Article

Japanese Company Presence in News on the Web: Graphical Study of English Language Markets

Theron E. Fairchild

Abstract

The study addresses the presence of Japanese companies in online English news markets globally. The primary instrument is graphical representation, based on searches conducted in the News on the Web (NOW) corpus. Unique keyword identifiers and quantifiable assessment point to word-frequency presence across 15 English language news markets. The study is a data treatment of recent history, and is just as much about assessing Japanese business presence as it is about the methods deployed to conduct the research. The study follows a type of historical science approach.

keywords: digital humanities, historical graphs, Internet news, Japanese business, Japanese companies, Japanese economy, web news

Introduction

The study is primarily a graphical demonstration of the presence of Japanese companies in English-language Internet news sources. The purpose was to visually address, rather than critically analyze, how keyword searches reveal useful information regarding business presence in web-based news. In this case, the demonstration was built through a sample of keywords, which pointed to some of the most-active Japanese companies in international markets.

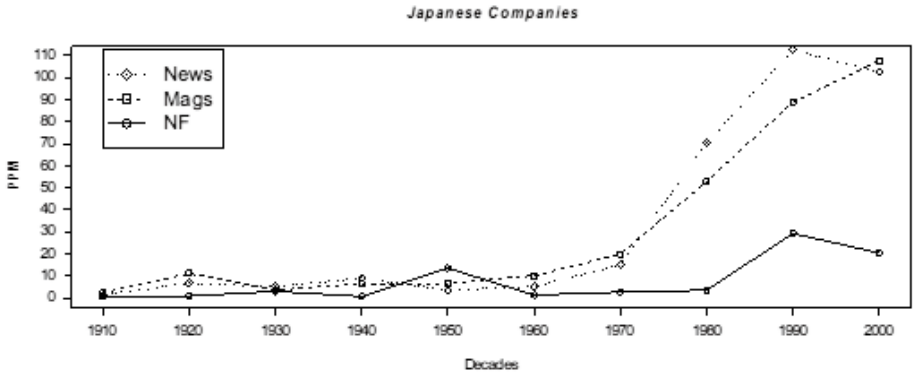
Past Research. In a previous paper [1], I investigated aspects of Japanese history through a quantifiable word-frequency assessment, which included general trends in the international development of Japanese businesses (see Figure 1). In the study, I cited the 1949 statistical work on postwar Japan, by economist J. B. Cohen [2], as an early, viable attempt to more fully understand a historical phenomenon. Though Cohen's efforts were uncommon for his era, advances in computers and data processing have recently allowed historical investigators to take advantage of methods that were previously reserved for scientists. In what could be tentatively called historical science, by merging the boundaries of science and the humanities, historians can take fuller advantage of quantifiable and graphical representations. This is important when understanding that historical events are

not static occurrences, but evolve from previous historical developments. Assessing these patterns of the past, and their potentially complex relationships with the present, are part of what it means to interpret history effectively.

With the evolution of computing, scholars in the humanities and social sciences are now able to exploit more data as part of their research. Such data can take the form of word-frequency characteristics in language. Analyzing language in this way has become more common in fields such as social psychology [e.g., 3], literary studies [e.g., 4], and history as a science [e.g., 5]. Specifically, distant reading, as it is tentatively called, refers to analyzing texts for how their respective subjects have evolved in the written record, and how the record itself has changed. In my previous research [1], I briefly examined word-frequency characteristics regarding Japanese history and economy, as they appear in published American English (Figure 1).

Figure 1

Japanese Companies: Frequency in American English Publications (20th Century)



Note. News = newspapers. Mags = magazines. NF = nonfiction books. Number of search terms for Japanese companies = 30. (Graphic from Fairchild, 2017) [1]

Present Research. In the current study, the focus again was on Japan, a nation with a strong presence in economic history and world affairs over the past century. Though earlier work looked at Japan and Japanese business in American English over that century, the current concentration was on Internet-based international news over the eight-year span of 2010 through 2017.

To conduct the examination, I explored News on the Web (NOW), the online corpus available through the Brigham Young University website (corpus.byu.edu) [6]. The corpus includes global web news sources, in the form of major online newspapers and magazines. The database builds from more than 5.5 billion words published between January 2010 and the present. The corpus has the unique feature of dividing its source samples by individual country of origin. These comprise 20 nations where English is either the primary language (i.e., Canada), or is a dominant official or second language (i.e., India & Malaysia) (see Table 2). This feature allows for isolating specific English news markets for comparison.

Through keywords that were operationalized as identifiers for Japanese companies, the study attempted to show how such companies appear in individual NOW markets. For the sample list of Japanese companies, the study employed a set of 66 keyword identifiers (see Table 1), which was designated *JapanCo*. The set includes such well-known names as Fujitsu, Nintendo, and Toyota, which signal online news items within the NOW corpus. Because the corpus could be searched by country, the frequency and density of *JapanCo*, as well as individual company names within the set, revealed profiles for how Japanese companies occur in each web news market, and how those markets are shaped. One of the questions asked, was do companies display similar or different profiles in certain types and sizes of markets?

Specifically, do web news markets that are geographically closer to Japan display more interest in Japanese companies than do markets in places farther away? Likewise, do news markets closer to Europe publish fewer articles regarding Japanese companies? To draw such comparisons, the study employed 53 keywords that uniquely identified European-based companies. The sample is referred to as *EuroCo*, and includes such well-known, international names as Chanel, Lloyds, and Volkswagen.

Graphical Representation

The primary purpose was to graphically represent the data, rather than critically evaluate it, given that evaluation requires a richer context and background for proper assessment. In either case, whether representing or evaluating, a principal research category was constructed. In this case, the search category was *JapanCo*, a set of 66 unique terms that corresponded to names of Japanese companies (see Table 1). The terms were selected from lists of top global companies, as conveyed in sources such as *Forbes* [7]. Each term had to point primarily to a Japanese company, thus terms like Suzuki were not part of the list, because the commonality of the name would point to a variety of news articles unrelated to the Suzuki Motor Corporation. For comparative purposes, a second category, *EuroCo*, was also constructed in the same way, which incorporated 53 unique terms that corresponded to specific names of European companies.

Table 1

Company Search Terms

Category	Search Terms in Category (Single-word unique search terms)
JapanCo = search terms for sampled variety of Japanese companies (66 terms)	Advantest, Aeon, Asahi, Astellas, Bandai, Bridgestone, Calbee, Casio, Daihatsu, Daikin, Daiwa, Denso, Docomo, Fanuc, Fujifilm, Fujitsu, Hitachi, Honda, Idemitsu, Isuzu, Itochu, Kawasaki, KDDI, Kenwood, Kikkoman, Kodansha, Komatsu, Konami, Konica, Kyocera, Marubeni, Mazda, Mitsubishi, Mitsui, Mizuho, Nidec, Nikon, Nintendo, Nippon[a], Nissan, Nomura[b], NSK, Omron, Panasonic, Pentax, Rakuten, Ricoh, Sankyo, Sanyo, Sega, Seiko, Shiseido, Softbank, Sony, Subaru, Sumitomo, Suntory, TDK, Toho, Toshiba, Toray, Toyo, Toyota, Uniqlo, Wacom, Yamaha
EuroCo = search terms for sampled variety of European companies (53 terms)	(Germany) Airbus, BASF, BMW, Bosch, Braun, Commerzbank, Daimler, Lufthansa, Merck, Siemens, Steinway, ThyssenKrupp, Volkswagen | (France) Carrefour, Chanel, Dior, Engie, Michelin, Paribas, Renault, Sanofi, Sodexo | (UK) Barclays, BP, HSBC, Lloyds, Prudential, Tesco, Vodafone | (Italy) Armani, Benetton, Eni, Fiat, Gucci, Unicredit, Versace | (Switzerland) Glencore, Nestle, Novartis, Rolex, Suisse | (Spain) Repsol, Telefonica | (Netherlands) Rabobank, Unilever | (Belgium) Anheuser-Busch | (Luxembourg) ArcelorMittal | (Denmark) Maersk | (Norway) Statoil | (Sweden) Ericsson, Volvo | (Finland) Nokia | (Russia) Gazprom
Note. Country names listed in *EuroCo* = location of home offices for the respective companies. [a] In News on the Web (NOW) sources, the search term Nippon points primarily to articles regarding Reliance Nippon Life Insurance, Ltd. (RNLI). It also covers All Nippon Airways (ANA) and Nippon Telegraph & Telephone (NTT). [b] In NOW sources, the search term Nomura points predominately to Nomura Holdings, Inc. (Nomura Securities Co.) and its subsidiaries.	

The database consulted for the study was the NOW corpus (News on the Web) [6]. The corpus is composed of over 5.5 billion words of data, drawn from hundreds of online news publications across 20 nations. The current research was limited to 15 of these nations (see Table 2), disregarding five small markets that seemed less relevant to the study. Because NOW is updated daily, most of the data for the study ranges from the beginning of the corpus (1 January 2010) to when the data were extracted for the study (25 December 2017). A few residual calculations were made from additional data, drawn two weeks later (8 January 2018), which negligibly affected the graphics and assessments.

Table 2

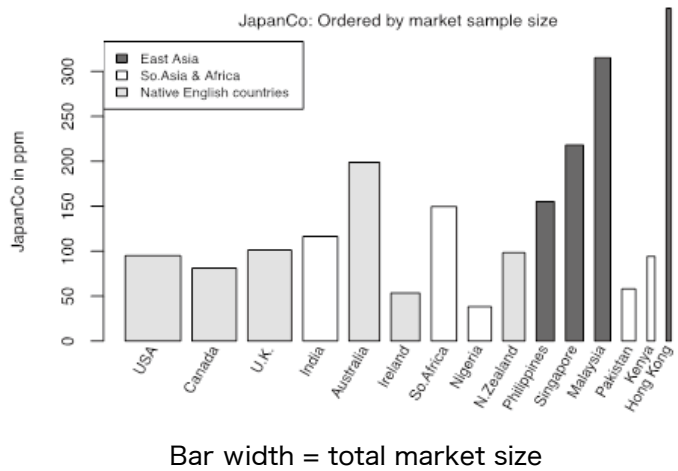
Japanese Companies in English Web News: By Market Size (2010-2017)

English NOW Market (2010-2017)	Market sample (in millions of words)	Frequency of *JapanCo*	ppm of *JapanCo*
United States	890.9	84,576	94.93
Canada	732.0	59,436	81.20
Great Britain (UK)	728.8	73,711	101.14
India	593.3	69,044	116.37
Australia	419.8	83,480	198.86
Ireland	397.2	21,251	53.50
South Africa	316.0	47,239	149.49
Nigeria	251.2	9,611	38.26
New Zealand	226.9	22,326	98.40
Philippines	186.4	28,960	155.36
Singapore	184.3	40,191	218.07
Malaysia	167.8	52,860	315.02
Pakistan	155.2	8,993	57.94
Kenya	80.2	7,559	94.25
Hong Kong	22.7	9,053	398.81
Grand total	5,352.7	618,290	94.25

Note. JapanCo = 66 single terms that point to Japanese companies. Frequency = raw occurrence of terms. Ppm = parts per million (i.e., density of *JapanCo* in each market).

Figure 2

Japanese Companies (JapanCo) in English Web News: By Market Size



Note. Bar width = size of NOW market (e.g., USA is largest English web news market). Bar height = density in ppm for in each market (e.g., Malaysia has 2nd highest density for *JapanCo*).

Table 3

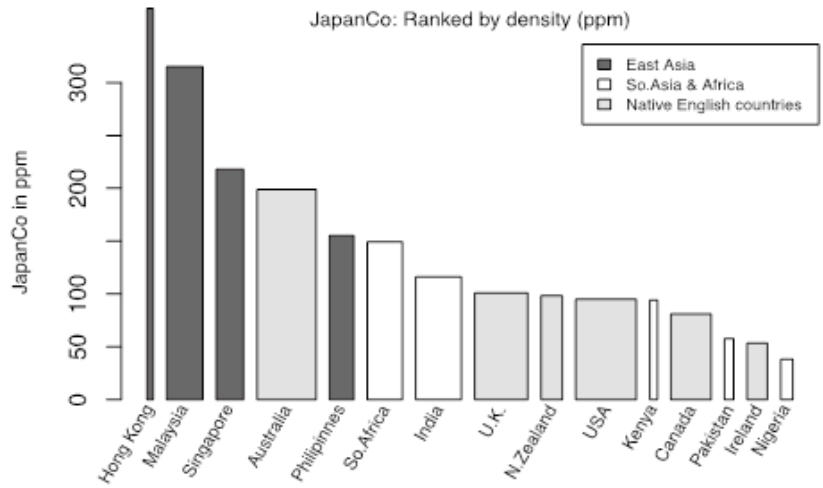
Japanese Companies (JapanCo) in Web News: Listed by Density (ppm)

English NOW Market (2010-2017)	PPM of *JapanCo*	Frequency of *JapanCo*	Market size (in mil. of words)
Hong Kong	398.81	9,053	22.7
Malaysia	315.02	52,860	167.8
Singapore	218.07	40,191	184.3
Australia	198.86	83,480	419.8
Philippines	155.36	28,960	186.4
South Africa	149.49	47,239	316.0
India	116.37	69,044	593.3
Great Britain (UK)	101.14	73,711	728.8
New Zealand	98.40	22,326	226.9
USA	94.93	84,576	890.9
Kenya	94.25	7,559	80.2
Canada	81.20	59,436	732.0
Pakistan	57.94	8,993	155.2
Ireland	53.50	21,251	397.2
Nigeria	38.26	9,611	251.2

Note. As with Table 2, these are composite scores for the eight-year period, 2010-2017.

Figure 3

Japanese Companies in English Web News: Density (ppm) over Frequency



Note. Bar width = total frequency of *JapanCo* per NOW market. Bar height = density in ppm for *JapanCo* in market (e.g., Hong Kong posts the lowest total frequency but the highest density).

Process. Once the *JapanCo* and *EuroCo* sets were constructed, each was imported as keyword search lists into the online BYU search engine for the NOW corpus. The lists were applied globally, and to each of the 15 individual nation web-news markets. The subsequent data were transferred into Microsoft Excel spreadsheets for management. All calculations were then conducted through Excel and the R Studio interface (an R language application). All figures were drawn primarily through graphical programming in R Studio.

Results. The total frequencies for *JapanCo* and *EuroCo* are depicted in Figure 4, and the top 30 search terms are listed in Table 4. Although *EuroCo* encompassed fewer terms than *JapanCo*, the fact that the European names were drawn from an entire continent, rather than a single nation, resulted in a greater frequency of *EuroCo* occurrences (868,529) compared to *JapanCo* occurrences (618,290) (as of 25 December 2017). As a result, an adjustment factor of 0.711882 was applied to draw fairer frequency and market-ratio comparisons in Figures 4 and 5. In particular, Figures 3 and 5 draw attention to the relationship between *JapanCo* in the news and the geographical location of the news sources.

Japanese and European Comparisons

The study also compared the way Japanese and European companies appear in NOW sources. The top 30 terms from each are ranked globally, in Table 4. The top three *JapanCo* terms in global web news were Sony, Toyota, and Honda, whereas the top three *EuroCo* terms were BMW, Barclays, and Nokia.

Table 4

Top 30 Search Terms: NOW Frequencies and ppm Ranked Globally

ppm	Frequency	*JapanCo*	Rank	*EuroCo*	Frequency	ppm
19.98	106,949	Sony	1	BMW	80,915	15.12
17.74	94,965	Toyota	2	Barclays	67,608	12.63
13.58	72,689	Honda	3	Nokia	59,942	11.20
11.24	60,151	Nintendo	4	Volkswagen	50,518	9.44
9.17	49,097	Nissan	5	BP (British Petrol.)	44,940	8.40
4.40	23,554	Mitsubishi	6	Renault	41,689	7.79
4.23	22,651	Mazda	7	Vodaphone	40,957	7.65
2.71	14,510	Toshiba	8	HSBC	37,258	6.96
2.65	14,161	Subaru	9	Suisse	35,213	6.58
2.61	13,968	Nomura	10	Tesco	34,026	6.36
2.57	13,744	Panasonic	11	Airbus	33,951	6.34
2.29	12,265	Yamaha	12	Volvo	28,649	5.35
2.16	11,553	Softbank	13	Fiat	22,286	4.16
1.86	9,973	Mizuho	14	Unilever	18,577	3.47
1.32	7,083	Sega	15	Chanel	16,930	3.16
1.29	6,899	Nikon	16	Bosch	15,133	2.83
1.23	6,565	Hitachi	17	Nestle	13,997	2.61
1.20	6,405	Fujitsu	18	Prudential	13,823	2.58
1.17	6,280	Kawasaki	19	Siemens	13,626	2.55
0.99	5,275	Isuzu	20	Lloyds	13,216	2.47
0.87	4,657	Nippon	21	Ericsson	13,146	2.46
0.75	3,994	Bridgestone	22	Glencore	12,150	2.27
0.73	3,930	Ricoh	23	Michelin	12,004	2.24
0.71	3,789	Sumitomo	24	Gucci	10,971	2.05
0.69	3,696	Mitsui	25	Dior	10,166	1.90
0.67	3,595	Asahi	26	Lufthansa	10,078	1.88
0.64	3,447	Fujifilm	27	Merck	9,675	1.81
0.63	3,355	Uniqlo	28	Paribas	9,330	1.74
0.62	3,341	Docomo	29	Braun	8,911	1.66
0.59	3,154	Konami	30	Rolex	8,811	1.65

Note. Frequency = global NOW occurrence of each search term (among the 15 nations sampled). Each ppm was calculated from the 15-nation market size of 5.35 billion total words.

Figure 4

Japanese and European Companies: Compared Total Frequencies

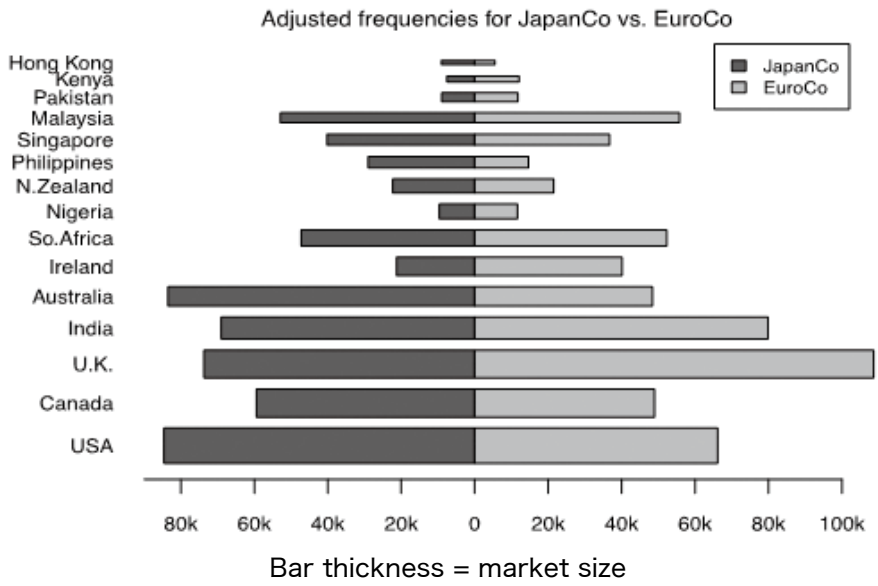
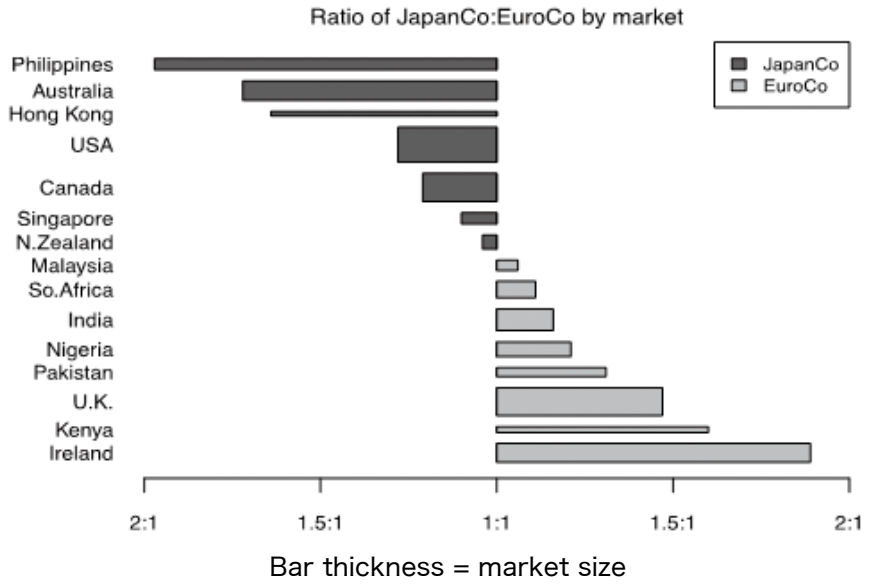


Figure 5

Japanese and European Companies: Preference Ordered by Market Ratio



Note. The axis indicates a compensated 1:1 occurrence between *JapanCo* and *EuroCo*.

The total frequencies for *JapanCo* and *EuroCo* are depicted in Figure 4. Although *EuroCo* encompassed fewer terms than *JapanCo* (see Table 1), the European terms represented a larger set, which is reflected in frequency and ppm in Table 4. To draw useful comparative representations between the two sets, a compensating calculation was

employed to center a 1:1 ratio axis.

Figures 4 and 5 indicate that news sources of the Pacific Rim tend to favor reports on Japanese companies, whereas online news from Europe, Africa, and South Asia show more preference toward European companies. This is perhaps unsurprising, but some of the comparative ratios, like those for the Philippines and Australia, were noticeably divergent from non-Pacific Rim areas. The graphs suggest that geographic location is relevant to which companies get reported in which news markets.

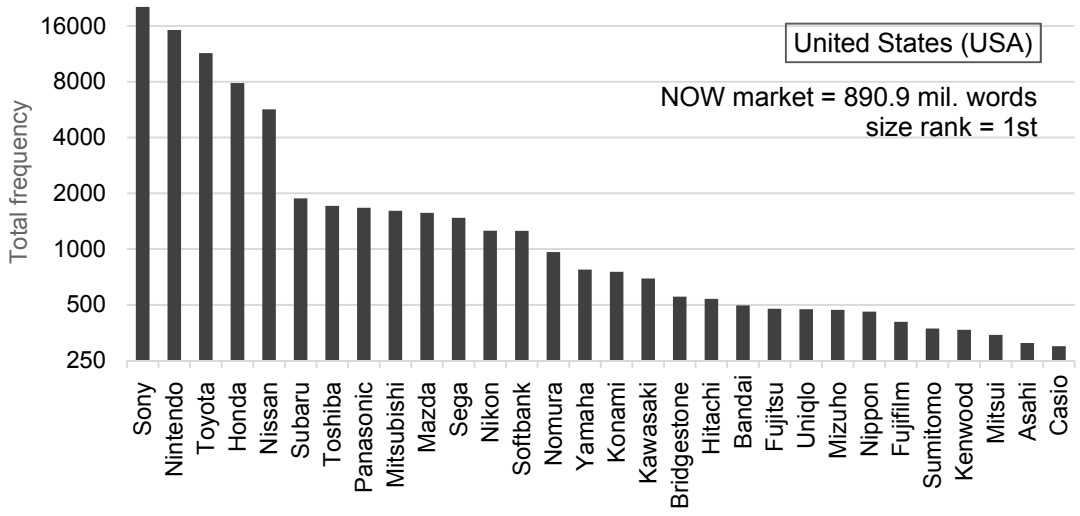
Individual Japanese Companies

Just as Table 4 indicated the top 30 search terms for *JapanCo* and *EuroCo* globally, Figures 6 through 11 show the top *JapanCo* selections from six web news markets. Three of these are for native English-language countries (Figures 6, 7, 8), and three are for non-native (Figures 9, 10, 11). The native group tended to be dominated by similar top performers like Sony and Toyota. By comparison, the rankings of search terms in the non-native markets tended to be more varied, with high-frequency showings by names such as Mizuho, Nomura, Softbank, and Yamaha. The sixth-place ranking of Yamaha in India was unsurprising (Figure 9), given the country's large, growing market for imported motorbikes, predominately from Japan [8]. Likewise, the common appearance of Nomura in Malaysia (Nomura Securities Malaysia) [9], as well as in Singapore (Nomura Trust, Ltd.) and Hong Kong (Nomura International, Ltd.), coincided with the company's global holdings, and indicated the tendency of these markets toward reporting financial news.

Subsequently, Figures 12, 13, 14, and 15 isolated four *JapanCo* terms (Toyota, Yamaha, Nintendo, and Fujitsu), and showed how each performed in all 15 news markets. The results were varied, and telling of how each company was reported in different online markets. For instance, Toyota's profile (Figure 12) matched expectations, somewhat dominant around the Pacific Rim. By comparison, the Yamaha frequency (Figure 13) was more focused in a varied handful of news markets, including India and South Africa, two places where only Japanese automotive companies tended to be evenly represented. Likewise, Nintendo (Figure 14) made its dominant showing in North America and other native-English new markets. By comparison, Fujitsu (Figure 15) displayed its own unique pattern, being heavily represented in Hong Kong and Singapore, where the company maintains important international divisions [10]. The name also held a sizeable news market share in the U.K., reflecting acquisitions by Fujitsu Ltd. of a number of European IT companies over the past two decades, through which Fujitsu conducts a great deal of its research and development [11].

Figure 6

JapanCo: Top 30 Total Frequencies in the United States



Note. Total frequency (vertical axis) uses a logarithmic scale (base 2).

Figure 7

JapanCo: Top 30 Total Frequencies in Great Britain

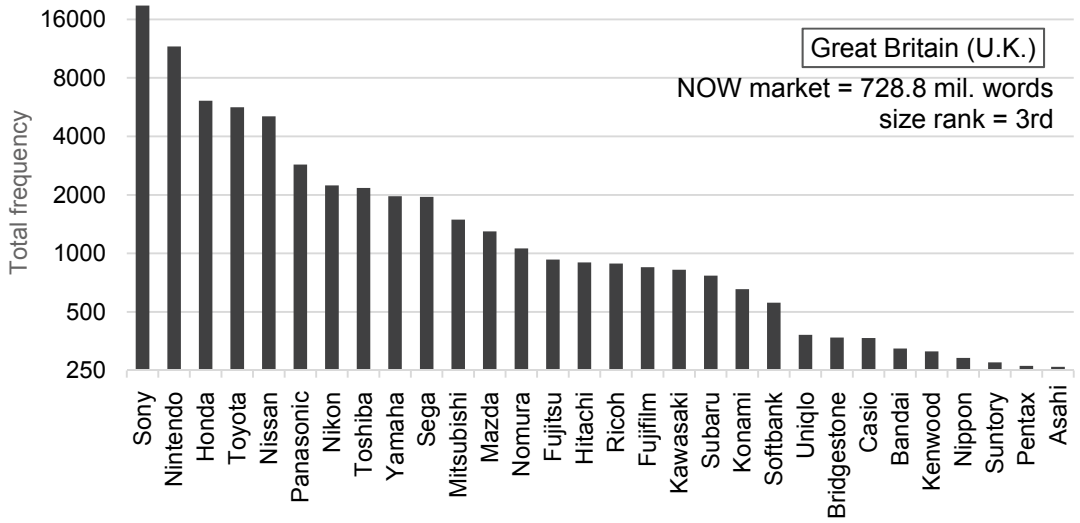


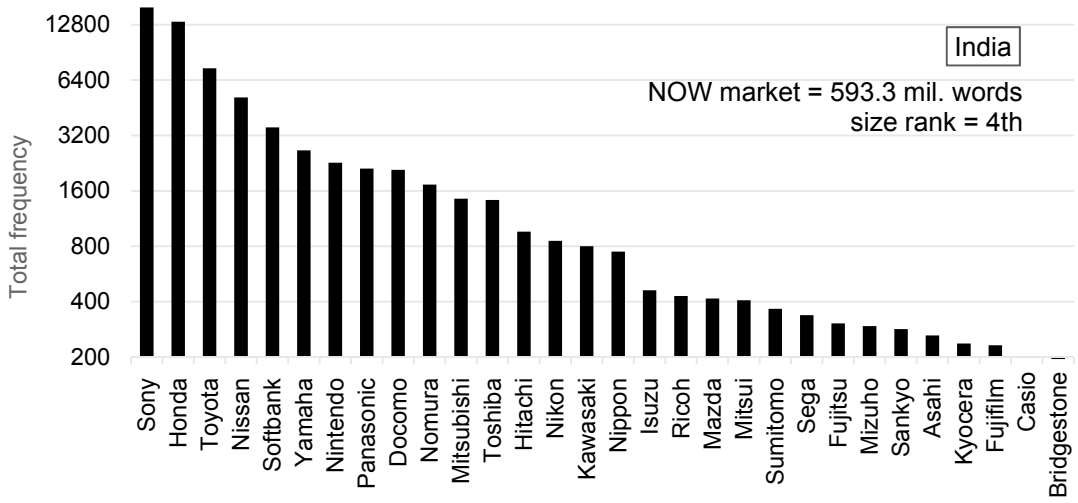
Figure 8

JapanCo: Top 30 Total Frequencies in Australia

Note. Total frequency (vertical axis) uses a logarithmic scale (base 2).

Figure 9

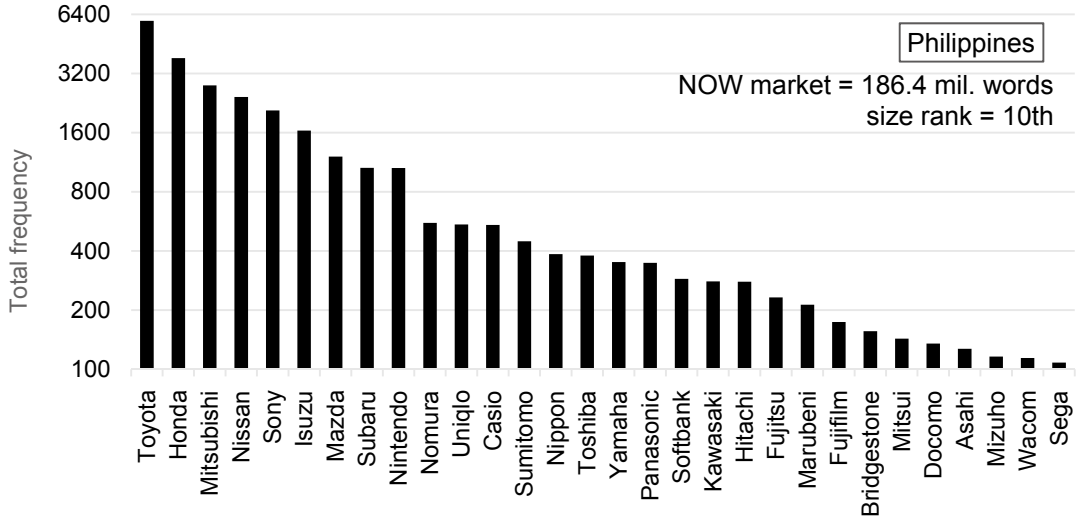
JapanCo: Top 30 Total Frequencies in India



Note. Total frequency (vertical axis) uses a logarithmic scale (base 2).

Figure 10

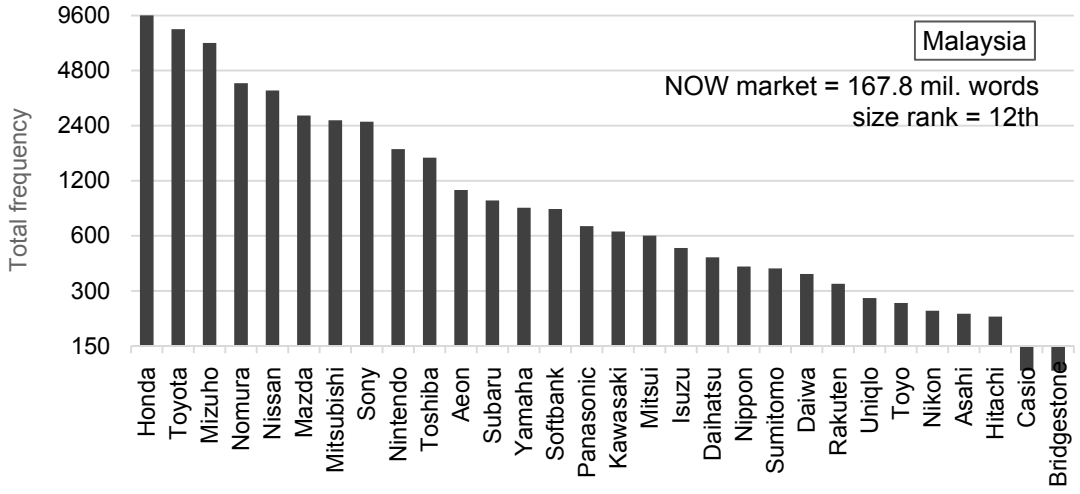
JapanCo: Top 30 Total Frequencies in the Philippines



Note. Total frequency (vertical axis) uses a logarithmic scale (base 2).

Figure 11

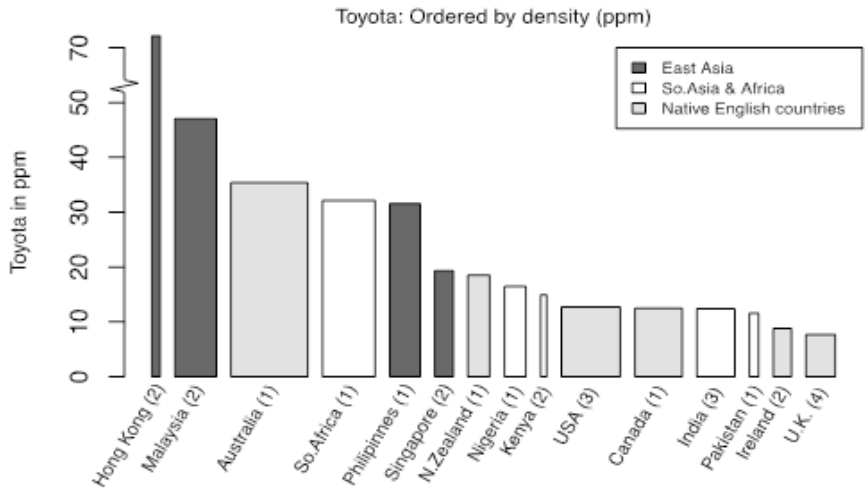
JapanCo: Top 30 Total Frequencies in Malaysia



Note. Total frequency (vertical axis) uses a logarithmic scale (base 2).

Figure 12

Toyota *in English Web News: By Density (ppm) (2010-2017)*

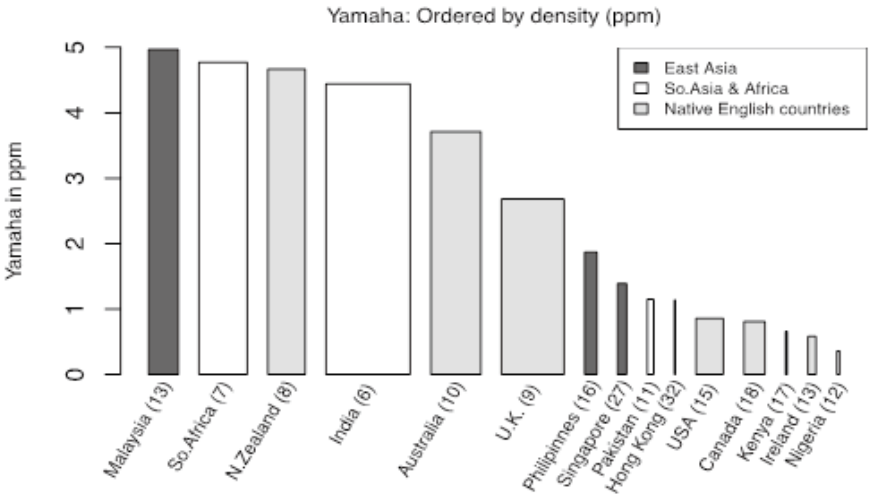


Bar width = total frequency of *Toyota* per market

Note. Bar width = total frequency. Bar height = density in ppm. Number next to country (in parenthesis) indicates the term's *JapanCo* market position (e.g., Toyota is the number-one *JapanCo* term in Australia, where it posts the largest total frequency and the third-highest density).

Figure 13

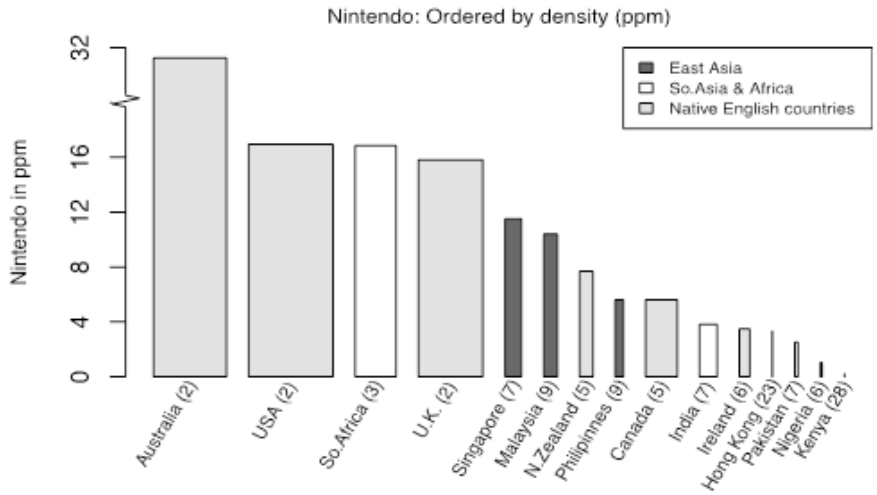
Yamaha *in English Web News: By Density (ppm) (2010-2017)*



Bar width = total frequency of *Yamaha* per market

Figure 14

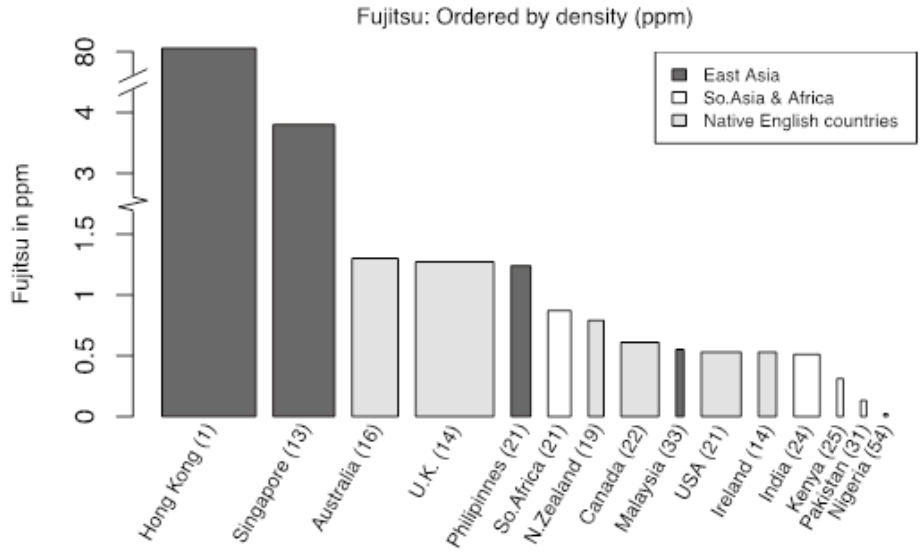
Nintendo *in English Web News: By Density (ppm) (2010-2017)*



Bar width = total frequency of *Nintendo* per market

Figure 15

Fujitsu *in English Web News:* By Density (ppm) (2010-2017)



Bar width = total frequency of *Fujitsu* per market

Note. Hong Kong *Fujitsu* ppm = 81.53; Singapore = 3.90; Australia = 1.30; etc.

Summary

One question for the study was whether geography has played a part in shaping the presence of Japanese companies in web news markets. Another was whether companies, and the *JapanCo* set as a whole, have responded differently in different news markets. As the tables and graphs suggested throughout, the answer to both questions appeared to be yes. Figures 3 and 5, in particular, visualized the relationship between *JapanCo* occurrences and geographical regions.

In conclusion, the data depicted throughout the study revealed a great deal of information through visual representation alone. Subsequent analysis, of news markets and of Japanese business presence in global news, could develop from what has so far been indicated. The method was also intended to illustrate how a certain type of research could benefit business, market, economic, and general historical studies. Application toward the evolution of historical science was implied.

References

[1] Fairchild, T.E. (2017). “Seeking Japan in American English: Potential History in Word Frequencies.” *International Management Review (Kanagawa University),* 54, 127-144.

[2] Cohen, J.B. (1949). *Japan’s Economy in War and Reconstruction.* University of Minnesota Press, Minneapolis, MN.

[3] Pennebaker, J.W. (2011). *The Secret Life of Pronouns: What Our Words Say about Us.* Bloomsbury Press, New York.

[4] Blatt, B. (2017). *Nabokov’s Favorite Word Is Mauve: What the Numbers Reveal about the Classics, Bestsellers, and Our Own Writing.* Simon and Schuster, New York.

[5] Aiden, E., & Michel, J.P. (2013). *Uncharted: Big Data and an Emerging Science of Human History.* Riverhead, New York.

[6] Davies, M. (2017). *NOW Corpus (News on the Web)* [Database]. Available online at http://corpus.byu.edu/coha/

[7] *Forbes.* (2017). “The World’s Biggest Public Companies” [Webpage]. Retrieved from http://www.forbes.com

[8] *IM (India Marks).* (2017). “Top Seven Imported Bikes in India.” [Webpage]. Retrieved from http://www.indiamarks.com

[9] *Bloomberg*. (9 January 2018). “Company Overview of Nomura Securities Malaysia Sdn. Bhd.” [Webpage]. Retrieved from http://www.bloomberg.com

[10] Fujitsu. (2017). Fujitsu U.K. [Website]. Retrieved from http://www.fujitsu.com/uk

[11] Fujitsu. (2017). Fujitsu Hong Kong [Website]. Retrieved from http://www.fujitsu.com/hk