The Fall of the Empire: The Americanization of English

Bruno Gonçalves [1,†], Lucía Loureiro-Porto[2], José J. Ramasco[3], David Sánchez[3]

1 Center for Data Science, New York University, 60 5th Avenue, New York, NY 10012, USA

2 Departament de Filologia Espanyola, Moderna i Clàssica, Universitat de les Illes Balears (UIB), 07122 Palma de Mallorca, Spain

3 Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), 07122 Palma de Mallorca, Spain

† E-mail: bgoncalves@gmail.com

arXiv:1707.00781v1 [cs.CL] 3 Jul 2017

Abstract

As global political preeminence gradually shifted from the United Kingdom to the United States, so did the capacity to culturally influence the rest of the world. In this work, we analyze how the world-wide varieties of written English are evolving. We study both the spatial and temporal variations of vocabulary and spelling of English using a large corpus of geolocated tweets and the Google Books datasets corresponding to books published in the US and the UK. The advantage of our approach is that we can address both standard written language (Google Books) and the more colloquial forms of microblogging messages (Twitter). We find that American English is the dominant form of English outside the UK and that its influence is felt even within the UK borders. Finally, we analyze how this trend has evolved over time and the impact that some cultural events have had in shaping it.

1 Introduction

With roots dating as far back as Cabot's explorations in the 15th century and the 1584 establishment of the ill-fated Roanoke colony in the New World, the British empire was one of the largest empires in Human History. At its zenith, it extended from North America to Asia, Africa and Australia deserving the moniker "the empire where the sun never sets". However, as history has shown countless times, empires rise and fall due to a complex set of internal and external forces. In the case of the British empire, its preeminence faded as the United States of America –one of its first colonies– took over the dominant role in the global arena.

As an empire spreads so does the language of its ruling class. Thanks to both its global extension, late demise, and the rise of the US as a global actor, the English language enjoys an undisputed role as the global *lingua franca* serving as the default language of science, commerce and diplomacy [1, 2] (see Fig. 1). Given such an extended presence, it is only natural that English would absorb words, expressions and other features of local indigenous languages resulting in dozens of dialects and topolects (language forms typical of a specific area) such as "Singlish" (Singapore), "Hinglish" (India), Kenyan English [3], and, most importantly, American English [4] a variety that includes within itself several other dialects [5, 6].

The transfer of political, economical and cultural power from Great Britain to the United States has progressed gradually over the course of more than half a century, with World War II being the final stepping stone in the establishment of American supremacy. The cultural rise of the United States also implied the exportation of their specific form of English resulting in a change of how English is written and spoken around the world. In fact, the "Americanization" of

(global) English is one of the main processes of language change in contemporary English [7]. As an example, if we focus on spelling, some the original differences between British and American English orthography (most of which are the result of Webster's reform [8]) are somehow blurred and, for instance, the tendency for verbs and nouns to end in ⟨-ize⟩ and ⟨-ization⟩ in America is now common on both sides of the Atlantic [9]. Likewise, a tendency for Postcolonial varieties of English in South-East Asia to prefer American spelling over the British one has been observed, at least, for Nigerian English [10], Singapore and Trinidad and Tobago [11], regarding spelling and lexis, for Indian English [12] and Bahamas [13], regarding syntax, and for Hong Kong [14], regarding phonology. In addition, a growing tendency for Americanization has been observed for Philippine English, which, despite being rooted in American English, has experienced a rise in the frequency of American forms [15]. Although this Americanization is found in different registers, web genres have been highlighted as a text-type where American forms are preferred [16]. Electronic communication has indeed been considered to play a role in linguistic uniformity [17]. It is in this sense that this paper will make a contribution to the study of the Americanization of English, since a corpus of 213,086,831 geolocated tweets will be used to study the spread of American English spelling and vocabulary throughout the globe, including regions where English is used as a first, second and foreign language.

The study of diatopic variation using Twitter datasets is a relatively new subject [18]. The use of geotagged microblogging data [19] allows to quantitatively examine linguistic patterns on a worldwide scale, in automatic fashion and within conversational situations. The global extension and the real time availability of the data constitute major methodological advantages over more traditional approaches like surveys and interviews [20]. Importantly, the resulting corpora are publicly available [21], although due to their nature most of the literature has been concerned with lexical variation (for an exception that addresses semantic and syntactic variation, see Ref. [22]). Thus, different variables can be mapped after carefully removing lexical ambiguities [23]. A Bayesian approach shows good agreement between baseline queries and survey responses [24]. Machine learning techniques applied to Twitter corpora reveal the existence of superdialects [25,26], which can be further analyzed with dialectometric techniques [27]. Linguistic evolution in social media appears to be strongly connected to demographics [28]. Age and gender issues can be additionally introduced in the analysis [29]. Moreover, an investigation of lexical alternations unveils hierarchical dialect regions in the United States [30]. Twitter can be also employed in the study of specific varieties departing from the standard form [31]. However, online social media are more suitable for a synchronic approximation to language variation. If one aims at understanding the diachronic evolution of language, we need a corpus well established over time. This is available with the Google Books database [32], which has already been used for the analysis of relative frequencies that characterize word fluxes [33,34] or the applicability of Zip's and Heaps's law with different scaling regimes [35]. Here, we will complement our Twitter study of the Americanization of English with an analysis of the dynamical process that is taking place since 1800.

In this paper we analyze how English is used around the world, in informal contexts, using a large scale Twitter dataset. Due to the written nature of our corpus we consider in detail both how vocabulary and spelling of common words varies from place to place in order to understand how American cultural influence is spreading around the world. We complement this synchronic analysis with a diachronic view of how the prevalence of British and American vocabulary and

spelling have evolved over time in British and American publications using the Google Books dataset.

2 Methods

Datasets

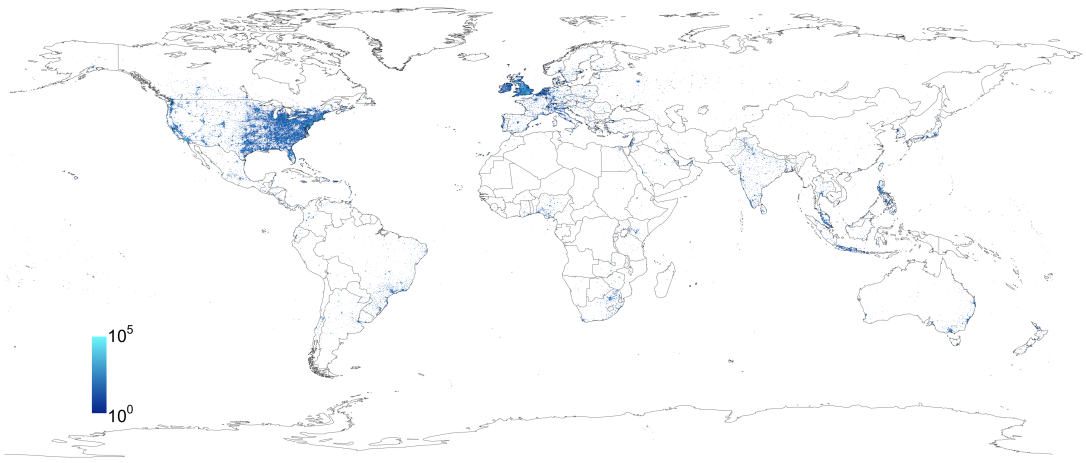

Figure 1. English tweets A heatmap showing the location of geolocated English tweets in our dataset that match our keywords.

The goal of this manuscript is to analyze how English is used across both time and space. We study the geographical variation of English by using the Twitter Decahose from which we collect [36] all tweets written in English between May 10, 2010 and Feb 28, 2016 that contain geolocation information. The language is detected using Chromium Compact Language Detection library as in Ref. [36]. The tweets are then mapped to a grid of cells of $0.25° \times 0.25°$ spanning the globe and resulting in $30,898,072$ tweets matching our list of words. A heatmap illustrating the geographical distribution of matching tweets is shown in Fig. 1. Out of the general dataset we further select those tweets with spelling and vocabulary features that allow us to discern the variety of English used (see below for a detailed description).

The temporal evolution of English is analyzed using the Google Books dataset [32] of books published by both British and American publishers. The dataset contains the number of times individual words were used in books scanned by Google and dating back to the 15th century. However, due to the poor statistics in earlier periods, we restrict our analysis to the period between 1800 and 2010. Both data sources are different in nature: Twitter contains more colloquial expressions, while the language recorded in the books is more formal. As a result, these two sources, in combination, can provide a useful perspective on the spatio-temporal patterns developed or developing in English.

Metrics

The polarization, V_w^c, for a concept w in cell c during the data collection period is defined as the ratio:

$$V_w^c = \frac{A_w^c - B_w^c}{A_w^c + B_w^c}, \tag{1}$$

where A_w^c (B_w^c) is the number of American (British) forms of the concept w observed in cell c. The polarization is then constrained to be in the $[-1, 1]$ domain, with -1 corresponding to purely British and 1 being purely American forms.

The polarization of each cell, V^c, is then determined by taking the average polarization over all words observed in cell c:

$$V^c = \frac{\sum_w V_w^c}{W^c}, \tag{2}$$

where W^c is the number of different words observed in cell c. Similarly, the polarization score of a country is defined as the average polarization taken over all the cells within that country. By considering the average polarization we are able to compare countries of varying sizes.

In the case of Twitter, the polarization signal is measured over the complete time period of the database since, as it is not long enough to allow for large variations in the language use patterns. On the other hand, when the time evolution of written language is considered with Google Books the space is not relevant, beyond the country of origin of the published book, and we add an index referring to the year y considered. The polarization V^y is then defined as:

$$V^y = \frac{\sum_w V_w^y}{W^y}, \tag{3}$$

where V_w^y is the concept polarization for year y and W^y refers to all the books published in the country considered, the US or the UK, during year y.

3 Results

In our analysis, we consider two factors of differentiation between American and British English: Spelling and Vocabulary with different word lists used for each case. The complete list of words and expression used in each case can be found in the online supplementary material. It is the result of compiling information in reference books [9] and online sources such as the Oxford Dictionaries[1]. The words in the list were subsequently checked in two widely-used representative corpora of British and American English [37, 38]. Only pairs of words in which one of the members exhibits a significantly higher frequency in either of the two varieties were considered for inclusion in the list. Inflectional forms (e.g., *solicitor*, *solicitors*, *solicitor's*, *solicitors'*) as well as derived (e.g., *amphitheater*) and compound forms were also included in the search (e.g., *sportscenter*).

Let us start by considering how the Vocabulary used for common terms such as *lorry/truck* or *motorway/freeway* changes around the world by defining the ratio of each cell as described above. The results are plotted in Fig 2. Unsurprisingly, we find that the British Islands are tendentially

[1] https://en.oxforddictionaries.com/usage/british-and-american-terms

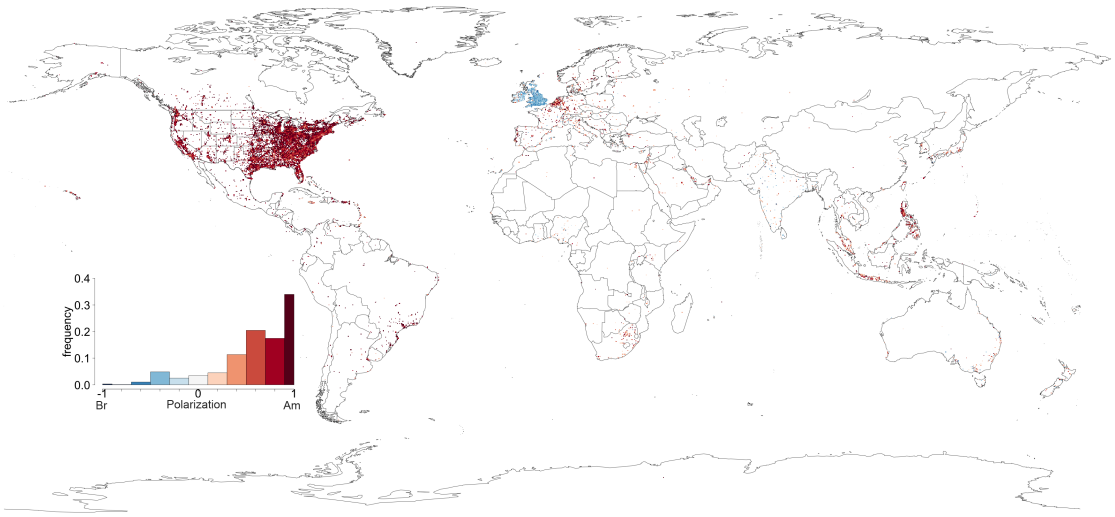

Figure 2. Vocabulary The polarization ratio of each cell around the world according to the vocabulary used within each cell. The inset barplot is an histogram of the number of cells as a function of the ratio.

blue while the United States is predominantly red as befits the representatives of each trend. Interestingly, Western Europe where English teaching has traditionally followed British norms the American influence is undeniable. Most areas are depicted in various shades of red while some of the largest international metropolises such as Madrid, Paris, Amsterdam, Berlin, Milan or Rome are visible in light shades, in no doubt due to their role as touristic and transportation hubs, see Fig. 3(left). A more marked British influence is easily seen in former colonies (see also Fig. 5) such as South Africa, Australia, New Zealand ("the only large areas in the Southern hemisphere where English is spoken as a native language" [9], and which have reached a very advanced phase of development, according to Schneider's 2007 Dynamic Model [39]) or India (where English is spoken as a non-native language, but which has followed an exonormative model, i.e., strongly based in the British rules [40]) displaying large areas of blue side by side with tell tale patches of white in the most international areas such as Pretoria, Melbourne, Sidney, Auckland, New Delhi or Mumbai. Furthermore, countries such as the Philippines (one of the few Postcolonial varieties of English with an American superstratum [39]), as well as Taiwan, South Korea and Japan (where English is spoken as a second language) attest their strong American influence with full displays of red.

Regarding Spelling, the case for American influence becomes even stronger as displayed in Fig. 4. The British Isles attain significantly lighter shades of blue as do the former British colonies with South Africa, Australia and New Zealand becoming predominately red. This dichotomy between spelling and vocabulary, illustrated in Fig. 3 for Europe, is perhaps a testament to the conflicting forces of traditional formal education and media influence. Individuals who studied in school systems that subscribe to the British form of English are more prone to continue writing words in the way they originally learned them. However, through the influence of American

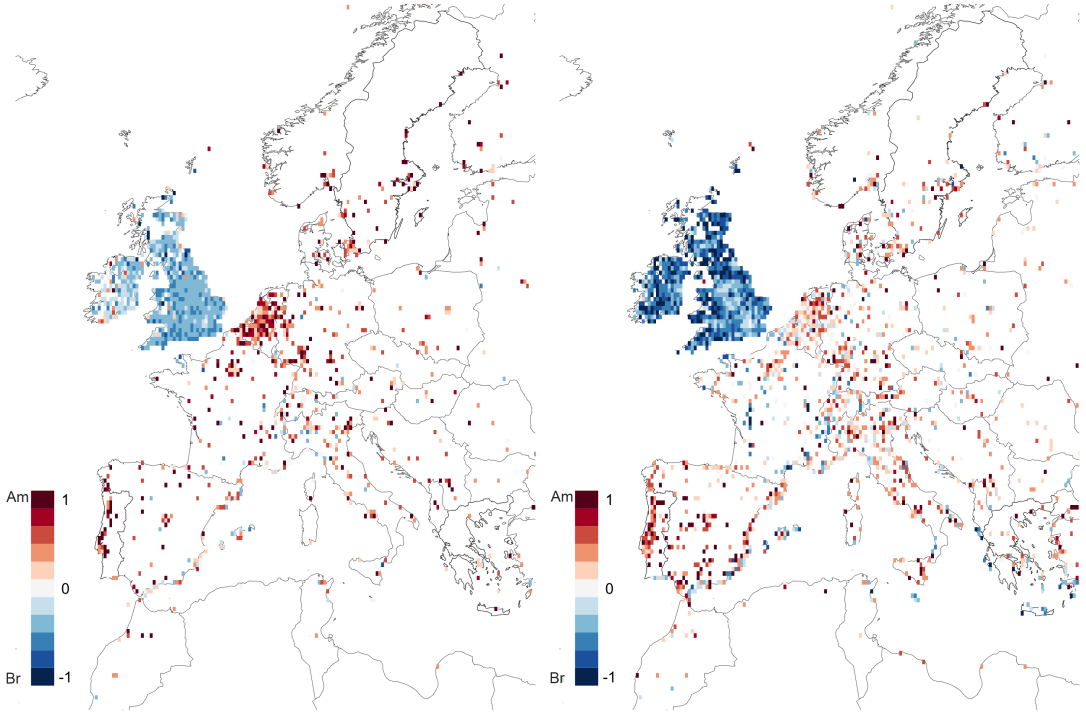

Figure 3. Europe Side by side comparison of the Vocabulary (left) and Spelling (right) results for countries in continental Europe. The tension between British spelling and American vocabulary is clearly visible by the shift towards lighter shades of blue and darker shades of red between the left and the right plots.

dominated television and film industries they have acquired new (American) vocabulary. This can be clearly seen in Fig. 5 where we plot the average polarization for both Vocabulary and Spelling for 30 countries around the world, including countries belonging to Kachru's [41] inner circle, i.e., where English is spoken as a native language (e.g., UK, Ireland), outer circle, i.e., where English is spoken as a second language (e.g., India, South Africa) and the expanding circle, i.e., where English is spoken as a foreign language (e.g., Portugal, Finland, Russia). Interestingly enough, in all expanding circle territories the American orthography and vocabulary dominate, and the same happens, obviously, in the United States and in the Philippines, a former American colony. The bottom part of the figure includes inner and outer circle varieties, where American vocabulary is also chosen over British forms, with the notable exception of India, UK and Ireland, whose green bars are always towards the left hand (British) side of the ratio spectrum. India's alignment with the UK is clearly the result of an exonormative model and postcolonial prescriptivism in this former colony of the United Kingdom [40, 42]. Surprisingly, we find that in some ex-colonies which still hold strong ties with the British empire, such as South Africa, Australia and New Zealand, the drift towards American vocabulary is unmistakable.

We now consider a temporal view of how English as a language is evolving. Using the word counts provided by the Google Books digitalization efforts, we measure the Vocabulary and

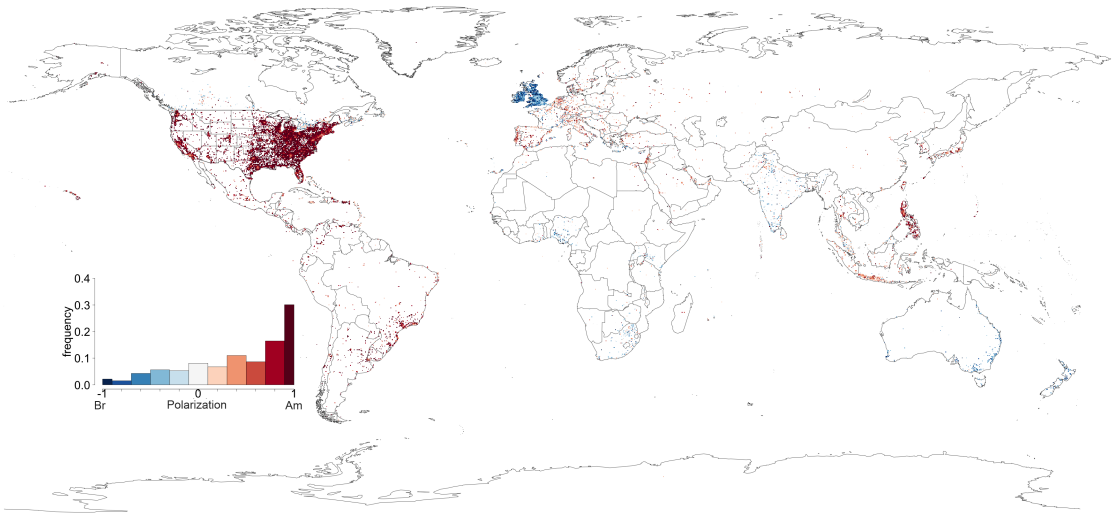

Figure 4. Spelling The polarization of each cell around the world according to the spelling used within each cell. The inset barplot is an histogram of the number of cells as a function of the ratio observed.

Spelling average ratio per year for books published by American and British publishing houses. An analysis of the resulting timelines as shown in Fig. 6 provides several interesting insights. First, we can see that the divergence in spelling between the American and British forms has significantly increased in the last 200 years. Indeed, from this time series we can pinpoint the beginning of the trend to around 1828 when Noah Webster published *An American Dictionary of the English Language* [43] with the explicit goal of systematizing the way in which English was written in America. As [44] puts it: "He is certainly responsible for establishing (though not inventing) the common differences between traditional British and American spellings" the final *-or* versus *-our* in *color*, *labor*, *savor*, and the like; *-er* versus French *-re* in *theater*, *center*, *meter*; and the simplification of final *-ck* as in *physic*, *music*, *logic*. This is now considered to have been the first American-English dictionary and it started the Merriam-Webster series of Dictionaries that is still dominant today. The US vocabulary curve follows a similar but less pronounced trend as it takes longer for new words to be created than for people to agree on a common spelling form.

Another interesting feature of these timelines is the pronounced "Britishization" of American English in the years following World War II as seen by the declining slope that extends until after 1960. This can likely be explained by the large influx of European migrants that moved to America in search of a better life away from a destroyed or warring Europe. In the immediate aftermath of WWII congress passed the War Brides Act in 1946 and the Displaced Persons Act in 1948 to facilitate the immigration to the US by the people affected by the war. It is estimated that between 1941 and 1950 over 1 Million people [45], mostly of European descent, immigrated to the United States that at the time had a population of 150 million. In the following decade, this number doubled to over 2 Million [46].

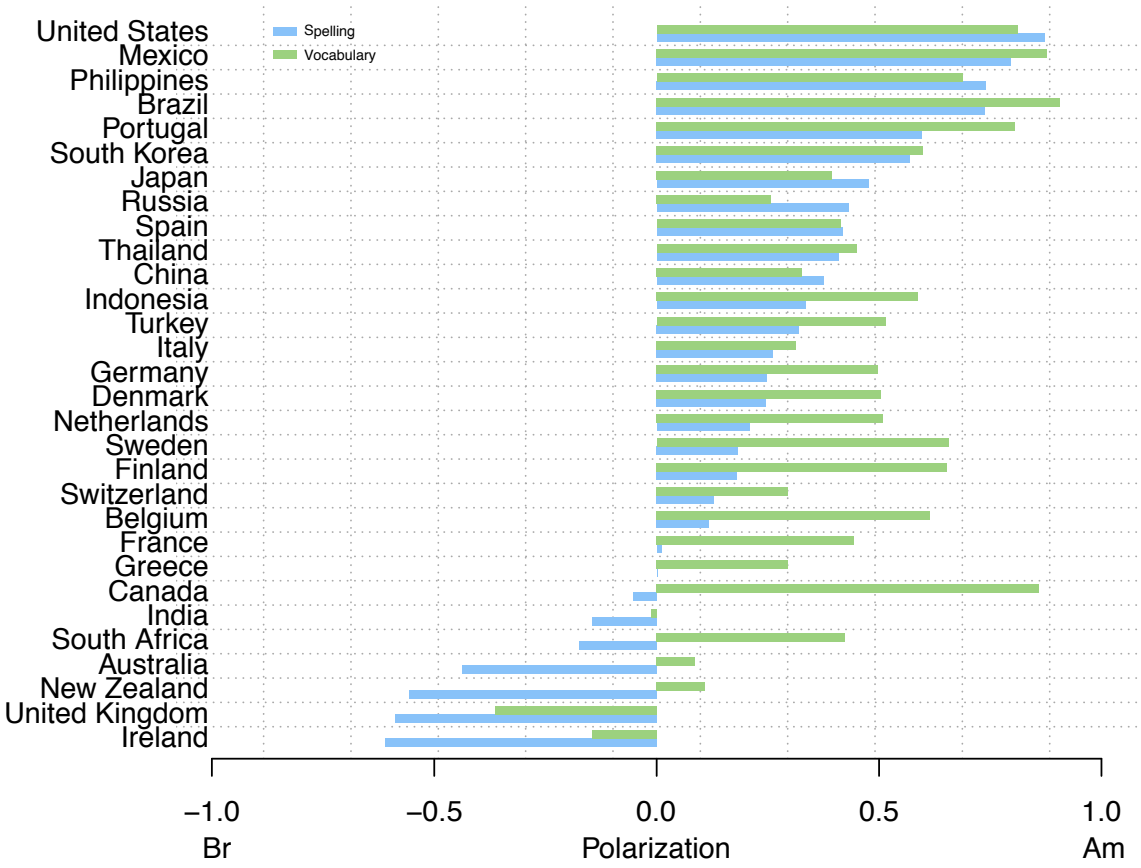

Figure 5. Vocabulary and Spelling polarization ratio by country

Interestingly, while the ratio timelines within the United Kingdom had been towards becoming ever more British, we find a significant change of trend in the last 20 years of our dataset, corresponding to the period after the fall of the Berlin wall and the end of the Cold War that left America as the world's only superpower. It is the status quo resulting from the aftermath of this trend that we are able to observe in the Twitter analysis above.

4 Conclusions

The way in which languages evolve in time and change from place to place has long been the focus of much interest in the linguistic community. With the advent of new and extensive corpora derived from large scale online datasets we are now able to take on a more quantitative approach to tackling this fundamental question. In this work we analyze two datasets that, when taken together, are able to provide a bird's eye view of the way English usage has been changing over time and in different countries.

The picture we are able to paint is particularly stark. The past two centuries have clearly resulted in a clear shift in vocabulary and spelling conventions from British to American. This

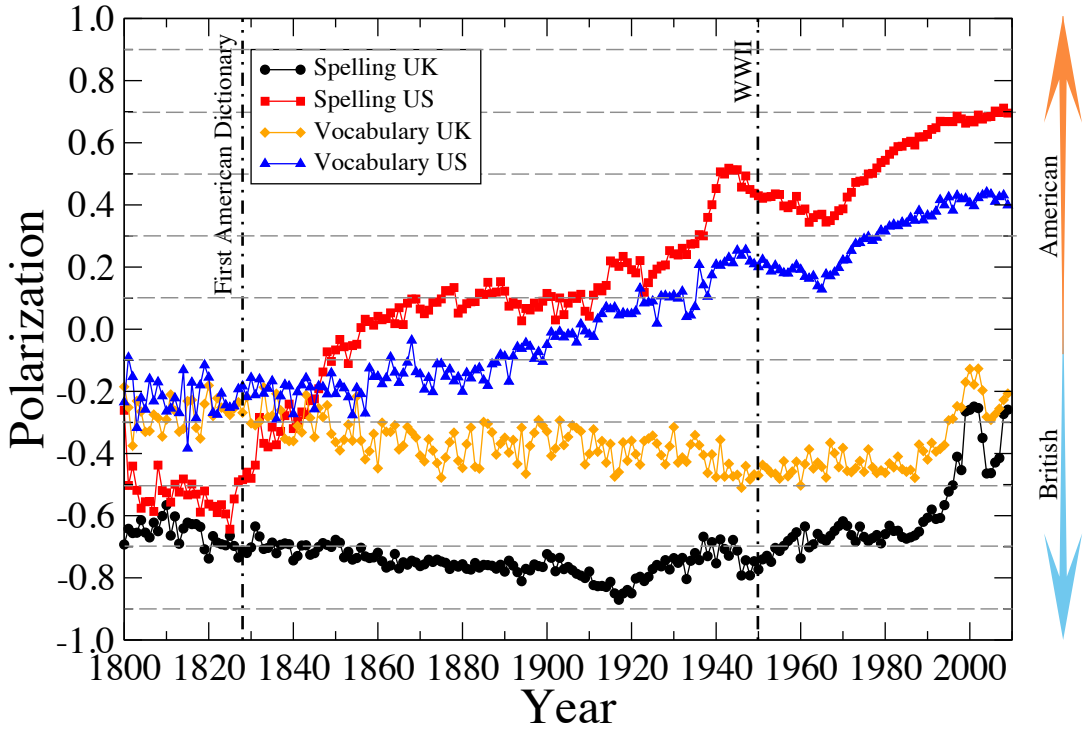

Figure 6. The Americanization of English over time Averaged polarization ratio of Vocabulary and Spelling for books published by US and UK publishing companies in the $1800 - 2010$ period.

trend is especially visible in the decades following WWII and the fall of Berlin Wall. Indeed, when we consider the current status quo as seen through the lens of Twitter, it becomes clear that only in the countries where British influence has been strongest, such as ex-colonies with a strong exonormative influence (in Schneider's terms [39]), are British conventions still dominant to some degree.

It should be noted that both datasets we utilize in our analysis are intrinsically biased. Books are typically written by cultural elites. Also, despite their increasing democratization, GPS enabled mobile devices are, in many countries, only available to middle and higher economic strata. As a result, there are certainly factors of linguistic evolution we are missing but the fact that both datasets agree on the general picture means that we are able to capture, at the very least, the underlying trends.

5 Acknowledgments

BG thanks the Moore and Sloan Foundations for support as part of the Moore-Sloan Data Science Environment at NYU. LL-P thanks the Spanish Ministry of Economy and Competitiveness for funding under the grants FFI2014-53930-P and FFI2014-51873-REDT.

References

1. Crystal D (2003) English as a Global Language. Cambridge University Press.
2. Jenkins J, Leung C (2013) English as a lingua franca. The Companion to Language Assessment IV 13: 1605.
3. Mesthrie R, Bhatt RM (2008) The Study of New Linguistic Varieties. Cambridge University Press.
4. Grieve J (2016) Regional Variation in Written American English. Cambridge University Press.
5. Pederson L (2001) The Cambridge History of the English Language, Cambridge University Press, chapter Dialects. p. 253.
6. Wolfram W, Shelling N (2016) American English: Dialects and Variation. Malden/Oxford: Wiley-Blackwell.
7. Leech G, Hundt M, Mair C, Smith N (2009) Change in Contemporary English: A Grammatical Study. Cambridge University Press.
8. Algeo J (2001) The Cambridge History of the English Language, Cambridge University Press, chapter External History.
9. Gramley S, Pätzold KM (2003) Survey of Modern English. Routledge.
10. Awonusi VO (1994) The Americanization of Nigerian English. World Englishes 13: 75.
11. Hänsel EC, Deuber D (2013) Globalization, postcolonial Englishes, and the English language press in Kenya, Singapore, and Trinidad and Tobago. World Englishes 32: 338.
12. Davydova J (2015) Indian English quotatives in a real-time perspective, Benjamins, chapter Indian English quotatives in a real-time perspective. p. 173.
13. Hackert S (2015) Pseudotitles in Bahamian English: A case of Americanization? Journal of English Linguistics 43: 143.
14. Hansen Edwards JG (2016) Accent preferences and the use of American English features in Hong Kong: a preliminary study. Asian Englishes 18: 197.
15. Fuchs R (Forthcoming) The Americanization of Philippine English –Recent diachronic change in spelling and lexis. Philippine ESL Journal .

16. Mukherjee J (2015) Response to Davies and Fuchs. English Worldwide 36: 34.

17. Venezky RL (2001) The Cambridge History of the English Language, Cambridge University Press, chapter Spelling. p. 340.

18. Nguyen D, Dogrüoz AS, Rosé CP, de Jong F (2015) Computational sociolinguistics: A survey. arxiv:150807544 .

19. Melo F, Martins B (2016) Automated geocoding of textual documents: A survey of current approaches. Transactions in GIS .

20. Chambers JK, Trudgill P (1998) Dialectology. Cambridge University Press.

21. Malmasi S, Zampieri M, Ljubesšić N, Nakov P, Ali A, et al. (2016) Discriminating between similar languages and Arabic dialect identification: A report on the third DSL shared task. Proceedings of the Third Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial3) : 1–14.

22. Kulkarni V, Perozzi B, Skiena S (2016) Freshman or fresher? quantifying the geographic variation of language in online social media. Proceedings of the Tenth International AAAI Conference on Web and Social Media .

23. Russ B (2012) Examining large-scale regional variation through online geotagged corpora. ADS Annual Meeting .

24. Doyle G (2014) Mapping dialectal variation by querying social media. EACL .

25. Gonçalves B, Sánchez D (2014) Crowdsourcing dialect characteriation through Twitter. PLoS One 9: E112074.

26. Gonçalves B, Sánchez D (2016) Learning about spanish dialects through Twitter. RILI 28: 65–75.

27. Donoso G, Sánchez D (2017) Dialectometric analysis of language variation in Twitter. Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial4) : 16–25.

28. Eisenstein J, O'Connor B, Smith NA, Xing EP (2014) Diffusion of lexical change in social media. PLoS ONE 9: E113114.

29. Pavalanathan U, Eisenstein J (2015) Confounds and consequences in geotagged Twitter data. Proceedings of the Conference on Empirical Methods in Natural Language Processing .

30. Huang Y, Guo D, Kasakoff A, Grieve J (2016) Understanding U.S. regional linguistic variation with twitter data analysis. Computers, Environment and Urban Systems 54.

31. Blodgett SL, Green L, O'Connor B (2016) Demographic dialectal variation in social media: A case study of African-American English. EMNLP .

32. Michel JB, Shen YK, Aiden AP, Veres A, Gray MK, et al. (2011) Quantitative Analysis of Culture Using Millions of Digitized Books. Science 331: 176–182.

33. Pedersen AM, Tenenbaum JN, Havlin S, Stanley HE, Perc M (2012) Languages cool as they expand: Allometric scaling and the decreasing need for new words. Sci Rep 2: 943.

34. Pechenick EA, Danforth CM, Dodds PS (2015) Is language evolution grinding to a halt?: Exploring the life and death of words in English fiction. arXiv:150303512 .

35. Gerlach M, Altmann EG (2016) Stochastic model for the vocabulary growth in natural languages. Phys Rev X 3: 021006.

36. Mocanu D, Baronchelli A, Perra N, Gonçalves B, Vespignani A (2013) The Twitter of Babel: Mapping world languages through microblogging platforms. PLoS One 8: E61981.

37. Davies M (2004) BYU-BNC (based on the British National Corpus from Oxford Universtiy Press). Available online at http://corpus.byu.edu/bnc.

38. Davies M (2008) The corpus of contemporary American English: 520 million words 1990-present. Available online at `http://corpus.buy.edu/coca`.

39. Schneider EW (2007) Postcolonial English. Varieties around the World. Cambridge University Press.

40. Schneider EW (2011) English around the World: An Introduction. Cambridge University Press.

41. Kachru BB (1985) English in the world: Teaching and learning the language and literatures, Cambridge University Press, chapter Standards, codification and sociolinguistic realism: the English language in the outer circle. pp. 11–30.

42. Collins P (2013) English modality: Core, Periphery and Evidentiality, Mouton de Gruyter, chapter Grammatical colloquialism and the English quasi-modals: a comparative study.

43. Webster N (1828) An American dictionary of the English language. S. Converse.

44. Cassidy FG, Hall JH (2001) The Cambridge History of the English Language IV: English in North America, Cambridge University Press, chapter Americanisms.

45. Census Bureau (1950) Statistical Abstract of the U.S. 1950.

46. Census Bureau (1960) Statistical Abstract of the U.S. 1960.

A Word Lists

Vocabulary

British	American
railway	railroad
MA dissertation, MA dissertations	MA thesis, MA theses
doctoral thesis, doctoral theses	doctoral dissertation, doctoral dissertations
draughts	checkers
abseil, abseils, abseiled, abseiling	rappel, rappels, rappelled, rappeled, rappelling, rappeling
antenatal	prenatal
anticlockwise	counterclockwise
aubergine, aubergines, aubergine's, aubergines'	eggplant, eggplants, eggplant's, eggplants'
barrister, barristers, barrister's, barristers', solicitor, solicitors, solicitor's, solicitors'	attorney, attorneys, attorney's, attorneys'
biscuit, biscuits, biscuit's, biscuits'	cookie, cookies, cookie's, cookies'
car park, car parks, car park's, car parks'	parking lot, parking lots, parking lot's, parking lots'
caster sugar, icing sugar	confectioner's sugar, powdered sugar
corn flour	corn starch
cupboard, cupboards, cupboard's, cupboards'	closet, closets, closet's, closets'
demister	defroster
drawing pin, drawing pins, drawing pin's, drawing pins'	thumbtack, thumbtacks, thumbtack's, thumbtacks'
Father Christmas	Santa Claus
handbrake, hand brake	emergency brake
hire purchase	installment plan
inside leg	inseam
mobile phone, mobile phones, mobile phone's, mobile phones'	cell phone, cell phones, cell phone's, cell phones'
motorway, motoways, motorway's, motorways'	expressway, expressways, expressway's, expressways', freeway, freeways, freeway's, freeways'
nappy, nappies, nappy's, nappies'	diaper, diapers, diaper's, diapers'
notice board, notice boards, notice board's, notice boards'	bulletin board, bulletin boards, bulletin board's, bulletin boards'
number plate, number plates, number plate's, number plates'	license plate, license plates, license plate's, license plates'
plasterboard, plasterboards, plasterboard's, plasterboards'	Wallboard, wallboards, wallboard's, wallboards'
polystyrene	styrofoam
porridge	oatmeal

perspex	plexiglass
Pushchair, pushchairs, pushchair's, pushchairs'	Stroller, strollers, stroller's, strollers'
rubbish	garbage
skirting board	baseboard
spring onion, spring onions, spring onion's, spring onions'	green onion, green onions, green onion's, green onions'
sticky tape	scotch tape
sweets	candy
torch, torches	flashlight, flashlights
tracksuit, tracksuits, tracksuit's, tracksuits'	sweatsuit, sweatsuits, sweatsuit's, sweatsuits'
trousers	pants
valuer, valuers, valuer's, valuers'	appraiser, appraisers, appraiser's, appraisers'
wellington boots, wellingtons	rubbers, rubber boots, rain boots
windscreen, windscreens, windscreen's, windscreens'	windshield, windshields, windshield's, windshields'
lorry, lorries, lorry's, lorries	truck, trucks, truck's, trucks'
chemist's	drug store, drug stores
elastic band, elastic bands, elastic band's, elastic bands'	rubber band, rubber bands, rubber band's, rubber bands'
estate agent, estate agents, estate agent's, estate agents'	realtor, realtors, realtor's, realtors'
cot, cots	crib, cribs
off-licence	liquor store
crayfish	crawfish
capsicum	bell pepper

Table 1. Vocabulary word list

Spelling

British	American
skilful	skillful
wilful	willful
fulfil, fulfils	fulfill, fulfills
instil, instils	instill, instills
appal, appals	appall, appalls
flavour, flavours, flavour's, flavours'	flavor, flavors, flavor's, flavors'
mould, moulds, mould's, moulds'	mold, molds, mold's, molds'
moult, moults, moulted, moulting	molt, molts, molted, molting
smoulder, smoulders, smouldered, smouldering	smolder, smolders, smoldered, smoldering
moustache, moustaches, moustache's, moustaches'	mustache, mustaches, mustache's, mustaches'

centre, centres, centre's, centres'	center, centers, center's, centers'
metre, metres, metre's, metres'	meter, meters, meter's, meters'
theatre, theatres, theatre's, theatres'	theater, theaters, theater's, theaters'
analyse, analyses, analysed, analysing	analyze, analyzes, analyzed, analyzing
paralyse, paralyses, paralysed, paralysing	paralyze, paralyzes, paralyzed, paralyzing
defence	defense
offence	offense
pretence	pretense
revelling, revelled	reveling, reveled
travelled, travelling	traveled, traveling
traveller, travellers, traveller's, travellers'	traveler, travelers, traveler's, travelers'
marvellous	marvelous
plough, ploughs, ploughed, ploughing	plow, plows, plowed, plowing
aluminium, aluminium's	aluminum, aluminum's
jewellery, jewellery's	jewelry, jewelry's
pyjamas	pajamas
whisky, whisky's	whiskey, whiskey's
neighbour*	neigbor*
honour	*honor*
colour	*color*
behaviour	*behavior*
labour	*labor*
humour*	humor*
favour	*favor*
harbour*	harbor*
tumour*	tumor*
vigour*	vigor*
rumour*	rumor*
rigour*	rigor*
demeanour	*demeanor*
clamour*	clamor*
odour*	odor*
armour*	armor*
endeavour*	endeavor*
parlour*	parlor*
vapour*	vapor*
saviour*	savior*
splendour*	splendor*
fervour*	fervor*
savour*	savor*
valour*	valor*

candour*	candor*
ardour*	ardor*
rancour*	rancor*
succour*	succor*
arbour*	arbor*
catalogue*	catalog*
analog*	analog*
acknowledgement*	acknowledgment*
goitre	goiter
foetus	fetus
paediatrician	pediatrician
oesophagus	esophagus
manoeuvr*	maneuver*
oestrogen	estrogen
anaemia	anemia

Table 2. Spelling word list

View publication stats