The *Fashionable Functions Reloaded.*
An Updated Google Ngram View of Trends in Functional Differentiation (1800-2000)*

Steffen Roth  
*ESC Rennes School of Business, France*

Carlton Clark  
*University of Wisconsin-La Crosse, USA*

Jan Berkel  
*Independent, Portugal*

**ABSTRACT**

Using the updated Google Book corpus dataset generated in July 2012, we analyze the largest available corpus of digitalized books to review social macro trends such as the secularization, politicization, economization, and mediatization of society. These familiar trend statements are tested through a comparative analysis of word frequency time-series plots for the English, French, and German language area produced by means of the enhanced Google Ngram Viewer, the online graphing tool that charts annual word counts as found in the Google Book corpus. The results a) confirm that the importance of the political system, religion, economy, and mass media features significant change in time and considerable regional differences and b) suggest that visions of economized or capitalist societies are intellectual artifacts rather than appropriate descriptions of society.

**Keywords:** functional differentiation; Google ngram; culturomics; capitalism; economization; secularization; politicization; mediatization.

**INTRODUCTION**

In this article, we draw on the second version of the Google Book corpus dataset to show how big data analysis may check and challenge old familiar self-definitions of modern society. The starting point of our venture is the distinction between autonomous function systems such as the economy, science, art, religion, etc. In fact, this form of functional differentiation is considered a core concept of modern societies (Leydesdorff, 2002; Beck et al., 2003; Berger, 2003; Vanderstraeten, 2005; Brier, 2006; Baecker, 2007; Kjaer, 2010; Berghaller & Schinko, 2011; Roth, 2015a). Without functional differentiation, there would be no difference between truth and money, a hospital would be considered the same as a bank, and there would be no sense in the critiques of doping, corruption, or the selling of indulgences. In like manner, the larger part of

---

contemporary definitions and criticisms of modern society would have to do without their most basic categories, since all observations of secularization, economization, and mediatization implicitly refer to an underlying concept of functional differentiation.

Though generally accepted, the idea that particular function systems are more relevant to society than others is not understood without ambiguity. On the one hand, in the light of the fundamental equivalence (Vanderstraeten, 2005; Jönhill, 2012) and autonomy (Tsivacou, 2005; Valentino, 2012) of the function systems there is no way of arguing that the political system or the economy is essentially more important than religion or sport, per se. On the other hand, there seems to be plenty of empirical evidence of such imbalances in terms of the just mentioned trend observations.

This contradiction can be resolved by stating that it is not despite, but because of their basic equivalence that function systems can be ranked at all because if the function systems were essentially unequal, they would already be ranked and, therefore, could no longer be ranked. In this sense, the function systems can be treated as nominal data that feature a skewed distribution whenever it comes to the analysis of concrete segments of society. Hence, modern societies so far have been defined in terms of different biases to particular function systems with the most prominent cases being the definition of society as capitalist (Roth, 2015b). Though there is still little consensus on the question of whether capitalism results either from the primacy of a particular form of politics or from the primacy of the economy (Risse, 2003; Wallerstein, 2003; Foucault, 2008; Urry, 2010; Lash, 2007; Madra & Adaman, 2014), most people would basically agree on the idea that present societies are subject to an economization of collective goals (Alexander, 1985). This “increasing influence of economic factors and values on the political agenda and other areas of society” (Blumler & Kavanagh, 1999:210) includes the economization of

- Health (Musick, 1999; Cartier, 2003; Ewert, 2009; Brown et al, 2011),
- Art (Velthuis, 2003; Behnke, 2007; Eikhof & Haunschmid, 2007; Kjaer, 2010; De Valick, 2014),
- Science (Penders et al., 2009; Berman, 2013),
- Education (Fludernik, 2005; Wilkesmann & Schmid 2012; Spring, 2015),
- Religion (Robertson, 1992; Wannenwetsch, B. 2008), and, as a matter of course,
- Society as a whole (Polanyi, 1957; Habermas & McCarthy, 1985; Schmidt, 1993; Altvater & Mahnkopf, 1996; Enderle, 1997; Chomsky, 1999; Sayer, 1999; Finch, 2007; Schimank & Volkmann, 2008).

As a result of this “economic turn” (Smart, 2003) or fetishization of the economy (Foucault, 2008), economization emerges so omnipresent and dominant that even the proliferation of economics is taken for an indicator of economization (Çalışkan & Callon, 2010) or a need of de-economization (Latour, 2004), respectively. Nonetheless, there is also discussion on further forms of trend statements and predictions, which includes sometimes concurrent, sometimes competing definitions of society as mediatized (Dennis, 1978; Eaman, 1987; Castells, 1996; Chomsky, 1997; Blumler & Kavanagh, 1999; Croteau & Hoynes, 2003; Schulz, 2004; Hjarvard, 2008, 2013; Mazzoleni, 2008; Moon, 2012; Esser & Strömbäck, 2014), politicized (Chomsky, 2000; Blumler & Kavanagh, 1999), militarized (Regan, 1994; Young, 2007; Levy, 2010), intellectualized (Alexander, 1985), or even aestheticized (Blumler & Kavanagh, 1999; Rocha de Oliveira, 2009). Anything seems to go as long as there is not too much religion involved. Both the
economization (Robertson, 1992) and the politicization (Thompson, 2006) of religion seem to be in line with the project of modernity. Any sign of religious recovery, however, apparently understood as a threat to modernity (Martin, 2005; Martin, 2011; Bracke, 2008), even though such a “de-secularization” would not challenge functional differentiation and modernity in a more profound way than a possible “economization of every sphere of existence” (Kane, 2010:81).

In view of the literature, the long-term trends in the discussion of –izations are indeed the politicization and, most prominently, economization, which take place against the background of a requiem for religion and are recently complemented by an emerging media boom.

The problem with the corresponding marginalization of the “other areas of society” (Blumler & Kavanagh, 1999:210) is not only in the fact that “(o)ther social institutions are seen (once again) as mere puppets in the hands of powerful economic trends and actors” (Stehr, 2002:4) or as subjects to the respective political counter-performances, but also in the empirical basis of this discursive bias. So far, the existence of trends is supposed rather than studied. As well, even the few contributions that call for more “observational research” (Blumler & Kavanagh, 1999:225) are so much focused on the analysis of dominant or strong function systems that they do hardly reflect why the respective focus system should be preferred to others and therefore lack the overview Poul Kjaer (2010:532) is interested in the analysis of function systems:

“(T)he development of a general theory capable of linking them systematically together. When observed in isolation, the mutual supportive character of these dimensions is not obvious. Only a more general conceptual framework will make it possible to empirically observe to what degree the observed phenomenon constitute or potentially will be capable of constituting a ‘higher order’ (...).”

In order to approach such a general framework and compensate for the lack of evidence for the respective diagnoses, the present article will consider both the general idea of biased constellations of function systems and the observation of particular trends in the significance of individual function systems. Both are considered hypotheses that still need to be defended.

Subsequent to a theory statement, the present article asks for the empirical validity of the most popular statements of trends in functional differentiation. Namely the secularization, politicization, economization and mediatization of society will therefore be re-conceptualized in terms of hypotheses and tested against the results of a Google Ngram Viewer analysis of the most frequent function system references in the updated second version of the Google Books corpus for the years 1800 through 2000. The results not only show that the importance of individual function systems changes over time and across language areas, but also give reason to believe that popular observations of an increasing importance of the economic systems are intellectual artifacts rather than appropriate (self-) descriptions of society.

**THEORY STATEMENT**

“Ever since there has been sociological theory it has been concerned with social differentiation” (Luhmann, 1990:423). Differentiation refers to an intrasystem process of subsystem formation (Luhmann, 1977, 1997b). The first known forms of subsystems within society were families and tribes. Families coexisted with other families of the same tribe in the same manner as tribes coexisted with other tribes. As a result, early societies are said to be differentiated into identical
and co-equal segments of society. However, in the course of the Neolithic revolution, the situation changed as soon as location decisions condensed to locational advantages and disadvantages in such a way that individual settlements ascended towards centers, while others turned into periphery. This center-periphery differentiation was soon complemented and superposed by stratification, which is commonly associated with the formation of hierarchical social orders, such as the Indian cast system or the Occidental Estates of the realm. Such stratified societies defined persons into ranked hereditary communities and allowed for only limited social mobility, if any at all. Conversely, in the European case, it was mobility that finally changed the static order. Be it the movable types of Gutenberg’s printing press, the Central-European rural exodus, or the fact that too many commoners had been granted entry into the gentry, in all cases, the constitutive distinction of nobles and commoners was for its own part superposed by a distinction so fundamental to modernity that animals can no longer be divided in the following way:

“(a) belonging to the Emperor, (b) embalmed, (c) tame, (d) sucking pigs, (e) sirens, (f) fabulous, (g) stray dogs, (h) included in the present classification, (i) frenzied, (j) innumerable, (k) drawn with a very fine camelhair brush, (l) et cetera, (m) having just broken the water pitcher, (n) that from a long way off look like flies” (Borges & Weinberger, 1999:231).

In fact, modern man considers it absurd to assume that a siren and a pig represent the same type of animal as soon as they both belong to the emperor or are drawn with the same kind of brush. “Modern society is no longer characterized by a stratification of lineage, clans and families, but by a differentiation of function systems” (Vanderstraeten, 2005:476). Today, society therefore cannot only be differentiated into a) similar and equal segments, b) similar and unequal centers and peripheries, and c) dissimilar and unequal strata, but into d) dissimilar and equal function systems, as well (see Figure 1).

![Figure 1. Social Differentiation (source: Roth, 2015b:113)](image)

Despite their obvious existence, and despite their importance to modern society, there is still little consensus on necessary and sufficient conditions for the definition of function systems. Nevertheless, when looking at existing working definitions and non-exhaustive lists of function systems (Reese-Schäfer, 1999:176f, 2007:120; Künzler, 1987:327, 1989:100f; Andersen, 2003:159; Stichweh, 2005; Baecker, 1994; Henkel, 2010:183; Luhmann, 1997a:11), a list of 10 function systems can be extracted (Roth and Schütz, 2015): Political system, economy, science, art, religion, legal system, health, sport, education, and mass media. What defines a function
system is “its function for society, which is constitutive for even the most basic legal operations and thus for the autopoiesis of the legal system. This is why the criterion of function systems is that they distinguish themselves by their relation to society, and not by their relation to themselves or to any particular subsystems of society” (Roth and Schütz, 2015:18). Each of the 10 function systems differentiates society by the binary re-coding of communication according to a specific symbolically generalized communication medium. Each function system applies only one single code, which it also applies exclusively. For example, science, and only science, is all about the medium truth, which is binary coded as true or untrue. Scientific programs and theories decide on when the code of science is properly applied. The function of science is to provide society with ongoing knowledge communication (see Tab. 1).

**Table 1: The Function Systems of Society (source: Roth and Schütz, 2015:24)**

<table>
<thead>
<tr>
<th>System</th>
<th>Code</th>
<th>Medium</th>
<th>Program (ex.)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Political System</td>
<td>Government/opposition</td>
<td>Power</td>
<td>Ideology</td>
<td>Limitation</td>
</tr>
<tr>
<td>Economy</td>
<td>Payment/non-Payment</td>
<td>Money</td>
<td>Price</td>
<td>Distribution</td>
</tr>
<tr>
<td>Science</td>
<td>True/untrue</td>
<td>Truth</td>
<td>Theory</td>
<td>Verification</td>
</tr>
<tr>
<td>Art</td>
<td>Innovative/imitative</td>
<td>Style</td>
<td>Fashion</td>
<td>Creation</td>
</tr>
<tr>
<td>Religion</td>
<td>Immanent/transcendent</td>
<td>Faith</td>
<td>Confession</td>
<td>Revelation</td>
</tr>
<tr>
<td>Legal System</td>
<td>Lawful/unlawful</td>
<td>Norm</td>
<td>Law</td>
<td>Standardization</td>
</tr>
<tr>
<td>Sport</td>
<td>Success/failure</td>
<td>Achievement</td>
<td>Goal</td>
<td>Mobilization</td>
</tr>
<tr>
<td>Health System</td>
<td>Ill/healthy</td>
<td>Illness</td>
<td>Diagnosis</td>
<td>Restoration</td>
</tr>
<tr>
<td>Education</td>
<td>Placeable/unplaceable</td>
<td>Vita</td>
<td>Curriculum</td>
<td>Formation</td>
</tr>
<tr>
<td>Mass Media</td>
<td>Informative/Non-informative</td>
<td>Medium</td>
<td>Topic</td>
<td>Multiplication</td>
</tr>
</tbody>
</table>

While it can be considered impossible to belong to both the estate of the nobles and the estate of the commoners or to two castes at the same time, persons can be included in all of the function systems. This multi-inclusiveness of function system applies to organizations as well, which are, also regarded as multi-referent (Tacke, 2001; Simsa, 2001), polyphonic (Thygesen & Andersen, 2007; Andersen & Born, 2007), or multimedia organizations (Roth et al., 2010; Roth 2012; 2014b).

Combined with the basic assumption that no function system is essentially more important than the other, the multimedia character of persons, organizations, and further segments of society are the basis of the observation of individual differences in the frequencies of function system references. It is precisely because of the mutual exclusiveness and functional equivalence of the individual function systems that individual persons, organizations, nation states, or cultural areas are likely to feature different frequency distributions of particular function system (p)references. “On this background, it is possible to understand asymmetries, crowding-out effects, and negative externalities between functionally differentiated spheres as a central source of tension and conflict in modern society” (Kjaer, 2010:494). The present article will explore these tension zones and, in doing so, regard society as a system of accessible communications. In this sense, the article takes the language border as a better distinction than national, i.e. geopolitical borders, and will refer to the English, French, and German language areas as societies.
HYPOTHESES

The basic assumption of the functional equivalent and mutually exclusive nature of function systems represents an excellent groundwork for the null hypothesis, which the present article proposes to test. Representing coequal nominal data, function systems can be assumed equally relevant to social systems. The null hypothesis is, therefore, as follows:

\((H_0)\) Function systems relevancies exhibit a uniform distribution in social system.

On the other hand, empirical research seems to indicate an unequal distribution of function system relevancies. The alternative hypothesis reads as follows:

\((H_1)\) Function systems relevancies exhibit an unequal distribution both in the course of time \((H_{1.1})\) and across cultures \((H_{1.2})\).

Pursuing the alternative hypothesis and linking to the discourse on major trends in functional differentiation, the article, furthermore, considers the assumed unequal distributions subject to constant change similar to “changes in economic cycles and international competition [that] create preferences for certain kinds of language and explanations” (Cornelissen and Kafouros, 2008:14). The second alternative hypothesis, thus, suggests testing the most popular common senses on trends in functional differentiation:

\((H_2)\) Societies feature relatively stable trends in functional differentiation, including \((H_{2.1})\) the secularization, \((H_{2.2})\) the politicization, \((H_{2.3})\) the economization, and \((H_{2.4})\) the mediatization of society.

The second hypothesis will be tested against both weak and strong definitions of the respective trends. Trends will, consequently, be analyzed with regard to their incidence in one or several cultural areas as well as to their individual course and their relative importance.

METHOD

The key assumptions proposed in this article are the temporal change of and the intercultural differences in the importance of particular function systems to particular societies. The major problem becomes how to measure this importance in a context that can be expected to allow for somewhat representative information on societies as a whole.

The importance of concepts is often defined in terms of the frequency of their occurrence in given corpora, which is considered “the simplest and most impartial gauge of word importance” (Kloumann et al., 2012:1) or the popularity of objects, ideas, and persons (Ophir, 2010; Bohannon, 2011). Given the scope of the present research question and the corresponding scale of the research field, research utilizing the Internet as the largest possible corpus would make sense. Despite the fact that web search engines are said to return word frequency estimates that are highly consistent with established methods (Blair et al., 2002), the problem with Internet word frequency analyses is that the analysis of this most recent media hardly allows for long-term trend analyses. Hence, the present discussion of popular trends in functional differentiation
analyses changes the frequency distributions of function system references in the second largest available text data collection.

Since its start in 2004, the Google Books project has digitalized some 15 million of the estimated 115 million books ever published. In 2007, a Harvard research team (Michel et al., 2011) recognized the research potential of the Google Books corpus, performed considerable quality checks, and finally reported the compilation of a representative corpus of more than five million books or 500 billion words covering seven language areas and a time span of 600 years. This corpus was later updated to version 2, issued in July 2012. The development of this enormous dataset soon raised hopes of a golden age of digital humanities (Johnson, 2010), which would open up new types of historical knowledge (Ophir, 2010); it has already given birth to the discipline of culturomics as “the application of high-throughput data collection and analysis to the study of human culture” (Michel et al., 2011:181). Most culturomic research addresses issues of literature, language, and culture studies (e.g. Gibbs & Cohen, 2011; Nicholson, 2012). Yet, there also have been attempts to bring culturomics to the political sciences, for example, by the use of the Google Book corpus for a retroactive forecasting of social movements such as the Arab Spring (Leetaru, 2011).

Figure 2. Word Frequency Shares of Deutschland, England, Frankreich, and USA in the 2009 and the 2012 German Language Google Books Corpus (source: own Google Ngram enquiries)
The access to the Google Books corpus is facilitated by the Google Ngram Viewer—an open-access interface that allows for *trending* (Manovich, 2012) in terms of the production of customized time-series plots for entered search terms.

In such way, the Google Ngram Viewer can be used to re-present well-known information in a compact and intuitive way (see Figure 2).

In contrasting the two charts combined in figure 2, we find that the patterns of the graphs remain relatively stable between the 2009 and the 2012 versions of the dataset; yet, there are visible differences, too, for example with regard to the performances of the USA and France in the 1990s. Moreover, figure 2 clearly illustrates how the importance of the concepts of Germany, England, France, and the USA varied between the year 1800 and 2000. The graph shows that it was not until the eve of World War I that the concept of Germany became dominant in German books. Furthermore, with regard to Germany, we find top peaks of popularity during the early Nazi era, the Wirtschaftswunder, and the German reunification as well as interim lows after the lost World Wars and in relation to the Protests of 1968. The overall trend, nonetheless, is positive, which is not the case for France and England. Both countries lost relevance after World War II and were overtaken by the USA in the mid-1980s. In this sense, the Google Ngram Viewer also allows for the qualitative analysis of quantitative data, e.g. in terms of the assumption that the USA became more important to Germany and perhaps was the most important occupying power.

Furthermore, the Viewer can be used to detect needs for research as much as for the developing of research questions (see Figure 3-4).

![Google Books Ngram Viewer](image)

*Figure 3. Word Frequency Shares of money, power, and love in the Overall English Language Google Books Corpus 2012 (source: own Google Ngram enquiry)*
Contrasting the German and the English distribution of the concepts of money (blue), power (red), and love (green) we find that love and money started to feature common characteristics in the English context as of the late 1940s, while love seems to be closer to power than to money in the German context, if at all. Far from representing proof for anything, the presented evidence can, nonetheless, be considered a reason for further research. One big advantage of the updated version of the Google Ngram Viewer is that the tool does to some extent allow for word context analyses now. This is useful in the case of power, since it helps with identifying the frequencies of non-functional meanings such as “electric power” and – if need be – with their exclusion from the sample (which, however, is not required in the present case because of the only marginal performances of non-political meanings of power). Again, a comparison of the present figures 3 and 4 with those in the original article (Roth, 2014) shows that the graphs perform almost identical in the English language case, whereas there are considerable, yet not fundamental, differences in the German language case.

In some cases, a Google Ngram View might even be used for validity tests of more or less prominent hypotheses (see Figure 5-7).
Figure 5. Word Frequency Shares of globali(z/s)ation and regionali(z/s)ation in the Overall English Language Google Books Corpus 2012 (source: own Google Ngram enquiry)

Figure 6. Word Frequency Shares of globalisation and régionalisation in the French Language Google Books Corpus 2012 (source: own Google Ngram enquiry)
For example, regionalization is sometimes said to be an effect of globalization (Kacowicz, 1999; Amin, 1999; Hurrell, 2007). The Google Ngram View of the respective concepts, however, does not support this idea. Rather, we find evidence for the opposite effect, being that in all the three language areas, the concept of regionalization enjoyed a modest popularity before globalization boomed, which is most evident in the French case, which featured a noticeable trend in regionalization between 1960 and 1975. It is not until the mid-1980s (US-American English) or the early 1990s (British English, French, German), respectively, that the globalization discourse reached the level of the respective regionalization discourses. Moreover, the findings suggest that the globalization discourse could have had its origins in the French discourse and the largest impact on the German discourse. In all these cases, the graphs feature similar patterns in both 2009 and 2012; what seems to have changed are only the absolute word frequencies (see e.g. the German graphs).

Testing its own hypotheses, the present article will focus on the analysis of trends in the 2012 versions of the English, French, and German Google Books 1-gram corpora 1800-2000.3

In this setting, the null hypothesis will be rejected in favor of the alternative hypothesis (H1) when the time-series plots produced by the Google Ngram Viewer feature an unequal distribution of the occurrence frequencies of all function system designations. Moreover, the plots will be analyzed with regard to variations of occurrences in the course of time (H1.1) and across cultures (H1.2).

The hypotheses H2.1-2.4 will be weakly confirmed if the output of the Google Ngram Viewer query supports the respective trend assumptions in at least two out of three language areas. In concrete terms, the plots should display a decrease in the importance of religion (H2.1: secularization) as well as an increase in the importance of the political system (H2.2: politicization), the economy (H2.3: economization), and the media (H2.4: mediatization),
respectively. A strong confirmation would, moreover, call for (the loss of) a dominant position in the texture of function system references for all language areas.

Meeting concerns that the function system designations might not represent the most frequent function system reference in the corpus, the Top 2000 1-grams of the English-language Google book corpus version 2012 will be scanned for further function system references. The list of the ten most frequent references to function systems relevant to the hypotheses H2.1-2.4 will also be entered into the Google Ngram Viewer and the result compared to the outcome of the query for the names of the function systems.

**RESULTS**

The ten function systems exhibit unequal occurrence frequencies that vary within and across all the language areas:

![Google Books Ngram Viewer](image)

**Figure 8. Word Frequency Shares of the Names of the Function Systems in the Overall English Language Google Books Corpus 2012 (source: own Google Ngram enquiry)**

In the English case (see Figure 8), at the beginning of the sample period, religion (not the legal system as erroneously reported in the original version) is the most dominant function system
followed by art and the political system. A major change to the earlier version of the article is that we replaced policy by political, which is a stronger and more adequate indicator of the importance of the political system. The performance of the political system towards the end of the sample period therefore appears even stronger than in the earlier version. The general picture, however, remains the same: the political system is dominant, followed by health, which overtakes education soon before the millennium. Currently, religion ranks ninth, outperformed even by the formerly marginal economy and mass media system. The performance of the economy is relatively less pronounced in the 2012 data as compared to the 2009 data.

A closer examination of the function systems relevant to the trend hypotheses 2.1-2.4 calls attention to the word frequency shares of the political system, the economy, science, religion, and the mass media system. The decline of religion ($\bar{r} = 0.27$ compared to 0.25 as visually assessed in the earlier version of this article)\textsuperscript{4} and the rise of the political system ($\bar{r} = 3.5/3.7$) are the most striking trends. Religion is outperformed by the political system during World War I and by science in the early 1930s, with the latter being a result of slow, but steady growth. At about the same point in time, the formerly marginal economy takes off ($\bar{r} = 3.57/5.00$ between 1930 and 1990). Yet, in 2000 economy and science share about the same numbers of mentions in the Google Books corpus (approximately 100ppm)\textsuperscript{5} and are, therefore, both lower-middle ranking function systems: The political system (320ppm), health (177ppm) and education (168ppm), and art (119ppm) occupy the first five ranks. On the verge of the millennium, religion (64ppm) was finally outpaced by the mass media (75ppm). Sport is last (16ppm).

In the French case (see Figure 9, next page), art initially ranks first in front of religion, justice, and the political system. Between World War I and World War II, it was overtaken by the political system, whose increase ($\bar{r} = 5.3/4.9$) to a frequency of 701ppm in 2000 is remarkable (the small counter-trend around 1980 is less pronounced and stable than in the 2009 data). The political system is ranked first uncontestably since the end of the 1920s. In 2000, the political system occurred three times more often than art. As of the late 1920s, after an increase of economic communication, the economy ranks third, closely followed by science, education, the legal system, health, and religion. Mass media and sport come last.

A comparison of the English and the French data uncovers the higher relevance the French function systems have in their corpus. The word "politique," however, must be considered an exception to this rule because it is both a noun and an adjective. Complemented by the adjective(s), the English and the German political performance would be about the same as the French. Still, it is notable that even "politique\text{\_\textsc{noun}}", the string used to count only cases in which the ngram "politique" appears as noun, remains the clear number one in the French corpus in the year 2000 (320ppm). Additionally, despite a considerable decline, second-ranked art (211ppm) is still more important to the French corpus than second and third ranked health and education are to the English.
The German data display the largest variances and the broadest scope of changes within the three language areas. In the early 19th century, Germany seems ‘governed’ by the legal system as well as science, art, and religion. The latter triad appears to share a common destiny in terms of their collective decline up until the 1960s, when science and art finally separated from the downturn of religion, which appears to have stopped since the 1980s. After a most notable increase ($\bar{t} = 7.9/7.0$), since the early 1960s, Politik (290ppm) has ranked first, followed by the legal system (249ppm) as well as art and science (141-161ppm). Economy (132ppm) is ranked fifth due to an increase in importance dating back to the late 19th century rather than to the late 1920s, like in the case of both the English and the French.
Figure 10. Word Frequency Shares of the Names of the Function Systems in the German Language Google Books Corpus 2012 (source: own Google Ngram enquiry)

However, there is also a media trend, which started in the mid-1960s, whose calculation in terms of a growth factor does not make much sense because, just like in the case of the English and the French, the media seems to appear from nowhere. In the German case, the same applies to the late 19th century rise of the economy. At least a comparison with the political curve seems revealing to some extent, as both curves feature a certain parallelism between World War I and the mid-1950s (which, however, appears clearer in the 2009 data).

Interim discussion

So far, the findings indicate both considerable changes within and significant differences between the language areas. Moreover, with regard to all language areas, the results seem to support the hypotheses H2.1 (secularization) and H2.2 ( politicization), even if there is some evidence for the fact that the trends have stopped towards the end of the sample period. Despite considerable increases in the frequencies of all the three designations of the economic system and the mass media system, both systems are still far from being dominant in their respective
language areas (see Figure 11). So far, the results would suggest rejecting the hypotheses H2.3 (economization) and 2.4 (mediatization), at least in their strong variant.

![Figure 11. Interfunctional Comparative Profiling of the English, French, and German Language Area](image)

**Results, continued**

In order to re-test the hypotheses, in particular H2.3 and 2.4, the present article presents the results of an additional analysis. We used a Python program coded by co-author Jan Berkel to extract a word frequency list of the entire English-language Google Ngram corpus data from 2012. For each of the function systems relevant to the hypotheses H2.1-2.4, the ten most frequent references in this list were manually identified and then entered into the Google Ngram Viewer to analyze their performance in the general English corpus using the Google Ngram Viewer (see, Figure 12-15).

Just like the performance of the concept itself, the broader view of religion also shows a downtrend of religion that is somehow moderated as of the second half of the 20th century (see Figure 12). The results are in line with the previous multi-language analysis. Contrasting the
The newly identified political references display a stable trend (see Figure 13). The most important political concept, power, is losing importance between 1800-2000, whereas the term political is becoming increasingly important in the 20th century. The overall results mean a significant relativization of the outcome of the previous analysis: The upswing of the concept of politicization is not attended by signs of an increasing importance of the political system, which nonetheless appears to be the most dominant function system given the very high word frequencies of the political keywords (ranging from near 100ppm to 570ppm).
Except for the term economic, which starts to rise at about the same time as the German word for economy, the economic frequencies seem to be stagnating or regressing since a sometimes rise in the first half of the 20th century (see Figure 14). Obviously, not starting until the early 1980s, the larger-scale discovery of an economization or commodification of the society might, therefore, have been late, wrong, or both.
As for the mass media system, many keywords feature an upward trend particularly in the second half of the 20th century (see Figure 15). Only “pp.”, the abbreviation for pages, displays a downtrend after 1980, thus maybe indicating a changing referencing culture. Apart from this, the findings reinforce the earlier results and add to a greater picture of the evolution of the media system. As a side note, we also find that the lower section of Figure 15 clearly corresponds to the change of the dominant meaning of the concept mediatization, which originally referred to a political reorganization of the Holy Roman Empire and now is used to speculate about socio-cultural change induced by an increasing influence of the mass media system.
Looking at the word frequency shares of each ngram-bundle, again, we find that the economy can hardly be considered more relevant than secularized religion or even the emerging mass media system. The still-dominant political terms achieve twice the scores of the economic terms.

**DISCUSSION**

The first striking finding of the presented analyses is that the Google Ngram Views of functional differentiation actually react to history, which is true with regard to both the interregional and the regional level. On the one hand, all language areas display a decline of religion as well as clear evidence of the impact of Black Friday on the takeoff of the political system and the economy. On the other hand, the results also reflect particularities related to the histories of the individual languages areas: The German time-series plot (see Figure 10) displays a veritable program change for an entire language area from a realm of priests, poets, and philosophers to a national
economy. This politico-economic campaign started soon after the death of Karl Marx, which is considerably earlier than in the other two cases. In the English and French language area, it was not until Black Friday that the economy became at least moderately relevant. In the English case, it seems to be the two World Wars, the Cold War, and de-colonization that led to an uptrend of the political system, which again stopped and moderately declined around GATT Uruguay and Perestroika. In the French case, the plot even seems to tell the story of the beginning and end of the Fourth Republic as a struggle of the political and the legal systems.

The results, therefore, not only support the hypothesis that modern societies feature an uneven distribution of function system references (H1), but also show that these distributions themselves are subject to changes in time and across cultural borders. The hypotheses H1.1 and H1.2 are, hence, supported by the results.

Both the word frequency plots of the names of the function systems (see Figure 8-10) and the broadband trending of the ten most frequent religious words in the Google Books corpus (see Figure 11) support the secularization hypothesis (H2.1). The dethroning of both the denomination and the most frequent religious terms is evident throughout and across the entire sample(s). Referring to both a state and a trend, the secularization hypothesis can, as a result, be corroborated both in the weak and the strong variant. While the first method also supports the idea of a politicization of the three societies, the politicization hypotheses (H2.2) must be relativized in view of the results of the second. In this sense, the hypotheses can only be defended because of the dominant position the political system (still) takes up in terms of word frequency shares in all of the three societies. In this sense, politicization is a state description rather than a trend statement. As a result, H2.2 is corroborated in its weak variant.

Despite the remarkable growth figures the economic word frequency shares feature in all of the three corpora, the results of both methods suggest rejecting the economization hypothesis (H2.3) in both the strong and the weak variant. In fact, in two cases, the period of the potential trend was rather short (E: 1930-1990; F: 1925-1975). The trends are stopped in all of the three language areas, in none of which the economy ever reached a dominant position throughout the entire 200 years. The only economic term that displays a potentially ongoing uptrend is the term business (see Figure 14). Even this statement, however, has to be relativized because it is true only in the English case.

The mediatization of society (H2.4), in return, however, seems to be an actual trend. Starting in the 1940s (English), 1960s (German), and 1980s (French), the terms media, Medien, and médias feature a modest, but constant uptrend. On a larger scale, the results are also consistent with the analysis of the function system denominations. While the media system never reaches a dominant position in either of the language areas, the trend remains uninterrupted throughout the entire sample period. Even books are still popular in books. The mediatization hypothesis is, as a result, weakly corroborated by the present Google Ngram views of functional differentiation.

LIMITATIONS AND FUTURE RESEARCH QUESTIONS

One of the most serious limitations of the first version Google Ngram Viewer was that it did not allow for case-sensitive queries in ngrams. Ngrams also could not be bundled into one single graph, just as the queries could not be for lemmas. Today, the enhanced version of the Google Ngram Viewer can be used to bundle individual ngrams into one graph, in which case, however, the also updated option of case-sensitive research does not apply. Furthermore, the enthusiasm for analyzing combined ngram graphs is limited by the fact that a query may apparently not
comprise more than round about 30 ngrams. Still, even in this limited form, the option to compare is attractive (see Figure 16):

![Google Books Ngram Viewer](image)

**Figure 16. Word Frequency Shares of Selected Ngram Bundles in the English Language Google Books Corpus 2012 (source: own Google Ngram enquiry)**

In order to create Figure 16, we entered the following string into the Google Ngram form:

\[(God+church+St.+Church+Christ),(power+political+government+States+war),(economic+business+money+company+cost),(information+Press+book+pp.+Journal)\]

In looking at the performances of the combined keywords relevant to our hypotheses, we find that today’s dominant position of the political system is a side-effect of secularization rather than an indicator of an increasing importance of political issues. Furthermore, we see that the economy is far from being a dominant factor. Even within this reduced sample of four out of ten function systems, it is third to the mass media system. This finding contradicts intuitions according to which our lives are dominated by the economic principle(s).

Figure 15 is a fine indicator that further combined ngram queries may well lead to extremely insightful results. Still, however, future research might also be interested in using or developing interfaces that allow for a visualization of trends of keyword bundles comprising more than 30 ngrams. In fact, such tools would be needed to create a compelling chart comparing the performances of robust bundles of 5-10 keywords of all of the 10 function systems of society.

Further critique may come into play with the fact that studying the key media of the Gutenberg Galaxy might not be the key to the emerging Internet society. However, there is evidence that online and offline contents do not differ all that much (Stern, 2004). At the same time, the advantages of a books corpus are evident: Book content can be considered subject to stricter selection and, therefore, a better indicator for importance. Moreover, the Internet corpus does allow for research within the sample period of several hundred years. Further research limits are, as a matter of course, related to restricted language competences and the corpus data size. In fact, it would be best to compare all language areas and analyze 2- or more-grams, e.g.
with regard to word co-occurrences, in order to get much deeper context-information. Future research on “Big Data” (Boyd & Crawford, 2012) in general and the present topic in particular could, therefore, call upon international cooperation and access to more powerful computer resources.

Future expeditions in the corpus might then not only open up further and more snapshots of trends in functional differentiation, but also allow for an answer to the question of whether or not functional differentiation is indeed a master trend in present societies. In this sense, an extended Google Ngram View of modern societies could inform on the actual state and trend of modernity itself.

**CONCLUSION**

The English, French, and German societies actually display politicization, at least in terms of a prevailing, though not trendy, high weight of political communication. A recently re-enforced mediatization can also be observed in all the three cases, albeit as a trend rather than a state description. Secularization seems to characterize both a state (as religion, actually, is not dominant anymore) and a trend, which, however, might have stopped during the last years of the sample period (George, 2005).

The biggest surprise resulting from the presented research certainly is that an analysis of the largest available text corpus does not corroborate the perhaps most prominent state description and trend prediction related to functional differentiation: The data does not point to an economized or capitalist society. Starting as a trend in line with the rising popularity of the early socialist movement, the “economization of society” never resulted in a high weight of the economy and stopped rather early (German: 1950ies, French: 1970ies, English: 1990ies).

If it is true that economists “perform” economies, then the presented results also support the idea that economization critics perform economization and that both parties jointly perform an economy bias in social sciences that is incongruent with the average to marginal relevance of the research object. This issue is particularly critical in the context of foresight and futures studies, where it is important to avoid third order risks of giving the right answers to the wrong questions (Roth and Kaivo-oja, in press). Our findings therefore suggest thinking twice before we continue considering economic preeminent issues of present and future societies. In this respect, it is also congruent that the different dictions and notations of economization or commodification are as marginal entries in the English language corpus as misspellings or exotic forenames are. This fact is emphasized not to imply an only marginal relevance of research in economic risks and benefits, but rather as suggestion to consider re-focusing research foci and drawing increased attention to function systems beyond the politico-economic double stars of social science. Maybe even the solution to the present “economic” “crises” is not in more, but rather in less attention to the economy (Roth, 2015b; Roth, in press). In any case, further critiques of economization or economic colonialism necessarily contribute to economization and are, therefore, (drivers of) the problem they try to solve. If high weights of particular function systems are indeed a problem, then vigilance is needed with regard to the political system rather than to the economy. Even more light, however, could be shed on those areas over which the strong interest in (the interplay of) the political and the economic system has casted large clouds throughout the last decades. Apart from a certainly necessary re-cultivation of neglected landscapes of functional differentiation, there is hardly any reason why researchers should be biased to particular function
systems and there is even less reason for resentments towards a mediocre function system or for complicity with the most dominant, respectively.

In this sense, the present article gives us a first indication of the transformative power of computer communication. If big data analysis actually corroborates the finding that the idea of an economization of society is an artifact rather than a fact, then this will indeed change the face of modern society.

REFERENCES


**Endnotes**

1 The fade-out of the fact that the proliferation of a scientific discipline represents a process of scientification rather than economization can also be observed in Anders Blok’s (2011) reply to Michel Callon’s (Callon, 2007; Callon, 1998; Çalışkan & Callon, 2009, 2010) performativity of economics program. Blok is right in pointing out that markets are not only performed by economists, but also by politicians, however, he also takes economists for representatives of the economy and hopes that an ethically guidance of both performers and counter-performers will “generate not only civilized markets, but also civilized politics” (Blok, 2011:271), in the end.

2 The present interpretation of word importance does not refer to where word importance is inversely related to word frequency (Baeza-Yates & Ribeiro-Neto, 1999). Such an indexical approach to word importance would only make sense if the present article was interested in comparing the discriminatory abilities of the concepts involved. All function systems, however, are on the same level of analysis and, therefore, feature the same degree of discriminatory power.

3 The data from 1500-1800 is likely to feature biases due to insufficient sample sizes: "The oldest works were published in the 1500s. The early decades are represented by only a few books per year, comprising several hundred thousand words. By 1800, the corpus grows to 98 million words per year; by 1900, 1.8 billion; and by 2000, 11 billion” (Michel et al., 2011:176).

4 Fold change of the relative word frequency calculated in terms of the ratio of the word frequency of 2000 and 1800: \( \tilde{r} = \frac{e(2000)}{e(1800)} \). Due to newly added features, the word frequencies did not have to be visually assessed anymore; rather, we could rely on precise measures as indicated by the enhanced Google Ngram Viewer.

5 The abbreviation ppm represents “parts per million”, i.e. the relative word frequency per million words in the Google Book corpus. The unit *per million* is used to avoid longer chains of digits after the decimal point. The most common English words account for estimated 45’000ppm in version 1 and measured 46’400ppm in version 2 of the data (4.5%/4.6%: *the*), 30’000ppm/29’400ppm (3.9%/2.9%: *of*) and 23’000ppm/22’800 (2.3%/2.3%: *and*). *Time*, the most frequent noun in the English-language Google Book corpus, has an appearance of 1’160ppm. *Political* (320ppm) hence appears about four times less than *time*.

6 In the original version of this article (Roth, 2014), this analysis was based on the English Google Million corpus (only, as other corpora proved unmanageable with the available hardware). First, the 10 separate fractions of this user-friendlier corpus were merged to a 8GB SPSS file. The corpus was then transformed into a ranked word frequency list of books published between 1800 and 2000. The 2,000 most frequent entries of this list were, thereupon, qualitatively analyzed for their function system references. As a first
result, some 230 words with a clear function system reference were identified. The only function system missing among the 2,000 most important words of the corpus was sport, with the designation itself as the first reference ranking far beyond the 6,000th position.

7 The code will be made available in a work-in-progress article entitled “Futures of a shared memory. A global brain wave measurement (1800-2000)”.

8 *politique* is both a noun and an adjective; the adding of adjectives like *légal/e* or *juridique*, however, has only minor effects on the big picture.

9 A search engine for the Google Books corpus designed by Marc Davies (2011-) allows for these and further options, however, for the American English corpus only.