitter through different languages would shed new light on multifaceted issues rarely explored. Given the majority of existing literature focuses on media content in one language – often English, research based on multiple languages of media content is in need and might unveil findings with greater validity.

In addition to serving as a hub many people rely on for crucial information, social media is simultaneously the venue where netizens opine, form groups, and share their views about the world’s countries. It is feasible for marketers, policy makers, and world leaders to gauge what the world’s citizens are thinking by collecting data stemmed from social media and mining the content to identify trends and patterns (Morozov 2011). Given the incessant interest in country image in the media (Servaes 2012), it would be rewarding to explore, first, how much Twitter users are thinking about each of the countries in the world. Thus, examining the frequency of content created and posted explicitly about individual countries would be a good start.

Social media content has been constantly monitored by private sectors and political entities worldwide. This new medium has also become a critical channel for governments to anchor target audiences and amplify instrumental national messages and images (Khatib, Dutton, and Thelwall 2012). However, the majority of social media platforms are proprietary and not readily available to researchers even though it is extremely important to systematically examine the content generated on social media. Given social media's increasing influence and its unique role in informing and connecting people around the world, it is imperative for researchers to investigate the social-mediated world as well as the underlying forces behind the content. So far, it is still a void in the literature of international communication.

Given this scenario, I set out to investigate the most fundamental questions about the world's coverage under the social media context. First of all, I would like to explore what countries are mentioned most frequently in a social network platform, Twitter? My further questions include: What differences are there among the tweets in major languages about the world's countries? What are the predictors of mentions of countries in tweets of different languages? The answers will shed light on the new era of international communication via social networks participated by worldwide citizens instead of exclusively news professionals. The findings of this investigation on social media-yielded country coverage will result in abundant implications and ramifications for the media, governments, world politics, as well as international relations. Since this research lies in uncharted waters, I based the following review of relevant literature on traditional media. The review of the usual suspects of news determinants in the traditional media hopefully would assist the inquiry and subsequently result in fresh findings on the social media.

## Review of literature

*The world, according to media.* The internet, along with other relevant technologies, generated high hopes for information and content about individual countries to flow more freely across borders. The traditional barriers and hurdles for news to circulate worldwide, such as geographic distance (e.g., Chang, Shoemaker, and Brendlinger 1987), limited space, time constraint, and gatekeepers' decision (Peterson 1981), matter much less with social media as the medium and average netizens as authors. Yet, studies showing high similarity of news between traditional media and web counterparts (Wu 2007) and significant intermedia agenda-setting between the media in elite and emerging countries (Guo and Vargo 2017) dampen the hope of the internet's independence of information flow.

In addition, empirical findings (e.g., Garrett 2009) indicated that social media can serve as echo chambers of mainstream media. Social media participants – due to average participants' limited resource and opinion-centered mode –

tend not to initiate their own investigation and post original discovery. Given the found tendency, it is worth investigating whether the pattern of countries mentioned in Twitter mirrors the pattern found on the traditional news about the world. The forces of international coverage had been extensively investigated and reported in the existing literature, which roughly fall into language and culture affinity, country attributes, and communication and economic resources.

**Language affinity.** It is reasonable that countries – like human beings – that share the same language communicate better. Therefore, languages can result in different versions of the world presented and discussed in social media. The past literature on news flow (Johnson 1997; Kariel and Rosenvall 1983; Kim and Barrett 1996) supports this point of view, indicating the impact of shared languages on the amount of information about certain countries flowing across national borders. The present study goes further by examining whether average Twitter users of different languages around the world actually echo the media professionals' inclination. As researchers (Hong, Convertino, and Chi 2011) already pointed out the importance of language determinant in Twitter, which provides a strong rationale to pursue this study.

Another aspect of language influence in social media can be ensconced in a larger context. The influence of a specific language on its users' understanding and interpretation of the world can be highly related to the Sapir-Whorf hypothesis (Kay and Kempton 1984), which posits language's impact on humans' experience and meaning extrapolation of their physical world. Social media output from various languages can provide a natural setting to test the linguistic influence on cognition and discourse about individual countries. Including the tweets in nine languages about the world's countries allows the researcher to explore the potential impact of language; it also permits examining the extent of similarity about the world to which the nine different languages present.

RQ1: How do tweets about countries of the world in different languages vary?

Cultural affinity can serve as a more general term that makes connection and communication work better. Cultural affinity not only encompasses language but also many other factors, including a country's colonial background. The common nexus with other countries may play an instrumental role in determining the discussion of countries in social media. Dupree (1971) and Skurnik (1981) found that countries within a colonial group tend to cover each other more. In this study, countries with one of the eight languages<sup>2</sup> as one of their official languages or belonging to one of the six colonial groups would be positively associated with their tweet mentions. Thus, two hypotheses are made:

H1: Shared language is related to the amount of tweet mention for individual countries.

H2: Belonging to the same colonial group is related to the amount of tweet mention for individual countries.

**Country attributes.** Given scarce literature regarding social media mentions about individual countries, news flow across national borders can provide a useful guidance. Although the mechanism of news coverage about individual countries differs from the counterpart of tweets, the similarities reside in the subject of the utterances centring on individual countries and in the nature of publicly displayed information. The determinants of international news also are related to the present investigation because attributes of a country can influence the decision making of news gatekeepers. Editors and foreign correspondents select and focus on countries they deem important and newsworthy; likewise, users of social media take on the role of gatekeepers and post in accordance with their perception of countries. Hence, the *attributes of countries* can be important to the amount of country mention on Twitter, including press and political freedom, population, and status in the world. It should be reasonable that some of the country attributes overlap with economic resources of a country, which will be elaborated in the next part.

A country's level of press freedom is associated with traditional news flow volume (Chang, Shoemaker, and Brendlinger 1987) – the freer the country, the more information it generates. A repressive environment for journalists to work in could be also hard for average Twitter participants – it would be extremely taxing to gather, investigate, and disseminate pivotal information via both traditional as well as social media. Another connected concept, political freedom, may influence the pattern of netizens' behaviour. When governments are overly sensitive on the information about their countries, negative impact on social media coverage and number of tweets is likely to occur.

Wallerstein's world systems theory (WST) and Galtung and Ruge's (1965) structural theory of foreign news can lend theoretical support here. Both maintained that a country's echelon in the world can generate a profound influence on its media coverage. By incorporating the impact of WST found in international news coverage (Chang 1998) and the updated framework of core, semi-periphery, and periphery countries (Babones 2005), the investigation of the structural influence in nine languages of the social media output can be theoretically fruitful. Specifically, it is conceptually meaningful to detect whether social-mediated content about individual countries actually transcends from the news structure, i.e., the countries in the core were more likely than the semi-peripheral and peripheral strata to be talked about in tweets. Additionally, the present study examined whether average social media participants of nine different languages share the mindset of mainstream journalists found in the West.

Given the aforementioned literature, I hypothesise that these country attributes can contribute to a country's presence on tweets. Instead of listing each country attribute, the following hypothesis serves the function of research guidance.

H3: The attributes of a country can predict its mention on Twitter.

**Economic resources.** A country's population and geographic size are two commonly identified factors that affect its chance of being covered in the traditional media. Studies on international news (e.g., Dupree 1971; Kareil and Rosenvall 1984; Westerståhl and Johansson 1994) found that these two factors are positively related with coverage; whereas others (e.g., Larson 1979) did not produce consistent and significant findings. These factors therefore merit further inspection in the new context.

A sizeable body of the literature (Ahern 1984; Kareil and Rosenvall 1984; Kim and Barnett 1996) indicates that a country's economic output can be conducive to how it is covered in other countries. It is worth investigating to what extent economic indicators are perceived important to Twitter participants. Aside from the macro-level statistics of economic resource, others (e.g., Wu 2000) also included the predictors of economic resource average citizens have and income equality, which seem pertinent when Twitter, a social media platform, is the subject of investigation. The booming of internet-based new technologies has yet made the divide between the haves and have-nots diminished (Blank 2017; Hargittai 2004). To participate in social media, users must be able to afford a slew of resources. Therefore, a country's aggregate economic resource and income equality may influence its citizens' participation and the mention of that country in social media. Given that, the following hypothesis centres on the influence of economic resource of countries.

H4: The economic resource of a country can predict its mention on Twitter.

**Communication resource.** The literature of news flow has identified the impact of communication resource on how a country is covered by others. For example, increasing transnational news flow was attributed to better infrastructure preparedness and technological availability (Larson 1979, 1984). The presence of news services also can elevate the odds of being covered worldwide (Wu 2000). Because Twitter is such a platform anyone in the world can use it as long as internet access is available, internet penetration of a country may contribute to the volume of tweets.

In addition to the access to the internet, telecommunication penetration should be positively linked to social media use as well (Seo and Thorson 2012). With significant social media activities taking place via mobile phones – particularly smartphones (Hwang and Park 2013), the penetration rate of mobile phones in any country will be positively related to the extent to which that country is discussed on Twitter. Simply put, the more available and affordable access people have to participate on Twitter, the more likely their countries will be mentioned on Twitter.

A common factor that may influence communication access is the freedom to participate in social media. Authorities can utilise an assortment of means to censor, block, and intercept sensitive information; others, like China and North Korea, ban U.S.-based social media entirely. Therefore, the level of

social media access a country's citizens enjoy may be highly relevant to the salience of that country mentioned on Twitter. However, this factor has been included in the category of country attributes and thus will not be repeated here.

H5: The communication resources of a country can predict its mention on Twitter.

## Method

This study incorporates national-level data of individual countries in the world during one-year time frame (October of 2015 to September of 2016). The total number of tweets each country received was obtained by manually searching the tweets via Crimson Hexagon, a social media analytics service, using each country's name in each of the 9 languages. For instance, Afghanistan, in addition to its English name, its Arabic name–أفغانستان, Spanish version–Afganistán, Portuguese version–Afeganistão, Japanese name–アフガニスタン, Russian name – Афганистан, and Chinese name–阿富汗, were entered one at a time for the year-long date range in all tweets. For countries with more than one commonly used name such as the United Kingdom and the U.K., separate queries were entered for each term.

To gauge volume of tweets, this study used the software ForSight in Crimson Hexagon. ForSight allows for monitoring, benchmarking, and tracking keywords on Twitter. Using the Guided Keywords function in the monitor, the researcher entered each country name, the date range, and each of the languages to retrieve output. ForSight then generated a report that provides the volume of tweets for each search, which was recorded one by one into a separate spreadsheet.

Many country-oriented statistics were downloaded from the Central Intelligence Agency's World Factbook (<http://www.cia.gov/library/publications/the-world-factbook/>). The values of individual countries include geographic size, official language, colonial past, GDP, and GINI coefficient. Media variables included are number of mobile telephone and number of internet users. Additional predictor variables from other sources include the ratings of political freedom and press freedom, collected from the Freedom House. The data of fixed broadband and mobile phone penetration in each country were obtained from the International Telecommunication Union.

Each country's ranking in the world system was also used to position individual countries within the world system as core, semi-periphery, or periphery countries (Wallerstein 1974). This structure was based on the strength of governments and economic systems in which 'core economies (high wages, high profits, capital-intensive) are bound in a lopsided exchange with the periphery economies (low wages, low profits, and labor-intensive)' (Burkhart and Lewis-Beck 1994, 903). Operationalisation of country identification was the result of synthesising two comprehensive studies. Chase-Dunn, Kawano, and Brewer (2000) evaluated national integration in the world system from 1795 to 1995.

The second (Babones 2005) examined the mobility of countries between the three levels from 1975 to 2002. Both studies were used to form a more updated categorisation system. Of the 210 countries included in the study, 21 belong to core countries, 24 are categorised as semi-periphery, and the rest are either periphery or not coded.

## Results

The result generated from the searches of each of the countries during the one-year period through the full Twitter fire hose generated a total of 2,749,973,772 tweets for the 210 countries included in this study. The number of tweets that mentioned countries in English (1143 million) outnumbers all other studied languages. The number of Spanish tweets (536 million) is the runner up, followed by the number of Japanese tweets (514 million). The least numbers of country tweets came from Chinese (10 million) and Russian (nearly half a million), which are surprising but might indicate the influence of governmental intervention on Twitter use. Looking at the ranking of countries mentioned in the tweets in all and each of the nine languages (see Table 1), one can instantaneously notice the level of diversity – particularly across the nine languages – is truly unprecedented and justifies the examination of the new medium. On the one hand, core countries like the U.S., France, and Germany are still on the lists; on the other, non-core countries – Thailand, Brazil, Egypt, Venezuela, India, Mexico, and Syria – are also salient on this social media platform. The results suggest that the pictures of the world on Twitter are dramatically different from language to language; moreover, as a whole, they greatly differ from the counterpart presented by traditional news media, which has been grounded on a long-standing world systems perspective.

Considering RQ1, from the top 20 countries illustrated in Table 1, it is readily apparent there is a wide range of prominent countries in different languages of tweets. This indicates that distinct interest and unique perspective stem from the very language used in the identical platform. Starting with tweets in English, five English-speaking countries – the United States, the United Kingdom, India, Canada, and Nigeria – were on the list of ten most mentioned countries. The other most frequently mentioned in English tweets are China, Japan, Jordan, and France. Another group of prominent countries belongs to the Middle-East region – in addition to Jordan, Syria, Turkey, Pakistan, Iran, and Saudi Arabia were all on top 20 list.

When the salient countries from tweets in other eight languages are examined, a dramatic shift emerges. Regional influence could be at work; cultural affinity seemingly exerts its influence too. Six of the 10 most tweeted countries in Chinese are in the Asia-Pacific region while the rest four are core Western countries. More than half of the top 20 Twitted countries in Spanish are in Latin America and the rest are traditional global powers such as the U.S.,



**Table 1.** The 20 most mentioned countries in tweets of 9 languages.

English	Chinese*	Spanish	Arabic	Portuguese
United States	China	Venezuela	Egypt	Brazil
United Kingdom	Japan	Mexico	Syria	Qatar
India	United States	Argentina	Yemen	United States
China	Hong Kong	Spain	Kuwait	Germany
Japan	Taiwan	Colombia	Iraq	France
Jordan	South Korea	United States	Iran	Argentina
Canada	United Kingdom	Chile	Turkey	El Salvador
France	North Korea	Cuba	Bahrain	Japan
Nigeria	Germany	Brazil	Qatar	China
Syria	Canada	Ecuador	Morocco	Iran
Australia	France	China	Lebanon	Italy
Turkey	Russia	France	Mali	Venezuela
Pakistan	India	Peru	Russia	Grenada
Iran	Singapore	Panama	United States	Spain
Germany	Turkey	Uruguay	Palestine	Peru
Mexico	Syria	Russia	Algeria	Mexico
Brazil	Vietnam	Syria	France	Russia
Italy	Thailand	Germany	Libya	Canada
Saudi Arabia	Iran	Italy	Israel	United Kingdom
Israel	Spain	Japan	Tunisia	Australia

\*Simplified Chinese characters were used for retrieval.

French	Russian	German	Japanese**
France	Russia	Germany	Thailand
Syria	Togo	Turkey	Aruba
Germany	Ukraine	Switzerland	United States
Turkey	Turkey	Iceland	Russia
Portugal	China	Japan	India
Belgium	Syria	Austria	France
Morocco	Egypt	France	Germany
Italy	France	Russia	Chile
Spain	Germany	China	Syria
Canada	Kazakhstan	Italy	United Kingdom
Algeria	Iran	Poland	Italy
Russia	Poland	Greece	Iran
Switzerland	Israel	Ukraine	Spain
China	Japan	Israel	Brazil
Monaco	Italy	Spain	Australia
Japan	Canada	Iran	Turkey
Israel	Spain	Sweden	Hong Kong
Gabon	Georgia	Nigeria	Canada
Tunisia	Azerbaijan	Portugal	Vietnam
Burundi	Iraq	Belgium	Philippines

\*\*In most cases, Hiragana and Katakana of country names were used for retrieval.

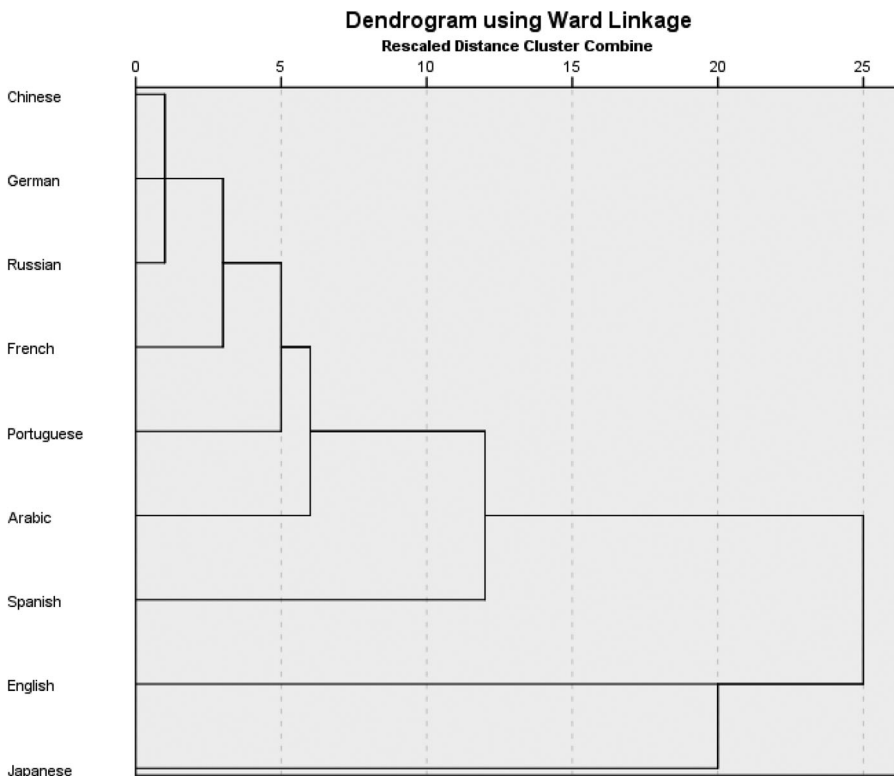
China, and France. It is intriguing to observe the 12 most mentioned countries in Arabic are in the Middle-East region – Russia and the U.S. only take the 13th and 14th spot, respectively.

As found in the English tweets where the U.S. garners most mentions, the most populous country of the world, China, is mentioned most frequently in the Chinese tweets. The pattern that the most populous country in the language group turns out the most mentioned one on Twitter is confirmed in five other languages: the Arabic, Portuguese, French, Russian, and German tweets have mentioned most frequently on Egypt, Brazil, France, Russia, and Germany, respectively. Shared language appears to play a role in shaping the list of most

mentioned countries in most of the languages examined. Most of the prominent countries covered by English, Spanish, Arabic, French, and Russian tweets appear to share each of the languages – a more rigorous regression test would help verify this language-tweet volume association.

It is important to note that of the 20 most frequently tweeted-about countries from the 9 languages examined, significant percentage of them are from periphery and semi-periphery countries. This interesting finding, situated from full fire hose access to one year of tweets, seems at odds with the existing literature that says news coverage about countries reflects the world system. Even if these findings are not perfectly generalisable to all social media, they are nonetheless indicative of the pictures of the world presented and discussed by average users in the social media era could be entirely different from the one gone by.

The second half of RQ1 aims to unveil the relationship among the tweets that mention countries in nine languages. Based on the result of a cluster analysis that used Ward distance as the measuring unit to examine the tweets of nine languages (see [Figure 1](#)), six of the languages are found to be within five units of Ward distance apart from each other – Chinese, German, Russian, Portuguese, French, and Arabic. The tweets in these six languages, as a cluster, are quite different from the tweets in the other three languages: Spanish (roughly



**Figure 1.** Linkage among the 9 languages on Twitter.

5 units of Ward distance away), Japanese (15 units away), and lastly English (20 units away). The first group of six languages, despite their visible differences in countries mentioned (as seen in Table 1), do not vary as much as their aggregate difference from the rest of three languages. In other words, the English tweets are significantly different from the rest of tweets in other languages, followed by the Japanese and Spanish tweets. This suggests there is a profound gap separating the English and non-English Twitter worlds. However, if the relationships among the country tweets in 9 languages are examined with rank-order correlation, then the seven highest correlation coefficients are the dyads of Chinese-English, Spanish-Portuguese, English-German, Chinese-German, German-Russian, Chinese-Russian, and Chinese-Japanese (highlighted in Table 2). The aforementioned result shows that the tweets in Arabic and French treat the world's nations differently from the rest.

To investigate further and provide more specificity and explication for identifying the factors that influence the presence of countries on Twitter, all four groups of independent variables were taken into consideration and modelled in an hierarchical ordinary least squares (OLS) regression model with the number of tweets for each country as the dependent variable. The results of all factors from regression model are summarise in Table 3, where readers can observe the four blocks of predictors representing language factor, country attributes, economic factors, and communication resource.

The regression results show that, once again, tweets of different languages are under different forces for their respective coverage of world's countries. The overall prediction model explains best for Chinese tweets ( $R^2 = .883$ ), followed by English tweets ( $R^2 = .623$ ) and Spanish tweets ( $R^2 = .477$ ). The prediction models for the tweets in four other languages (Arabic, Russian, German, and Portuguese) are moderately robust; whereas the overall model for French tweets is weak, but still statistically significant. The only prediction model that does not result in significant level is for Japanese tweets, perhaps due to the lack of shared language variable.

If one focuses on each block's net contribution, little consistency across the nine languages emerges. Regarding the language block (first panel), all except

**Table 2.** Correlation among the 9 languages of tweets.

	English	Portuguese	Spanish	Chinese	Arabic	German	French	Russian	Japanese
English	1								
Portuguese	.621**	1							
Spanish	.603**	.754**	1						
Chinese	.762**	.667**	.632**	1					
Arabic	.626**	.462**	.395**	.692**	1				
German	.742**	.627**	.620**	.735**	.565**	1			
French	.590**	.546**	.569**	.614**	.623**	.617**	1		
Russian	.596**	.523**	.504**	.724**	.520**	.730**	.547**	1	
Japanese	.658**	.601**	.592**	.718**	.553**	.660**	.513**	.593**	1

Presented in the table are *Spearman's rho*.

\*\*Significant at .001 level (2-tailed).

**Table 3.** Predicting country mention on Twitter.

Predictor variables	English tweet	Chinese tweet	Arabic tweet	Spanish tweet	Portuguese tweet
Language <sup>#</sup>	.113	.241***	.435***	.529***	.160
British colonial past	.002				
Spanish colonial past				-.021	.143
Portuguese colonial past					
R <sup>2</sup> Change	.015	.288***	.255***	.353***	.105***
Core country	.265***	.089*	-.065	.035	-.047
Semi-periphery country	.123	-.089*	.133	.292**	.188*
Press freedom	-.224	-.040	-.698	-.106	-.590
Political freedom	.332	.053	.864*	.163	.548
R <sup>2</sup> Change	.272***	.081**	.040	.068**	.067*
GDP	1.060***	-.097	.302	.049	-.155
GDP per capita	.025	-.061	.089	-.001	.073
Population	1.477	-2.169***	1.230	1.612	-.103
Size	.141	-.346***	.236	.352*	.335*
Gini	-.035	.037	-.155	-.139	.016
R <sup>2</sup> Change	.297***	.376***	.040	.027	.082*
Internet penetration	-.659*	1.462***	.121	.856*	.496
Mobile phone penetration	-1.328	1.673**	-1.714	-2.622*	-.356
R <sup>2</sup> Change	.039**	.138***	.011	.029*	.010
R <sup>2</sup>	.623	.883	.347	.477	.263
Sig	< .001	< .001	< .001	< .001	< .001

# One of the eight languages as official language of the country was entered into the regression model when tweets in that particular language were examined.  
NOTE: Beta coefficients are presented in the cells.  
\**p* < .05; \*\**p* < .01; \*\*\**p* ≤ .001.

Predictor variables	Russian tweet	French tweet	German tweet	Japanese tweet
Language <sup>#</sup>	.230*	.420**	.507***	
French colonial past		-.208		
German colonial past			.021	
Russian East Block past	-.028			
R <sup>2</sup> Change	.147***	.084**	.265***	
Core country	-.118	.221*	-.040	.084
Semi-periphery country	.011	.042	.080	.177
Press freedom	-.288	-.820	-.869*	-.347
Political freedom	.273	.932*	.924*	.503
R <sup>2</sup> Change	.009	.075*	.030	.062
GDP	-.290	-.028	.081	.252
GDP per capita	-.020	.005	-.062	.047
Population	1.264	-1.321	-1.655	-1.389
Size	.789***	-.223	-.187	.222
Gini	-.157	-.121	-.223*	.006
R <sup>2</sup> Change	.201***	.015	.055	.011
Internet penetration	.820*	.169	.235	-.774
Mobile phone penetration	-2.123	1.305	1.529	2.054
R <sup>2</sup> Change	.024	.015	.026	.022
R <sup>2</sup>	.382	.190	.376	.095
Significance	< .001	< .001	< .001	.331

# One of the nine languages as official language of the country was entered into the regression model when tweets in that particular language were examined.  
NOTE: Beta coefficients are presented in the cells.  
\**p* < .05 \*\**p* < .01 \*\*\**p* ≤ .001

English<sup>3</sup> turn out statistically significant predictors; that is, shared language does not pave the way for those English-speaking countries to appear more frequently in English tweets even though sharing any of the other six languages certainly do.

Given the majority of the examined languages provide evidence for the influence of shared language on country mention, H1 is supported. This finding suggests either English as a universal language on Twitter or Twitter users of the other languages are more inward looking – or both. English tweets' distinctiveness echoes the very finding from the cluster analysis. Another major difference in predicting English and other language tweets stems from the influence of world system standing. It is clear that core countries are more likely to be mentioned in English tweets ( $B = .265$ ) than in tweets of other languages. This is a uniquely elitist tendency of English tweets about the world.

Also within the first block, belonging to the same colonial group, unlike shared language, does not generate significant result. Not only do all six colonial groups fail to produce statistical significance, some of them even yield negative – despite statistically insignificant – coefficients in predicting country mention in Twitter. Based on this finding, H2 is rejected. The colonial background's influence on mediated presence of countries turns out truly a thing of the past.

Moving down to the block of country attributes, of the nine languages, Arabic, Russian, German, and Japanese tweets are not explained well by country attributes. Overall, country attributes seem to generate a positive yet very moderate influence on tweet volume with the exception of English tweets. Within the block, world system ranking of countries does not always translate well into tweets about countries – core countries fare better only in English, French, and Chinese tweets. Interestingly, semi-periphery countries receive more tweets in both Spanish and Portuguese tweets. Freedom levels of country merely lead to sporadic influence across the nine languages. Overall, given country attributes in the five of nine languages (based on the block's  $R^2$  values) result in statistical significance in predicting tweets, H3 is supported.

H4 posits the influence of a country's economic clout on tweets volume. Of the nine languages, only four – Chinese ( $\Delta R^2 = .376$ ), English ( $\Delta R^2 = .297$ ), Russian ( $\Delta R^2 = .201$ ), and Portuguese ( $\Delta R^2 = .082$ ) – indicate the influence from economic resource on tweet volume. Based on this finding, H4 is rejected. It is intriguing that the five predictors in this block do not generate any meaningful pattern. A country's GDP only predicts well on English tweets, whereas geographic size renders larger tweet amount only for Russian, Spanish, and Portuguese. Lastly, the economic factors have little to do with the tweets in Arabic, French, and Japanese.

As for the influence of communication resources, the block predicts significantly for Chinese ( $\Delta R^2 = .138$ ), English ( $\Delta R^2 = .039$ ), and Spanish ( $\Delta R^2 = .029$ ) tweets. With only three of the nine languages showing significance of communication predictors, H5 is rejected. Internet penetration is conducive to a country's mention in Chinese, Spanish, and Russian tweets; interestingly, it negatively influences tweet volume in English. Mobile phone penetration also helps country mention in Chinese tweets, but it renders negatively on Spanish tweet volume.

## Discussion

This study sought to extend the inquiry of country representation in the media and examine which countries are more salient on Twitter in nine common languages and which factors can help explain why certain countries are more noticeable on the popular platform. In so doing, this study sheds light on contemporary and transformative media representation and presentation of the world and updates the theory of underlying forces that shape individual country presence on social media. This study on mediated world not only tackles historically important frameworks of international news flow, world systems theory, country attributes, and economic and communication resources, it also incorporates the pivotal language factor, which fills the void of past scholarship.

The finding that the most prominent countries on Twitter are dramatically different from those regularly reported by mainstream media shows that social media (at least Twitter) in different languages have presented distinct worlds to the world. This phenomenon not only challenges the dominance of Western media conglomerates on the representation of various countries but also shifts and shapes the worldviews of social media users. News content in English used to be the primary researched language, whose content supposedly travels across national borders and reaches worldwide audiences. Today, English is merely one of many widely conducted languages in social media – a fact that truly commands a more inclusive approach. However, the English language's potential connectedness (Hale 2012) with other languages in social media is worth investigating. The shifted media-scape as well as the differed mediated worlds cannot be vividly illustrated without this multi-language study.

The findings also provide implications for stakeholders in international news. Countries, for one, not only should be keen to their portrayal on traditional print and broadcast media but also heed to the discourse on a host of social media platforms – including Twitter. Both information and viewpoints opined by participants toward any given country on social media matter. The lessons for public diplomacy practitioners to take away from this study, likewise, are vital. Missing the reportage in any key language would prove detrimental. On the other hand, users of social media ought to be acutely aware of the different pictures of the world the interactive platform delivers for them. Despite the free, more democratic, and engaging façade, it is premature to conclude whether the 'socially mediated' worlds are closer to the real world in terms of comprehensiveness, depth, value, and interpretation than the ones presented by traditional news media such as broadcasters and newspapers.

While the issue of information gap between the covered countries and the neglected ones may remain in the future, the unveiled foci of tweets in different languages suggest that social media may have led to the abatement of a single, operational world system that was observed in the twentieth century.

Indeed, the role of social media in cultivating increased level of national visibility has taken on additional importance as the internet 2.0 fulfils an ascending role in sharing and mobilisation (Gurevitch, Coleman, and Blumler 2009) of pivotal information all over the world. Just as older media technologies such as print and television structure political campaigns, knowledge, and interaction (Postman 1985), online media are reshaping political participation and cultural engagement worldwide. The global impact of the internet being both more accessible and socially interconnected with culturally relevant (or language-centered) content, however, has yet been analysed. For example, the fact that Japan is a country with one of the largest Twitter participants (more than 50 million) translates not only to the country's prominence in tweets of every language, but also sends Thailand, a much tweeted country in Japanese, to the top of most mentioned country. This multi-language comparison of tweets makes a unique contribution to understanding the bounds of any given language in getting the world informed but also why and how that may be the case. It provides a glimpse of how user-generated, social network-based country coverage may appear in the future.

This study sets out to explore the predictors of tweet volume of world's countries and concludes with the shared language factor as the most pervasive. Other predictors – bracketed in country attributes, economic and communication resource – are less consistently significant and their influences vary from language to language. This line of research definitely needs further examination. For one thing, the accuracy and aspects of a given country mentioned in the tweets have yet been examined in this study. More nuanced inspection of the country mention would be immensely helpful. Moreover, the affective dimension of mentioned countries in tweets is equally significant – e.g., whether the mention of a country is positive, negative, or mixed; what aspects of the country are invoked and emphasised. Therefore, these are the shortcomings of the present study and can be the direction of future studies on social media.

Another potential research agenda of Twitter mentions is whether and to what extent the tweets inform more people and elite decision makers in various countries and subsequently drive real world change. It would be interesting to examine whether the heavily discussed countries and their issues (in either negative or positive light) lead to more attention among the world's leaders and whether globally intertwined issues such as climate change, trade, refugees, and drug trafficking can result in more cooperation and collaboration. As a global medium, Twitter does have a lot of potential capacity toward facilitating the dialogue of concerned citizens and improving extant conditions.

The material as well as the mediated worlds, have not been created equal (Chang 1998). Country attributes are still pivotal to tweet mention, as shown in the examined languages. The present study, however, suggests that a more diverse, less hierarchical version of the world has been presented, which is

positive. On the other hand, it remains unclear how many individuals around the world access these tweets and what kind of impact socially mediated messages about the world have generated on the world – for example, whether and to what extent agenda-setting effect of tweets on people’s perception of nations happens (Wanta, Golan, and Lee 2004). Due to the lack of opinion assessment, the above examination of tweets’ agenda-setting effect cannot be executed. The present study only demonstrates shifts in focus and dissemination of information about countries. Although a watershed of transnational information flow appears to have emerged, this study only verifies documented determinants of country coverage also exert their influence on Twitter.

As a starting point for theory development in the social media era, this study has analysed predictors of country-related tweets in nine languages. One of the primary findings supports the Sapir-Whorf hypothesis in that the language factor plays a significant role in shaping the structure of social media coverage about the world. With the exception of English, the language conducted by Twitter users influences the countries they mention, reflect on, discuss about, and share – i.e., to a great extent, the choice of language influences the version of social mediated world. One distinction between this study and others in supporting the Sapir-Whorf hypothesis is that for this case the socio-cultural heritage in a given language appears to trigger the impact on users’ cognitive evidence on social media. Future studies may extend the number of languages and incorporate geographic locales of social media users to confirm the finding.

It is intriguing that economic resource and clout of countries that used to dominate international news in traditional media are only sporadic in the tweets of the nine examined languages. Compared to the findings of country coverage in traditional news, the overall results of this study suggest that the gatekeepers of social media are different and that the new platform may have brought on a democratisation effect with respect to country mentions. Given the inconsistent and moderate impact of communication resource on country mentions, it is hopeful that the barrier of internet-related technologies and access would not present a roadblock for this welcoming trend.

There are limitations of this study that should be reported. For a better understanding of country representation in social media, other widely adopted platforms such as Facebook, Instagram, WeChat, and YouTube merit scholars’ attention too. Visuals in which countries are mentioned and presented would be equally important to investigate. In addition, the interaction and networks on social media that are formed across national borders should be included in future endeavours. It could be rewarding to see how the connection of social media content is shaped and transforms how people in different locales understand and discuss the world. This study is only a start to examine countries presented and discoursed in the social media arena.



## Notes

1. <https://www.statista.com/statistics/262946/share-of-the-most-common-languages-on-the-internet/>
2. Japan is the only country in the world that recognizes Japanese as its legal language. Therefore, only 8 languages were included in the regression analyses in this study.
3. Again, the language block was not included in the prediction model of Japanese tweets.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Notes on contributor

*H. Denis Wu* is a professor of communication at Boston University (Boston, Massachusetts, USA). He has published extensively in the areas of international communication and political communication. Via this study, he explored how people from different parts of the world may be concerned and connect with the world distinctly. It is hopeful that the findings reported here may contribute to improving world communication and peace.

## ORCID

*H. Denis Wu*  <http://orcid.org/0000-0002-9657-1734>

## References

- Ahern, T. J. 1984. "Determinants of Foreign Coverage in Newspapers." In *Foreign News and the New World Information Order*, edited by RL Stevenson and DL Shaw, 217–236. Ames, Iowa: Iowa State University Press.
- Babones, S. J. 2005. "The Country-Level Income Structure of the World-Economy." *Journal of World-Systems Research* 11 (1): 29–55.
- Blank, G. 2017. "The Digital Divide among Twitter Users and its Implications for Social Research." *Social Science Computer Review* 35 (6): 679–697. doi:10.1177/0894439316671698.
- Burkhart, R. E., and M. S. Lewis-Beck. 1994. "Comparative Democracy: the Economic Development Thesis." *American Political Science Review* 88 (04): 903–910.
- Central Intelligence Agency. The World Factbook. <https://www.cia.gov/library/publications/the-world-factbook/>.
- Chang, T.-K. 1998. "All Countries Not Created Equal to Be News: World System and International Communication." *Communication Research* 25 (5): 528–563. doi:10.1177/009365098025005004.
- Chang, T.-K., P. J. Shoemaker, and N. Brendlinger. 1987. "Determinants of International News Coverage in the U.S. Media." *Communication Research* 14 (4): 396–414. doi:10.1177/009365087014004002.
- Chase-Dunn, C., Y. Kawano, and B. D. Brewer. 2000. "Trade Globalization Since 1795: Waves of Integration in the World-System." *American Sociological Review* 65 (1): 77–95.
- Dupree, J. D. 1971. "International Communication: View From "A Window on the World"." *International Communication Gazette* 17: 224–235.

- Galtung, J., and M. H. Ruge. 1965. "The Structure of Foreign News." *Journal of Peace Research* 2: 64–91.
- Garrett, R. K. 2009. "Echo Chambers Online?: Politically Motivated Selective Exposure among Internet News users1." *Journal of Computer-Mediated Communication* 14 (2): 265–285. doi:10.1111/j.1083-6101.2009.01440.x.
- Golan, G. J. 2010. "Determinants of International News Coverage." In *International Media Communication in a Global Age*, edited by G. Golan, T. Johnson, and W. Wanta, 125–144. New York: Routledge.
- Groshek, J. 2010. "A Time-Series, Multinational Analysis of Democratic Forecasts and Internet Diffusion." *International Journal of Communication* 4: 142–174.
- Guo, L., and C. J Vargo. 2017. "Global Intermedia Agenda Setting: A Big Data Analysis of International News Flow." *Journal of Communication* 67: 499–520.
- Gurevitch, M., S. Coleman, and J. G. Blumler. 2009. "Political Communication—Old and New Media Relationships." *The ANNALS of the American Academy of Political and Social Science* 625: 164–181.
- Hale, S. A. 2012. "Net Increase? Cross-Lingual Linking in the Blogosphere." *Journal of Computer-Mediated Communication* 17 (2): 135–151. doi:10.1111/j.1083-6101.2011.01568.x.
- Hargittai, E. 2004. "Internet Access and use in Context." *New Media & Society* 6: 137–143.
- Hong, L., G. Convertino, and E. H. Chi. 2011. *Language matters in twitter: A large scale study*. Paper presented at the Fifth international AAAI conference on weblogs and social media.
- Hwang, Y., and N. Park. 2013. "Digital Divide in Social Networking Sites." *International Journal of Mobile Communications* 11 (5): 446–464. doi:10.1504/IJMC.2013.056955.
- Johnson, MA. 1997. *Predicting news flow from Mexico*. " *Journalism & Mass Communication Quarterly*. Vol. 74, 315–330.
- Kareil, H. G., and L. A. Rosenvall. 1984. "Factors Influencing International News Flow." *Journalism Quarterly* 61: 509–516.
- Kariel, HG, and LA Rosenvall. 1983. "Cultural affinity displayed in Canadian daily newspapers." *Journalism Quarterly* 60: 431–436.
- Kay, P, and W Kempton. 1984. "What is the Sapir-Whorf hypothesis?" *American Anthropologist* 86: 65–79.
- Khatib, L., W. Dutton, and M. Thelwall. 2012. "Public Diplomacy 2.0: A Case Study of the U.S. Digital Outreach Team." *The Middle East Journal* 66 (3): 453–472.
- Kim, K, and GA Barnett. 1996. "The determinants of international news flow: A network analysis." *Communication Research* 23 (3): 323–352.
- Larson, J. F. 1979. "International Affairs Coverage on US Network Television." *Journal of Communication* 29 (2): 136–147.
- Larson, J. F. 1984. *Television's Window on the World: International Affairs Coverage on the U.S. Networks*. Norwood, NJ: Ablex.
- Lotan, G., E. Graeff, M. Ananny, D. Gaffney, I. Pearce, and d. boyd. 2011. "The Arab Spring The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions." *International Journal of Communication* 5: 1375–1405.
- Macbride, S. 1980. *Many Voices, One World: Towards a New, More Just, and More Efficient World Information and Communication Order*. New York: UNESCO.
- Masmoudi, M. 1979. "The New World Information Order." *Journal of Communication* 29 (2): 172–179. doi:10.1111/j.1460-2466.1979.tb02960.x.
- Matsa, KE, and E Shearer. 2018. *News use across social media platforms 2018: Most Americans continue to get news on social media, even though many have concerns about its accuracy*. Pew Research Center.
- Mitchell, A., and D. Page. 2015. *Millennials & Political News: Social Media – the Local TV for the Next Generation?* [www.pewresearch.org](http://www.pewresearch.org).

- Morozov, E. 2011. *The net Delusion: The Dark Side of Internet Freedom*. New York: Public Affairs.
- Peterson, S. 1981. "International News Selection by the Elite Press: A Case Study." *Public Opinion Quarterly* 45 (2): 143–163. doi:10.1086/268647.
- Postman, N. 1985. *Amusing Ourselves to Death*. London: Methuen.
- Seo, H., and S. J. Thorson. 2012. "Networks of Networks: Changing Patterns in Country Bandwidth and Centrality in Global Information Infrastructure, 2002–2010." *Journal of Communication* 62 (2): 345–358. doi:10.1111/j.1460-2466.2012.01631.x.
- Servaes, J. 2012. "Soft Power and Public Diplomacy: The New Frontier for Public Relations and International Communication Between the US and China." *Public Relations Review* 38: 643–651.
- Skurnik, W. A. E. 1981. "Foreign News Coverage in six African Newspapers: The Potency of National Interests." *International Communication Gazette* 28: 117–130.
- Taubman, G. 1998. "A Not-So World Wide Web: The Internet, China, and the Challenges to Nondemocratic Rule." *Political Communication* 15 (2): 255–272. doi:10.1080/10584609809342369.
- Wallerstein, I. 1974. *The Modern World-System: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. New York: Academic Press.
- Wanta, W., G. Golan, and C. Lee. 2004. "Agenda Setting and International News: Media Influence on Public Perceptions of Foreign Nations." *Journalism & Mass Communication Quarterly* 81 (2): 364–377. doi:10.1177/107769900408100209.
- Westerståhl, J., and F. Johansson. 1994. "Foreign News: News Values and Ideologies." *European Journal of Communication* 9 (1): 71–89.
- White, C. S. 1997. "Citizen Participation and the Internet: Prospects for Civic Deliberation in the Information Age." *The Social Studies* 88 (1): 23–28. doi:10.1080/00377999709603741.
- Wu, H. D. 2000. "Systemic Determinants of International News Coverage: a Comparison of 38 Countries." *Journal of Communication* 50 (2): 110–130. doi:10.1111/j.1460-2466.2000.tb02844.x.
- Wu, H. D. 2007. "A Brave new World for International News? Exploring the Determinants of the Coverage of Foreign News on US Websites." *International Communication Gazette* 69 (6): 539–551. doi:10.1177/1748048507082841.