

# The Themes of Entrepreneurship Discourse: A Data Analytics Approach

*Philip T. Roundy<sup>1</sup>, Arben Asllani<sup>2</sup>*

---

## **Abstract**

*Scholars are devoting heightened attention to the language of entrepreneurship and to its influence on the cognition, behaviors, and outcomes of entrepreneurs and their stakeholders. However, the primary themes that constitute entrepreneurs' language are unexamined. In this partially-inductive study, we identify the most common themes in entrepreneurship discourse and explore how they have changed over time. To map the themes in entrepreneurs' language, we use data analytic techniques coupled with text mining algorithms to analyze a longitudinal corpus of entrepreneurial discourse. Our findings reveal five dominant and recurring themes in entrepreneurship discourse – marketing activities, technology-oriented entrepreneurship, digital entrepreneurship, professional investment, and new venture entrepreneurship – and illustrate how these themes are evolving. By examining the key themes in the discourse of entrepreneurs and charting their transformation over time, our study makes theoretical and methodological contributions to entrepreneurship research. We identify the areas where the academic literature seems to be lagging practitioner discussions and suggest that scholars should evaluate research for how closely topics are calibrated with the main themes in the discourse of entrepreneurs. Our findings also produce practical implications for entrepreneurs by identifying the main themes receiving attention, which allows entrepreneurs to evaluate if the topics that comprise their day-to-day discourse align with the themes emphasized in the larger body of entrepreneurship discourse.*

**Keywords:** *entrepreneurship; entrepreneurial communication; discourse; text analysis; data analytics*

---

<sup>1</sup> Philip T. Roundy, Summerfield Johnston Centennial Scholar, UC Foundation Assistant Professor, Department of Marketing and Entrepreneurship, Gary W. Rollins College of Business, University of Tennessee (Chattanooga), 615 McCallie Avenue, Chattanooga, TN 37403-2598, e-mail: philip-roundy@utc.edu

<sup>2</sup> Arben Asllani, Marvin E. White Professor of Management, Department of Management, Gary W. Rollins College of Business, University of Tennessee (Chattanooga), 615 McCallie Avenue, Chattanooga, TN 37403-2598, e-mail: beni-asllani@utc.edu

*Received 8 March 2018; Revised 18 June 2018, 10 July 2018; Accepted 18 July 2018.*

## INTRODUCTION

---

Entrepreneurship scholars are embracing the “linguistic turn” in organization studies and the social sciences (Alvesson & Karreman, 2000; Hjorth & Steyaert, 2004; van Werven, Bouwmeester, & Cornelissen, 2015). Language shapes perceptions, actions, and the outcomes of entrepreneurship by influencing entrepreneurs’ cognitive processes (Cornelissen & Clarke, 2010; Kor, Mahoney, & Michael, 2007), resource acquisition strategies (Roundy, 2014), and stakeholders’ evaluations (Martens, Jennings, & Jennings, 2007; Parhankangas & Ehrlich, 2014). Entrepreneurs’ language-use manifests in the discourse constructed during the entrepreneurial process and used to describe the novel organizations, products, and initiatives that entrepreneurs create (Clarke & Cornelissen, 2014). Entrepreneurs’ language also influences the processes of attention, identity construction, legitimation, and sensemaking, which, in turn, shape entrepreneurs’ performance (Roundy, 2016). However, the themes of entrepreneurs’ language, how they appear in discourse (i.e., the contextualized language used in talk or text; Linell, 2010), and how they change over time, are not clear.

Despite the strides made by studies of entrepreneurs’ language, research has not attempted to identify the common themes in entrepreneurial discourse. Scholars generally adopt an interpretivist approach (cf. Leitch, Hill, & Harrison, 2010), which involves examining how discourse is constructed and interpreted during social interactions. The focus of this work is capturing rich representations of higher-level discourse constructs, such as narratives and stories, rather than understanding word-, phrase-, or theme-level language. Instead, research primarily emphasizes how entrepreneurs use language and the outcomes of language-use and does not devote attention to the content and structure of entrepreneurial discourse (e.g., Lounsbury & Glynn, 2001). This represents an important omission in studies of entrepreneurs’ language because without a detailed understanding of the themes of entrepreneurial discourse it is difficult to identify the topics that are at the center of entrepreneurs’ communications and attention.

To address these omissions in prior research, in this study we examine two related questions: *what are the themes that comprise entrepreneurship discourse and how have these themes changed over time?* To explore these questions, we use a partially-inductive methodology (cf. Gioia, Corley, & Hamilton, 2013), coupled with research from linguistics and entrepreneurship, to analyze the themes that are present in a corpus of entrepreneurship discourse. Specifically, we combine MapReduce programming, a Big Data methodology (cf. Asllani, 2014), with traditional statistical methods to develop a text mining algorithm that generates insights into the contextualized themes of entrepreneurship

---

discourse. We identify the most common themes in the entrepreneurship lexicon and examine the extent to which they change over time.

Our study design and findings respond to calls for research at the intersection of data analytics and entrepreneurship (e.g., George, Haas, & Pentland, 2014). A greater understanding of the themes of entrepreneurship discourse represents a contribution to entrepreneurship scholarship and has implications for entrepreneurs and policymakers because it sheds light on the topics currently receiving the most attention in entrepreneurship practice, including technology-oriented entrepreneurship, digital entrepreneurship, marketing activities, professional investment, and new venture entrepreneurship. These themes were identified inductively, rather than making *a priori* assumptions about the issues that matter to entrepreneurs. This is an important distinction because it places the focus on the major themes comprising practicing entrepreneurs' discourse (i.e., practitioner discourse or discourse-in-use) rather than the themes comprising entrepreneurship scholars' discussions (i.e., academic discourse). As our findings suggest, the themes in academic and practitioner discourse are not perfectly aligned and divergences exist.

We structure the remainder of the paper as follows. First, we provide an overview of prior studies at the intersection of entrepreneurship, language, and discourse. We devote extended attention to the substantive omissions in this research that our study aims to address. We then describe the study's research design, methods, and our findings. The paper concludes with a discussion on the implications, limitations, and future directions of our research on entrepreneurship discourse.

## LITERATURE REVIEW

---

The linguistic (or "discursive") turn in the social sciences (e.g., Harre, 2008) emphasizes the power of language to shape how reality is perceived, interpreted, and described. Social scientists' growing interest in language is motivated, in part, by the linguistic paradigm in philosophy, which laid the foundations for studying the influence of language on human cognition (Wittgenstein, 1922; cf. Lycan, 2012). Disciplines as disparate as law and criminal justice (e.g., Maynard, 1988), medicine (e.g., Greenhalgh, 1999), public health (e.g., Greene & Brinn, 2003), and agriculture (e.g., Morgan, Cole, Struttmann, & Piercy, 2002) find that language-use is not "just talk" but can influence decision making, the persuasiveness of communication, the transfer of knowledge, and how people and organizations are evaluated (e.g., Breunig & Roberts, 2017). For example, scholars studying environmental policy

decisions find that the language used to frame policies influences decision making, persuasion, and evaluation (cf. Feindt & Oels, 2005). Rydin (1999), for instance, examines the language of sustainability-focused environmental policies and, quoting Edelman (1988, p. 103), argues that environmental policy is influenced by “language games that construct alternative realities, grammars that transform the perceptible into non-obvious meanings, and language as a form of action that generates radiating chains of connotations while undermining its own assumptions and assertions.” The language contained in types of discourse, such as narratives, is so influential it has been argued that “all of our knowledge is contained in stories and the mechanisms to construct and retrieve them” (Schank & Abelson, 1995, p. 1). Because of the role of language in the construction and transmission of human culture, scholars even argue that a more accurate name for the human race is *homo narrans*, that is, “narrative humans” (Niles, 1999).

The growing attention to linguistic issues in other social science disciplines spurred organizational researchers to consider the role of language in business contexts. Language can manifest in organizations in any form that discourse can take (Chatman, 1980), including direct inter-personal interactions or written texts. Studies examine the role of language in micro-phenomena, such as employee identity construction and sensemaking, and macro-oriented phenomena, such as organizational change and legitimation (cf. Vaara, Sonenshein, & Boje, 2016). In exploring these phenomena, studies analyze the language used in texts such as annual reports (e.g., Subramanian, Insley, & Blackwell, 1993), shareholder letters (Jameson, 2000), earnings press releases (e.g., Henry, 2008), and corporate websites (Pollach, 2003).

### **The power of language in entrepreneurship**

Entrepreneurship is the creation and pursuit of innovative opportunities to produce value for society (cf. Gartner, 1990; Shane & Venkataraman, 2000). Scholars focus on entrepreneur- and venture-level characteristics, such as alertness to new opportunities and bricolage activities (Roundy, Harrison, Khavul, Pérez-Nordtvedt, & 2017; Zollo, Rialti, Ciappei, & Boccardi, 2018) and, recently, on the system-level forces that support and promote regional entrepreneurial activities (Golejewska, 2018; Nicotra, Romano, Del Giudice, & Schillaci, 2018). Across these levels of analysis, scholars are devoting growing attention to how entrepreneurs construct, convey, and interpret their actions through language because of its central role in the entrepreneurship process (e.g., Clarke & Cornelissen, 2014; Roundy, 2016). These studies find that entrepreneurs’ language-use can impact identifying and constructing opportunities (Gartner, Carter, & Hills, 2003), developing business models (London, Pogue, & Spinuzzi, 2015), persuading stakeholders

to provide support (Spinuzzi, 2017), developing pitches, and pursuing investment (Parhankangas & Renko, 2017; Spinuzzi et al., 2015).

However, most entrepreneurship research examining discourse does not examine the specific words and themes that constitute the language of entrepreneurs. For example, Nicholson and Anderson (2005) analyze the role of discourse in sensemaking and sensegiving about entrepreneurship. They examine how the language about entrepreneurship contained in myths and metaphors presented in a British newspaper influences the image of entrepreneurship portrayed to readers. Similarly, Steyaert (2007, p. 463) argues that the social construction of entrepreneurship is conceptualized through “a myriad of linguistic forms and processes,” including discourse (Perren & Jennings, 2005), dramatization (Downing, 2005), metaphors (Dodd, 2002), and storytelling (Pitt, 1998). Roundy (2014) examines how the narratives constructed by social entrepreneurs influences their ability to secure professional investment. Although these studies increase understanding about how entrepreneurs use language to construct discourse and communicate, they do not examine specific word- or theme-level patterns. These studies also do not base their findings on a large corpus of text; instead, they focus on the discourse of small samples of entrepreneurs and ventures, rather than examining a broad sample of discourse across sectors.

A study by Parkinson and Howorth (2008) is an exception. They interview social entrepreneurs and then use corpus linguistics software and critical discourse analysis to identify common linguistic themes such as “local issues,” “collective action,” “geographical community,” and “local power struggles.” Moss, Renko, Block, and Meyskens (in press) and Parhankangas and Renko (2017) also examine word-level linguistic characteristics in their analyses of how entrepreneurs communicate about their ventures on crowdfunding platforms. They find that entrepreneurs’ linguistic styles impact audiences’ resource allocation decisions.

These studies and others (e.g., Lounsbury & Glynn, 2001; Martens et al., 2007) improve our understanding of the role of language and discourse in entrepreneurial activities. However, important issues remain unaddressed. First, as described, scholars examining entrepreneurial discourse primarily adopt interpretivist and social constructivist perspectives (Fenton & Langley, 2011) that are based on ethnographic and qualitative methods. Interviews are often used to capture language. However, as Achtenhagen and Welter (2007) argue, “the use of language in entrepreneurship research has potential far beyond the use of interviews” (193). Entrepreneurship researchers generally do not use quantitative methods focused on measuring and mapping the precise composition of language. Studies are also not based on a large corpus of text, in part, because analyzing such data is challenging using hand-coding

methods, which is the primary methodology in prior work. Scholars also tend to examine entrepreneurs' language in specific, localized settings (e.g., a specific organization or city); however, the national (and international) discourse about entrepreneurship has not been examined. These represent important omissions in prior research because the primary themes of entrepreneurship, and the topics receiving attention by entrepreneurs, are not clear without analyzing the precise content of entrepreneurial language and without examining the meta-discourse about entrepreneurship. The study described in the next section seeks to address these omissions in entrepreneurship research.

## RESEARCH METHODS

---

To answer our guiding research questions (i.e., what are the most prominent themes in entrepreneurship discourse and how have these themes evolved over time), we used a Big Data programming approach (MapReduce) and text mining software to analyze a large corpus of web content. Big Data is defined as data with the following characteristics: high volume, velocity, and variety (Katal, Wazid, and Goudar, 2013). Big Data is generated by sources such as social networks, web server logs, web page content, banking transactions, and financial markets. A unique set of processing and storage techniques are used to handle the challenges of collecting and analyzing Big Data (Asllani, 2014; White, 2012). Linguistic data can be analyzed with text mining methodologies, described in detail in the next section, which are used to process large amounts of text and to identify non-obvious patterns in a corpus (i.e., a collection of text; Feldman & Sanger 2007). Text mining reveals patterns and quantifies emerging keywords and phrases, which provide insight into a corpus's linguistic structure and themes (Baker et al. 2008; Morley & Bayley, 2009).

Due to the complexity and size of our dataset, we created a modified version of a traditional word-count algorithm (Dean & Ghemawat, 2008). Using a word-count algorithm with a large corpus can be challenging because it requires significant time to process the text in the corpus. We modified a MapReduce algorithm (described in detail in the next section) to run in a distributed file system (a Hadoop cluster with four nodes) and to perform the embarrassingly parallel computations in reduced time. "Embarrassingly parallel computing" is a programming concept used to describe computation problems that can be divided into a large number of parallel tasks with little effort (Herlihy & Shavit, 2012). Our word-count algorithm is a typical parallel computing task, which is used to make data analysis more manageable.

---

## Research design

The lack of prior theoretical work on the themes of entrepreneurial discourse suggests the appropriateness of exploratory, partially-inductive research design. Inductive research is appropriate when it is not clear *a priori* what specific constructs (or, in our study, words and themes) should be measured. Inductive studies generate data-driven theoretical and empirical insights rather than testing *a priori* theoretical frameworks. With a purely inductive design, the researchers design a study with limited (or even no) preconceptions about how a phenomenon works and allow the data to guide what questions are asked and, ultimately, what theories are informed.

Since we use guiding research questions about the themes of entrepreneurial discourse to focus our analysis, our study is appropriately described as *partially*-inductive (cf. Gioia, Corley, & Hamilton, 2013). A benefit of this approach is that it limits the influence of the preconceived notions and assumptions of the researchers about what themes are important—or should be important—in entrepreneurship. Minimizing the influence of such assumptions is critical because one of the main aims of the study is to understand if the themes of practitioner discourse align with, diverge from, or challenge the main topics examined by entrepreneurship scholars. If instead, we tested for themes identified from the entrepreneurship literature *a priori*, we would be unlikely to uncover themes that are unique to practitioner discourse.

In addition to the distinction between deductive and inductive approaches, there are also important differences between qualitative and quantitative methods for text analysis (cf. Berelson, 1952; Roberts, 2000). A text can be analyzed using qualitative methods that rely on researchers hand-coding texts for themes and subthemes (cf. Bowen, 2009). The advantage of this approach is that the researcher is directly analyzing the data, rather than using a computer-automated text analysis (CATA) program, which allows for rich and nuanced analysis of the data (Graebner, Martin, & Roundy, 2012). The chief downside of the qualitative approach, and the primary reason we adopted quantitative methods, is that hand-coding is a time-intensive process best-suited to relatively small datasets and corpora of text (Laver, Benoit, & Garry, 2003; Monaghan, Chater, & Christiansen, 2005). As described below, our dataset and research design produced a large corpus comprised of several million words and over three thousand web pages. It would have been very cumbersome to hand-code such a large dataset. Another advantage of quantitative text analysis approaches is that they are “hands-off” in that they rely on algorithms, not subjective perceptions, to identify common words and themes.

## Data collection

Our data source was the 2016 “Forbes Best 100 Websites for Entrepreneurs.” The “Forbes Best...” is a list of website selected annually (since 2013) by Forbes writers. The websites are selected for their:

“ability to address a range of topics of interest to entrepreneurs. Frequent posts and content quality helps get a nod. The list is a combination of practical tools – sites to crowdfund like Rock The Post or AngelList, or sites with educational resources, like Stanford’s eCorner – and inspirational advice from bloggers like Seth Godin and Steve Blank.” (Forbes, 2013).

We chose the “Forbes Best...” list, rather than compiling our own list of websites, to limit idiosyncratic researcher (and academic) bias and because the Forbes list seemed to represent a broad range of entrepreneurial discourse (e.g., discourse about starting a venture, acquiring funding, selling, and scaling). Also, Forbes relied on nominations from the entrepreneurship community to compile the list, asking for websites “that can address a wide range of topics, like how to start up, establish your brand, build a bang-up team and secure that seemingly elusive round of capital” (Forbes, 2015). The fact that Forbes “crowdsourced” at least some of the list suggests that the list contains websites that are, in fact, important to entrepreneurs. Although there are other lists of “top entrepreneurship sites” (e.g., Entrepreneur.com’s “8 successful online entrepreneurs you should be following”), the Forbes list was the most wide-reaching and comprehensive we could find.

In selecting the “Forbes Best ...” list, we analyzed sites to ensure that they represented forums for entrepreneurial discourse. We ensured that entrepreneurship was the primary focus of the sites, rather than a niche interest. We also examined each site at different points in its history to ensure that the focus of the domain name had not changed. One of the reasons we ultimately selected the Forbes list is because most of the sites were structured as blogs (i.e., rather than reproducing a story from another source, each posting had an identifiable author with a point of view) and readers could comment on each posting, which allowed for two-sided, interactive communication (a dialogue).

We constructed a corpus of text by sampling discourse from each of the websites at two different dates, per year, for a 16-year period (2001-2016). Using the Internet Archive ([www.archive.org](http://www.archive.org)) and its “Wayback Machine” feature, for each website two “snapshots” of the discourse content were captured from each year. A list of the uniform resource locators (URLs) for each site and each snapshot was generated. We then downloaded the web content into a Hadoop Distributed File System (HDFS) containing the text

from each site. The content of the websites was downloaded using the *wget* utility, defined as:

```
$ wget -l 2 -i url_list
```

where:

- **\$** is the prompt in the Linux environment terminal;
- **wget** is a freely-available utility for downloading files from the web that supports HTTP, HTTPS, and FTP protocols (i.e., the protocols that allow data communication on the web), and retrieval through HTTP proxies. **wget** is non-interactive, meaning that it can operate in the background of other operations. The command creates local versions of remote websites which are submitted to the HDFS for further processing;
- **-l 2** indicates level 2 inclusion in the download process. Level 1 of a URL represents the main page of the website and is normally named *index.html*. Level 2 represents the webpages that are linked to the main page;
- **-i** indicates the input, which can be found in the file named *url\_list*;
- **url\_list** is a text file containing the list of web page addresses from which the content should be downloaded.

We then created a MapReduce program to read the text between <body> and </body> tags in the index file of the website. Table 1 provides a summary of our data collection methodology.

Overall, we downloaded 3,434 webpages spanning 2001 to 2016 and used this data for the text mining methodology. On average, 215 unique webpages (from the Fortune 100 Best websites) were downloaded each year. The number of webpages is not equivalent to the number of websites because, as described, we analyzed data two levels deep (i.e., the main page for each website and the pages linked to the main page). That is, for a year in which all of the Fortune 100 websites are available at least 200 webpages were analyzed (the 100 websites at two points during the year). Finally, the number of webpages analyzed per year increased over time (as more webpages became available in recent years); however, we normalized our findings by year totals. These methods generated a corpus of entrepreneurial discourse of over 3 million words (3.55 gigabytes of raw text).

## Data analysis

After constructing the corpus of entrepreneurship discourse, our analysis consisted of two parts: (1) identifying the major themes and (2) charting the trends of themes over time.

**Table 1.** Summary of data collection and analysis steps

Methodological step	Description
<b>Data collection</b>	
Identified the data source	“Forbes Best 100 Websites for Entrepreneurs”
Created text corpus	Used the Internet Archive to find the URLs of each website at two points per year from 2001-2016 Used the wget utility (Linux command) to capture and download the text of the websites of the selected URLs two-levels deep Created a corpus of 3,434 files (approximately 3 million words)
Stored and organized data	Stored the downloaded text in a Hadoop Distributed File System (HDFS) with four clusters
<b>Data analysis</b>	
Cleaned the corpus	Used a modified MapReduce program to eliminate common words (“stop words”), HTML tags, and other symbols
Identified the most common words	Used a modified MapReduce program to identify the most common words and phrases
Identified the most common themes	Used exploratory factor analysis to identify themes in the most common words in the corpus.
Examined changes in the themes over time	Calculated the average frequency index for each theme during a given year

We began by modifying a MapReduce algorithm (Dean & Ghemawat, 2008) to count the frequency of each word in the corpus. The program also eliminated common words (e.g., “the,” “and”), HTML tags, and other symbols. Figure 1 contains pseudo code for the MapReduce program. The MapReduce algorithm was executed in a Hadoop cluster with four nodes. The most frequently used words for each year were selected and processed to eliminate duplicates. We also created obvious groupings (e.g., combining words like *knowledge* and *information* into *information*) and identified words sharing the same stem (e.g., *finance*, *financial*, and *financing*). Table 2 contains the full list of 126 words used in the factor analysis described below.

### Identification of themes

We used exploratory factor analysis (EFA; Fabrigar & Wegener, 2011) to identify themes in the most commonly occurring words in the corpus. Table 2 shows the overall model parameters for the EFA.

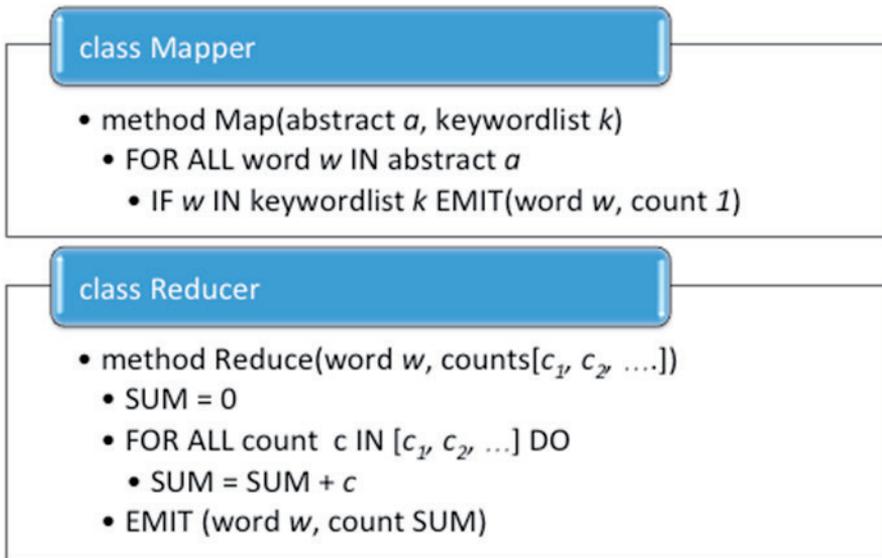
**Table 2.** The words of entrepreneurship discourse

acquisition	development	LLC	small business labs
advertising	digital marketing	market	small business
analytics	downline	marketing	small business administration
angel	due diligence	merger	sociable
angel investors	edge	mobile	social
application	entrepreneur	money	social enterprise
appraisal	entrepreneurial ecosystem	movable type	social entrepreneur
asset	Facebook	multi-level marketing	social good
barter	family	network marketing	social innovation
benefits	fast company	networking	social media
big data	feed	new venture	social network
bootstrap	financing	offices	software
business	focus	online	sole proprietorship
business advice	funded	opportunities	Stanford
business blogger	funds	option	startup
business filings	game	outsourcing	startup community
business incubator	general partnership	partnership	startup lawyer
business valuation	home based business	patent	startup lessons learned
capital	idea	people	startups
coaching	independent contractor	planning	stock
company	innerpreneur	player	story
computer	innovation district	product	strategic alliance
consumer direct marketing	internet	public relations	summary
copyright	intrapreneur	resources	team
corporation	investors	sales	tech
creator	joint venture	SBA	tech crunch
customer	limited liability company	services	technology
data	limited partnership	share	venture
data analytic	trademark	shift	venture blog
technorati	Twitter	valley	venture capital
trade	success	VC	
women	line of credit	small	

Once we identified the most frequent keywords, we calculated the frequency index  $f_{ij}$  of each key word  $i$  in webpage  $j$  as follows:

$$f_{ij} = \frac{F_{ij}}{T_j} \quad (1)$$

where  $F_{ij}$  is the frequency of keyword  $i$  in  $j$  and  $T_j$  is the total number of words in webpage  $j$ . To calculate  $F_{ij}$  and  $T_j$  we ran the MapReduce algorithm for each full webpage, with the keyword list as an input to the program.



**Figure 1.** Modified MapReduce program used to identify frequent words

The Kaiser-Meyer-Olkin (KMO) value of 0.70 indicates that our data is suitable for factor analysis (Cerny & Kaiser, 1977). Bartlett's test of sphericity tests the hypothesis that the variables are unrelated and, thus, unsuitable for structure detection and factor analysis. A low significance value (<0.001) indicates that factor analysis is, in fact, useful with our data (Snedecor & Cochran, 1989).

Table 4 contains the factor correlation matrix. Five independent factors – themes – of entrepreneurship discourse were identified. Table 5 contains the strongest-loading words on each of the five themes. In the factor analysis, words with loadings of .30 and greater were retained (following the recommendation of Brown, 2006).

**Table 3.** Model validity for factor analysis

Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy		.702
Bartlett's Test of Sphericity	Approx. Chi-Square	40282.785
	df	903
	Sig.	.000

**Table 4.** Factor correlation matrix

Factor	1	2	3	4	5
1	1.000	-.146	.009	.026	.205
2	-.146	1.000	.327	-.145	-.168
3	.009	-.327	1.000	.290	.142
4	.026	-.145	.290	1.000	-.070
5	.205	-.168	.142	-.070	1.000

**Note:** Extraction Method: Principal Axis Factoring; Rotation Method: Promax with Kaiser normalization.

**Table 5.** Exploratory factor analysis

Factor	Professional investment	Technology-oriented entrepreneurs hip	Digital entrepreneurship	New venture entrepreneurship	Marketing activities
venture capital funds	0.849				
venture capital vc	0.814				
	0.776				
venture capital	0.587				
vc	0.575				
technology		0.906			
software		0.843			
services		0.635			
shift		0.56			
Twitter		-0.416			
share			0.670		
social			0.601		
Facebook			0.464		
team			0.401		
mobile			0.355		
people			0.349		
startups				0.829	

Factor	Professional investment	Technology-oriented entrepreneurs hip	Digital entrepreneurship	New venture entrepreneurship	Marketing activities
angel				0.800	
startup				0.777	
small_business					0.534
data					-0.468
online					0.375
marketing					0.373
sba					0.338
sales					0.328

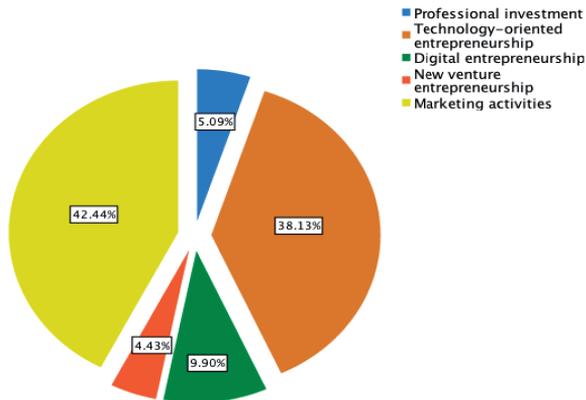
**Note:** Extraction Method: Principal Axis Factoring; Rotation Method: Promax with Kaiser Normalization; Rotation converged in 7 iterations.

## FINDINGS

The study aimed to identify the key themes in entrepreneurship discourse and to examine if these themes changed over time. In the following sections, we describe the five most common themes and their main characteristics.

**Marketing activities.** The most commonly occurring theme in entrepreneurship discourse, appearing in over 42% of websites included in the corpus (Figure 2), is comprised of keywords such as *marketing*, *sales*, and (customer) *data*. Given the focus of the words that loaded on this factor, we labeled this theme *marketing activities*.

Many of the foundational writings about entrepreneurship are from the work of economists (e.g., Cantillon, 1730; Knight, 1921; Say, 1816; Schumpeter, 1934). As entrepreneurship developed into an established academic field, management became its “home” discipline (Shane & Venkataraman, 2000). However, there is a growing stream of research at the intersection of marketing and entrepreneurship (cf. Hills & LaForge, 1992; Hills & Hultman, 2011). This work takes a “demand-side” perspective that emphasizes how entrepreneurs’ market their ventures to consumers (e.g., Priem, Li, & Carr, 2012), rather than a “supply-side” perspective focusing on the characteristics of entrepreneurs (Kaish & Gilad, 1991).



**Figure 2.** The representation of themes in entrepreneurship discourse

It is notable that the discourse of actual entrepreneurs reflects the increasing academic emphasis on entrepreneurs' marketing practices. This theme indicates that while it is important for entrepreneurs to create cutting-edge products and technologies, entrepreneurs are increasingly doing so by adopting a customer-centric mindset and using strategies (like design thinking; Elsbach & Stigliani, 2018) to understand consumers and gather customer data.

**Technology-based entrepreneurship.** The second most common theme in the corpus of discourse revolved around a cluster of words and phrases involving *technology-based entrepreneurship*. This theme appeared in over 38% of websites. The highest factor loadings in this category included words such as *technology, software, services* (as in "cloud-based services" and "software as a service"), and *technology shift*.

In the period studied (2001-2016), there is a growing focus in research and practice on technology entrepreneurship (Ratinho, Harms, & Walsh, 2015; Shane & Venkataraman, 2003). Technology entrepreneurship is at the intersection of two phenomena: technological innovation and entrepreneurship (Mosey, Guerrero, & Greenman, 2017). It involves the pursuit of an opportunity that "assembles and deploys specialized individuals and heterogeneous assets that *are intricately related to advances in scientific and technological knowledge* for the purpose of creating and capturing value for a firm" (Bailette, 2012: 9; emphasis added).

Individuals engaged in technology entrepreneurship assemble “resources and structures to exploit emerging technology opportunities” (Liu et al., 2005). Scholars acknowledge that technology entrepreneurship is not only a source of product innovation and technological advancement but serves as a potent mechanism for generating economic development (Bailetti, 2012). Findings suggest that technology entrepreneurship is also now a central theme in practitioner entrepreneurship discourse.

**Digital entrepreneurship.** A distinct theme also emerged around *digital entrepreneurship*, which included words such as *social (media)*, *share*, *Facebook*, and *mobile*. Digital entrepreneurship is a specific type of technology entrepreneurship focused on the pursuit of opportunities related to products and services based on digital media and other information technologies (Davidson & Vaast 2010: 2; Nambisan, 2017). This theme, which appeared in approximately 10% of websites in the corpus, includes the host of new business models being created around social media activities (cf. Hanna, Rohm, & Crittenden, 2011; Khajeheian, 2013) and corresponds to the digitalization of many industry sectors (Autio, Nambisan, Thomas, & Wright, 2018).

**Professional investment.** Another theme is comprised of keywords, such as *venture*, *capital*, *funds*, and *VC*, and phrases like *venture capital*. Because of the shared focus of these words, we labeled this theme “*professional investment*.” Professional investors, such as venture capitalists, are commonly-pursued by entrepreneurs as early-stage sources of funding that can complement (and come at a later stage than) other sources of startup funding, such as family and friends, angel investors, crowdfunding, and an entrepreneur’s personal wealth (Ascher, 2012; Gompers & Lerner, 2001; Wong, Bhatia, & Freeman, 2009). The importance of early-stage professional investment in supporting the scaling of high-growth ventures makes it unsurprising that discussions about such investment are one of the primary themes of entrepreneurship discourse. In sectors in which entrepreneurs pursue exponential (“hockey stick”) growth, such as internet technology, early-stage professional investment often represents a key source of funding that gives entrepreneurs access to the funds they need to develop their products, engage in R&D, hire a sales force, and create a marketing campaign (e.g., Davila, Foster, & Gupta, 2003). As Figure 2 illustrates, the venture capital theme was present in approximately 5% of discourse in the corpus. This percentage may reflect that, while professional investment is an important topic amongst some types of entrepreneurs, only a small

percentage of entrepreneurs are creating the types of fast-scaling ventures that need or can generate the type of returns that appeal to such investors. ***New venture entrepreneurship***. A final theme was comprised of words, like “startup,” which are a direct reference to new businesses and the creation of new organizations. Words associated with this theme were only present in less than 5% of the discourse, which might seem surprising given it is a corpus of entrepreneurship discourse; but there are at least two explanations for the theme’s low frequency relative to other common themes. First, words that are directly related to the creation of new organizations, such as “new venture,” might not need to be explicitly stated because the discourse was collected from entrepreneurship websites. In other words, there may be an implicit understanding that conversations are about activities involved in the creation of new firms and, thus, it is not necessary to overly use words like “startup” or “new venture” (e.g., articles about marketing challenges in new ventures, might simply refer to “marketing challenges” because the understanding is that the focus is new firms).

More subtly, the low prevalence of the *new venture entrepreneurship* theme, relative to the other themes, may reflect the fact that entrepreneurship is increasingly not confined to the creation of new organizations (Morris & Jones, 1999). Rather, contemporary definitions of entrepreneurship (and “entrepreneurial”) emphasize that entrepreneurship is the creation of innovative organizations, products, or initiatives that create value (Nasution et al., 2011; Roundy, Bradshaw, & Brockman, 2018). Pursuing opportunities for innovations that produce value can be done outside the startup context, such as in established organizations (cf. work on corporate entrepreneurship; Kuratko, Hornsby, & Covin, 2014; Zarei, 2017), or as part of causes, movements, or other types of temporary organizations that do not require the establishment of formal (fully-incorporated) ventures (Burke & Morley, 2016). Entrepreneurship discourse reflects these broader views of entrepreneurial phenomena.

### **The evolution of themes in entrepreneurial discourse**

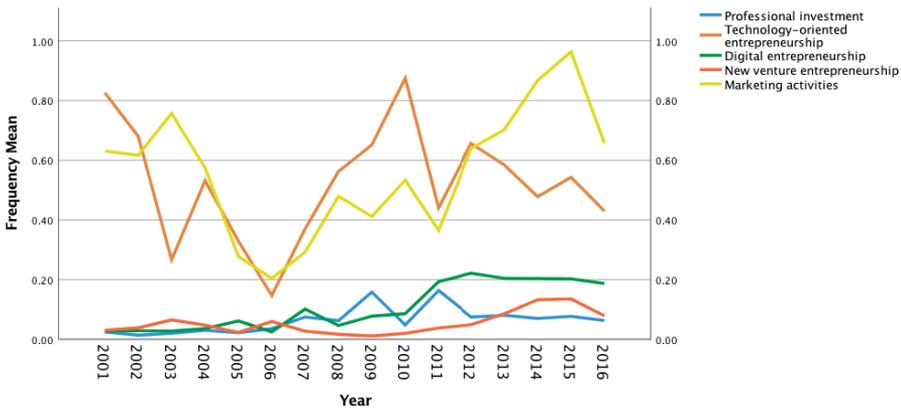
To examine how the themes identified in the previous section changed over time, we calculated the average frequency index for each theme during a given year, as:

$$F_{yt} = \frac{\sum_{i=1}^{|T_t|} f_{iy}}{|T_t|} \quad (2)$$

where:

- $t$  = the theme number ( $t = 1, 2, \dots, 5$ )
- $y$  = the year ( $y = 2001, \dots, 2016$ )
- $F_{yt}$  = the frequency of theme  $t$  in year  $y$
- $T_t$  = the set of keywords in theme  $t$ .
- $|T_t|$  = the cardinality of set  $T_t$ , that is, the number of keyword each set  $T_t$
- $f_{iy}$  = frequency of keyword  $i$  in year  $y$

Figure 3 represents the frequency of each theme during the 2001-2016 period.



**Figure 3.** The evolution of themes in entrepreneurship discourse

The figure indicates that the five themes can be further classified into two superclusters consisting of *marketing activities* and *technology-based entrepreneurship*, which during the span of the study were the most frequently-occurring themes in entrepreneurship discourse, and *digital entrepreneurship*, *professional investment*, and *new venture creation*, which were less dominant (occurring in less than 20% of the corpus) but have a continuous (albeit slightly increasing) presence during the past 16 years. One way to interpret these findings is that they indicate that marketing and technology are at the core of discourse about entrepreneurship while conversations about digital entrepreneurship, investment, and new venture activity are supplemental themes.

Several additional trends emerge when examining the themes separately. For instance, “digital entrepreneurship” steadily increased from 2001 to 2010,

presumably as the social media sector grew in prominence. From 2010-2012, there was a steep increase in digital entrepreneurship discourse, which has since leveled off. One possible explanation for the plateauing of the theme is that as social media platforms like Twitter and Facebook have become ubiquitous, the creation of business models and innovations based on digital technologies became an accepted part of entrepreneurship and, hence, a theme in entrepreneurship conversations that receives less attention. Furthermore, it is intuitive that *technology-based entrepreneurship* is a more common theme over time than *digital entrepreneurship* because the former is a more general type of entrepreneurship that includes a wider range of business models, industries, and products. Similarly, *marketing activities* is a more commonly occurring theme than professional investment because all ventures must interact with customers, but a smaller percentage pursue (and receive) professional investment. Overall, entrepreneurs' language reflects what is occurring in both the startup community and the general marketplace.

---

## DISCUSSION

The role of language in constructing and describing entrepreneurial activities is a topic receiving increased interest (cf. Clarke, Cornelissen, & Healey, in press; Spinuzzi, 2016). The theme-level content of entrepreneurship discourse is, however, not fully understood. Two overriding questions guided our study: what are the primary themes of entrepreneurship discourse? Moreover, how have these themes changed over time? Below, we summarize the answers we uncovered and examine the contributions and implications of our findings to scholars and practitioners.

### Contributions to scholarship

Despite growing attention to the discourse of entrepreneurs, we know surprisingly little about the specific themes that constitute their language. In this study, we identify the five most common themes in entrepreneurship discourse (marketing activities, technology entrepreneurship, digital entrepreneurship, professional investment, and new venture entrepreneurship) during the past 16 years. In doing so, we uncover, arguably, the most frequently discussed topics among entrepreneurs and the issues that they are giving the greatest attention. By creating a corpus from a range of national and international websites (from the Forbes Best 100 Websites for Entrepreneurs), we were able to identify the key themes in general entrepreneurship discourse, rather than focusing on the discourse tied to a specific subset of entrepreneurs, organizations, or industries. We were also able to approach the analysis

without *a priori* assumptions about what themes are most important to practicing entrepreneurs. By identifying the word- and phrase-level patterns that create distinct themes in entrepreneurship language, we make several conceptual and empirical contributions to entrepreneurship research.

First, our findings provide empirical support for intuitive trends in entrepreneurship, such as the rise of technology and digital entrepreneurship. To the extent that entrepreneurship discourse both reflects and helps to construct what is given attention (e.g., Logan, 1999), the themes we identify represent the issues that entrepreneurs devote most of their attention to discussing. Related to this point, the findings also call into question whether the concepts receiving the most attention from scholars are the main topics comprising entrepreneurship discourse. For most of the themes, there is alignment between the existence of a robust stream of academic research and a vibrant practitioner discourse (e.g., technology entrepreneurship; professional investment; new venture entrepreneurship).

However, for two themes – marketing activities (in an entrepreneurship context) and digital entrepreneurship – the academic literature seems to be lagging practitioner discussions, which suggests that more research is needed on these aspects of entrepreneurship. For instance, the stream of research that has developed at the marketing and entrepreneurship “interface” (e.g., Hills & Hultman, 2011), the creation of academic organizations focused on this topic (e.g., the Entrepreneurial Marketing special interest group (SIG) in the American Marketing Association), and the scholarly events dedicated to marketing issues in entrepreneurship (e.g., the Global Research Symposium on Marketing and Entrepreneurship), are all making in-roads in drawing attention to the importance of marketing in entrepreneurial activities. However, in many respects, this research is still considered a “niche” topic within the broader academic conversation about entrepreneurship. Our findings suggest that marketing issues are front-and-center in practitioner discourse and should occupy a more central position in academic conversations.

Furthermore, it is useful to think about what the two dominant themes in entrepreneurship discourse – technology entrepreneurship and marketing – represent. On a deeper level, the creation of new technologies is core to what entrepreneurs do and represents a primary form of “value creation” (e.g., Lepak, Smith, & Taylor, 2007). The introduction, development, and delivery of innovative technologies is central to the function that entrepreneurs serve in the marketplace. However, for entrepreneurs to be financially viable, they must also engage in “value capture” (Fayolle, 2007), which involves “the appropriation

and retention by the firm of payments made by consumers in expectation of future value from

consumption” (Priem, 2007, p. 220). Marketing activities are key to capturing value (Mizik & Jacobson, 2003). Thus, the dominant themes in entrepreneurship discourse reflect the two guiding logics – value creation and value capture – that entrepreneurs must manage.<sup>3</sup>

An interesting, although counter-intuitive, finding is the lack of evidence in practitioner discourse for some of the main themes in entrepreneurship research. Most notably, the topic of “opportunity,” and the examination of how entrepreneurs construct, discover, and develop new opportunities, is one of the most intensely researched topics in the entrepreneurship discipline (cf. Short, Ketchen, Shook, & Ireland, 2010). The word opportunity (and its variants), however, did not load on any of the five main themes we identified. There are at least two explanations for this result. First, opportunity may be a concept so pervasive in entrepreneurship, and so fundamental to the phenomenon, that entrepreneurs do not find it necessary to draw explicit attention to it. If so, then there is an unstated assumption among entrepreneurs that most conversations involve some aspect of turning an opportunity into a viable business. In contrast, “opportunity” may instead be a concept that scholars devote significant time to understanding while entrepreneurs focus on more concrete topics and practices (Gartner, Stam, Thompson, & Verduyn, 2016). Entrepreneurs may not spend time thinking and discussing concepts like opportunity because they are viewed as ethereal and not directly involved in day-to-day entrepreneurial activities. Our findings suggest that research is needed to examine the degree to which the opportunity concept plays a role in the practices of entrepreneurs.

The prevalence of the “digital entrepreneurship” theme, particularly post-2010, suggests that scholars should devote more attention to the growing digital infrastructure (Nambisan, 2017) and how it is changing entrepreneurial activities. For instance, research is needed on how entrepreneurs harness “technological affordances (Gibson, 1977) created by digital technologies and infrastructures,” and how the digitalization of the economy represents an “economy-wide redesign of value creation, delivery, and capture processes” (Autio et al., 2018: 74). At the same time, scholars should be attuned to changes in the tenor of entrepreneurial (and consumer) discourse about digitization as there may be a growing dialogue about the negatives of the digitalization of society and a developing counter-cultural movement away from digital to analog (e.g., Sax, 2016). Overall, our findings contribute to entrepreneurship research by serving as a reminder that scholars should be aware of the main themes in discourse about entrepreneurship to ensure that their research has some relevance to practitioners (cf. Vermeulen, 2007).

---

<sup>3</sup> We thank an anonymous reviewer for suggesting this line of thinking.

Our study also has methodological implications. Most research on entrepreneurship and discourse employs qualitative methods, such as interviewing and ethnographic observation, and utilizes small samples comprised of entrepreneurs from the same organization, industry, or geographic area. Our findings illustrate the use of quantitative, computer automated text analysis (CATA) and a “Big Data” approach (Asllani, 2014). Our methodology allowed us to construct a broadly-representative corpus of entrepreneurship discourse comprised of over 3 million words and over 3000 unique webpages. To the best of our knowledge, we are the first scholars to use this type of methodology in the context of entrepreneurship discourse. Our methods, which we describe in detail and can be followed by other researchers, represent an innovative approach to analyzing entrepreneurs’ language.

### **Implications for practitioners**

Research examining entrepreneurship discourse consistently finds that the language entrepreneurs use to conceptualize and describe their ventures matters. Language is not merely a reflection of cognition or behaviors; it can shape thinking and action (Lewis, 1966). For this reason, if entrepreneurs want to participate in conversations about entrepreneurship (e.g., when pitching their ventures or when gathering information from other members of their entrepreneurial ecosystem; Roundy, 2016), it is important for them to be aware of the main themes in entrepreneurship discourse so that they can tailor their language accordingly.

The content of the specific themes we identify also has implications for entrepreneurs. For example, entrepreneurs should acknowledge the important role played by marketing and what can be gained by taking a consumer perspective. Although this might seem like an obvious insight, many entrepreneurs, because of their backgrounds in non-business disciplines such as engineering and computer science, adopt a product- rather than customer-focus (Rosen, Schroeder, & Purinton, 1998). However, as evidenced by the high frequency of discussions about marketing and consumer activities, entrepreneurs are devoting an increasing amount of their discourse to marketing issues. At the same time, even though it was one of the least common of the five primary themes, discussions about professional investment still appeared in between 5% and 18% of website discourse. Given the extremely small percentage of firms that qualify for and receive professional investment (cf. Rao, 2013), this theme may actually be *over*-represented in entrepreneurs’ conversations. That is, entrepreneurs may be too concerned with discussing “how to attract venture capital” rather than pursuing other funding options such as bootstrapping or crowdfunding (e.g., Belleflamme, Lambert, & Schwienbacher, 2014). Thus, entrepreneurs could use our findings to assess

---

what they are spending their time discussing and to assess whether other topics should be the focus of their attention and discourse.

### **Limitations and directions for future research**

Despite the contributions of our research, it was not without limitations, which serve as directions for future research. First, our sample was comprised entirely of discourse from entrepreneurship websites. Although our sample produced a large corpus, it is not exhaustive of all types of entrepreneurship. Thus, while the corpus is representative of larger conversations about entrepreneurship, there may be some groups that are not part of these conversations. For example, there are some types of entrepreneurs, such as traditional small business entrepreneurs, that may be less likely than entrepreneurs who are growing rapidly-scaling ventures to take part in the discussions of the websites we examined. Furthermore, the “Forbes Best 100...” list is only a sample of global entrepreneurship discourse and has the limitation of only representing English-speaking journals. Research is needed examining the discourse of entrepreneurs outside the Western context.

In addition, as we have noted, our corpus is comprised of discourse from practitioners and does not reflect academic discourse about entrepreneurship. An important direction for future studies is formally analyzing the extent to which discourse contained in scholarship about entrepreneurship is lagging (or leading) practitioner entrepreneurship discourse. To explore this issue, researchers could create a corpus, similar to the one constructed for this study, but comprised of a collection of academic entrepreneurship articles from the same period as our study (e.g., all articles published in a particular journal or set of journals). Our text mining methodology could then be used to identify the main themes in academic entrepreneurship discourse to determine how they have changed over time and how much scholarly discourse matches or diverges from practitioner discourse.

An additional avenue for future research is to go beyond examining themes to analyze the deeper-level linguistic characteristics of entrepreneurship discourse. For example, CATA software, such as the Linguistic Inquiry and Word Count (LIWC) program, could be used to examine the social and psychological properties of entrepreneurial discourse, including its emotionality and concreteness (cf. Pennebaker et al., 2001).

## **CONCLUSION**

---

Entrepreneurship is increasingly viewed as a potent engine for unlocking economic potential and generating value. Language is involved in all facets

of entrepreneurship, including when entrepreneurs “develop an innovation, look at possible markets, conduct market research, seek intellectual property protection, develop a business model, describe a product, identify a value proposition, and pitch to stakeholders” (Spinuzzi, 2016, p. 316). Thus, it is important to understand what comprises entrepreneurial discourse. The study described in this paper represents the first steps toward mapping entrepreneurship discourse and identifying its key themes. We hope that our findings stimulate thought, debate, and ultimately future research, which produces a deeper understanding of the language of entrepreneurs.

## References

- Achtenhagen, L., & Welter, F. (2007). Media discourse in entrepreneurship research. In H. Neergaard, J.P. Ulhøi (Eds.), *Handbook of Qualitative Methods In Entrepreneurship Research* (pp. 193-215). Northampton, MA: Edward Elgar Publishing, Inc.
- Alvesson, M., & Kärreman, D. (2000). Taking the linguistic turn in organizational research: Challenges, responses, consequences. *Journal of Applied Behavioural Science*, 36(2), 136-58.
- Ascher, J. (2012). Female entrepreneurship – An appropriate response to gender discrimination. *Journal of Entrepreneurship, Management, and Innovation*, 8(4), 97-114.
- Asllani, A. (2014). *Business Analytics with Management Science Models and Methods*. Upper Saddle River, NJ: Pearson Education Incorporated.
- Autio, E., Nambisan, S., Thomas, L. D., & Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal*, 12(1), 72-95.
- Bailetti, T. (2012). Technology entrepreneurship: Overview, definition, and distinctive aspects. *Technology Innovation Management Review*, 2(2): 5-12.
- Baker, P., Gabrielatos, C., Khosravinik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse and Society*, 19(3), 273-306.
- Belleflamme, P., Lambert, T., & Schwienbacher, A. (2014). Crowdfunding: Tapping the right crowd. *Journal of Business Venturing*, 29(5), 585-609.
- Berelson, B. (1952). *Content Analysis in Communication Research*. Glencoe, IL: Free Press.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40.
- Breunig, K. J., & Roberts, H. (2017). Money talks: Communication patterns as knowledge monetization, *Journal of Entrepreneurship, Management, and Innovation*, 13(3), 71-94.
- Brown, T.A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.

- Burke, C.M., & Morley, M.J. (2016). On temporary organizations: A review, synthesis and research agenda. *Human Relations*, 69(6), 1235-1258.
- Cantillon, R. (1931). *Essai sur la nature du commerce en general [Essays on the nature of trade in general]*, H. Higgs (Ed. and translator). London, UK: Macmillan.
- Cerny, C.A., & Kaiser, H.F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioral Research*, 12(1), 43-47.
- Chatman, S.B. (1980). *Story and Discourse: Narrative Structure in Fiction and Film*. Ithaca, NY: Cornell University Press.
- Clarke, J.S., & Cornelissen, J.P. (2014). How language shapes thought: New vistas for entrepreneurship research. In J.R. Mitchell, R.K. Mitchell, B. Randolph-Seng (Eds.) *Handbook of Entrepreneurial Cognition* (pp. 383-397). Cheltenham, UK: Edward Elgar Publishing Limited.
- Clarke, J.S., Cornelissen, J.P., & Healey, M.P., Actions speak louder than words: How figurative language and gesturing in entrepreneurial pitches influences investment judgments. *Academy of Management Journal*, forthcoming.
- Cornelissen, J. P., & Clarke, J. S. (2010). Imagining and rationalizing opportunities: Inductive reasoning and the creation and justification of new ventures. *The Academy of Management Review*, 35(4), 539-557.
- Davidson, E., & Vaast, E. (2010). Digital entrepreneurship and its sociomaterial enactment. *The 43rd Hawaii International Conference on System Sciences (HICSS)*, IEEE, 1-10.
- Davila, A., Foster, G., & Gupta, M. (2003). Venture capital financing and the growth of startup firms. *Journal of Business Venturing*, 18(6), 689-708.
- Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107-113.
- Dodd, S.D. (2002). Metaphors and meaning: A grounded cultural model of US entrepreneurship. *Journal of Business Venturing*, 17(5), 519-535.
- Downing, S. (2005). The social construction of entrepreneurship: Narrative and dramatic processes in the coproduction of organizations and identities. *Entrepreneurship Theory and Practice*, 29(2), 185-204.
- Edelman, M. (1988). *Constructing the Political Spectacle*. Chicago: University of Chicago Press.
- Elsbach, K. D., & Stigliani, I. (2018). Design thinking and organizational culture: A review and framework for future research. *Journal of Management*, 44(6), 2274-2306.
- Fabrigar, L.R. and Wegener, D.T. (2011). *Exploratory Factor Analysis*. New York: Oxford University Press,
- Fayolle, A. (2007). *Entrepreneurship and New Value Creation: The Dynamic of the Entrepreneurial Process*. Cambridge, UK: Cambridge University Press.
- Feindt, P. H., & Oels, A. (2005). Does discourse matter? Discourse analysis in environmental policy making. *Journal of Environmental Policy & Planning*, 7(3), 161-173.

- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data*. New York, NY: Cambridge University Press.
- Fenton, C., & Langley, A. (2011). Strategy as practice and the narrative turn. *Organization Studies*, 32(9), 1171-1196.
- Gartner, W. B. (1990). What are we talking about when we talk about entrepreneurship?. *Journal of Business Venturing*, 5(1), 15-28.
- Gartner, W.B., Carter, N.M., & Hills, G.E. (2003). The language of opportunity. In C. Steyaert, D. Hjorth (Eds.), *New Movements in Entrepreneurship* (pp. 103-124). Cheltenham, UK: Edward Elgar.
- Gartner, W.B., Stam, E., Thompson, N., & Verduyn, K. (2016). Entrepreneurship as practice: Grounding contemporary practice theory into entrepreneurship studies. *Entrepreneurship & Regional Development*, 28(9-10): 813-816.
- George, G., Haas, M.R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321-326.
- Gibson, J.J. (1977). The theory of affordances. In R.E. Shaw & J. Bransford (Eds.), *Perceiving, Acting, and Knowing* (pp. 67-82). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gioia, D.A., Corley, K.G., & Hamilton, A.L. (2013). Seeking qualitative rigor in inductive research. *Organizational Research Methods*, 16(1), 15-31.
- Golejewska, A. (2018). Innovativeness of enterprises in Poland in the regional context. *Journal of Entrepreneurship, Management, and Innovation*, 14(1), 29-44.
- Gompers, P., & Lerner, J. (2001). The venture capital revolution. *Journal of Economic Perspectives*, 15(2), 145-168.
- Graebner, M. E., Martin, J. A., & Roundy, P. T. (2012). Qualitative data: Cooking without a recipe. *Strategic Organization*, 10(3), 276-284.
- Greene, K., & Brinn, L.S. (2003). Messages influencing college women's tanning bed use: Statistical versus narrative evidence format and a self-assessment to increase perceived susceptibility. *Journal of Health Communication*, 8(5): 443-61.
- Greenhalgh, T. (1999). Narrative based medicine in an evidence based world. *The British Medical Journal*, 318(7179), 323-325.
- Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're all connected: The power of the social media ecosystem. *Business Horizons*, 54(3), 265-273.
- Harre, R. (2008). The discursive turn in social psychology. In D. Schiffrin, D. Tannen, & H. E. Hamilton, *The Handbook of Discourse Analysis* (pp. 688-706). Massachusetts: Blackwell Publishers Ltd.
- Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication*, 45(4), 363-407.
- Herlihy, M., & Shavit, N. (2012). *The Art of Multiprocessor Programming, Revised Reprint (revised ed.)*. Elsevier.

- Hills, G. E., & LaForge, R. W. (1992). Research at the marketing interface to advance entrepreneurship theory. *Entrepreneurship Theory and Practice*, 16(3), 33-60.
- Hjorth, D., & Steyaert, C. (2004). *Narrative and Discursive Approaches in Entrepreneurship*. Cheltenham, MA: Edward Elgar.
- Jameson, D. A. (2000). Telling the investment story: A narrative analysis of shareholder reports. *Journal of Business Communication*, 37(1), 7-38.
- Kaish, S., & Gilad, B. (1991). Characteristics of opportunities search of entrepreneurs versus executives: Sources, interests, general alertness. *Journal of Business Venturing*, 6(1), 45-61.
- Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data: Issues, challenges, tools and good practices. In *Contemporary Computing (IC3), 2013 Sixth International Conference on* (pp. 404-409). IEEE.
- Khajeheian, D. (2013). New venture creation in social media platform; Towards a framework for media entrepreneurship. In M. Friedrichsen & W. Mühl-Benninghaus (Eds.) *Handbook of Social Media Management* (pp. 125-142). Springer, Berlin, Heidelberg.
- Kor, Y. Y., Mahoney, J. T., & Michael, S. C. (2007). Resources, capabilities and entrepreneurial perceptions. *Journal of Management Studies*, 44(7), 1187-1212.
- Knight, F.H. 1921[2012]. *Risk, Uncertainty and Profit*. Mineola, NY: Dover Publications.
- Kuratko, D. F., Hornsby, J. S., & Covin, J. G. (2014). Diagnosing a firm's internal environment for corporate entrepreneurship. *Business Horizons*, 57(1), 37-47.
- Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, 97(2), 311-331.
- Leitch, C.M., Hill, F.M., & Harrison, R.T. (2010). The philosophy and practice of interpretivist research in entrepreneurship: Quality, validation, and trust. *Organizational Research Methods*, 13(1), 67-84.
- Lepak, D. P., Smith, K. G., & Taylor, M. S. (2007). Value creation and value capture: A multilevel perspective. *Academy of Management Review*, 32(1), 180-194.
- Lewis, C.S. 1966. *On Stories: And Other Essays in Literature*. Orlando, FL: Harcourt, Inc.
- Liu, T.H., Chu, Y.Y., Hung, S.C. & Wu, S.Y., 2005. Technology entrepreneurial styles: A comparison of UMC and TSMC. *International Journal of Technology Management*, 29(1-2), 92-115.
- Linell, P. (2010). Communicative activity types as organizations in discourses and discourses in organizations. In S. K. Tanskanen, M. L. Helasvuo, M. Johansson, & M. Raitaniemi, *Discourses in Interaction* (pp. 33-59). Amsterdam: John Benjamins Publishing Company.

- London, N., Pogue, G., & Spinuzzi, C. (2015). Understanding the value proposition as a co-created claim. In *Professional Communication Conference (IPCC), 2015 IEEE International* (pp. 1-8). IEEE.
- Lounsbury, M., & Glynn, M. A. (2001). Cultural entrepreneurship: Stories, legitimacy, and the acquisition of resources. *Strategic Management Journal*, 22(6-7), 545-564.
- Lycan, W. G. (2012). *Philosophy of Language: A Contemporary Introduction*. New York: Routledge.
- Martens, M. L., Jennings, J. E., & Jennings, P. D. (2007). Do the stories they tell get them the money they need? The role of entrepreneurial narratives in resource acquisition. *Academy of Management Journal*, 50(5), 1107-1132.
- Maynard, D. W. (1988). Narratives and narrative structure in plea bargaining. *Law and Society Review*, 22(3), 449-481.
- Mizik, N., & Jacobson, R. (2003). Trading off between value creation and value appropriation: The financial implications of shifts in strategic emphasis. *Journal of Marketing*, 67(1), 63-76.
- Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, 96(2), 143-182.
- Morgan, S. E., Cole, H. P., Struttman, T., & Piercy, L. (2002). Stories or statistics? Farmers' attitudes toward messages in an agricultural safety campaign. *Journal of Agricultural Safety and Health*, 8(2), 225.
- Morley, J., & Bayley, P. (2009). *Corpus-assisted Discourse Studies on the Iraq Conflict: Wording the War*. New York, NY: Routledge.
- Morris, M. H., & Jones, F. F. (1999). Entrepreneurship in established organizations: The case of the public sector. *Entrepreneurship Theory and Practice*, 24(1), 71-91.
- Mosey, S., Guerrero, M., & Greenman, A. (2017). Technology entrepreneurship research opportunities: insights from across Europe. *The Journal of Technology Transfer*, 42(1), 1-9.
- Moss, T. W., Renko, M., Block, E., & Meyskens, M. (2017). Funding the story of hybrid ventures: Crowdfunder lending preferences and linguistic hybridity. *Journal of Business Venturing*, forthcoming.
- Nambisan, S. (2017). Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepreneurship Theory and Practice*, 41(6), 1029-1055.
- Nasution, H. N., Mavondo, F. T., Matanda, M. J., & Ndubisi, N. O. (2011). Entrepreneurship: Its relationship with market orientation and learning orientation and as antecedents to innovation and customer value. *Industrial Marketing Management*, 40(3), 336-345.
- Nicholson, L., & Anderson, A. R. (2005). News and nuances of the entrepreneurial myth and metaphor: Linguistic games in entrepreneurial sense-making and sense-giving. *Entrepreneurship Theory and Practice*, 29(2), 153-172.

- Nicotra, M., Romano, M., Del Giudice, M., & Schillaci, C. E. (2018). The causal relation between entrepreneurial ecosystem and productive entrepreneurship: A measurement framework. *The Journal of Technology Transfer, 43*(3), 640-673.
- Niles, J. D. (1999). *Homo Narrans: The Poetics and Anthropology of Oral Literature*. Philadelphia, PA: University of Pennsylvania Press.
- Parhankangas, A., & Ehrlich, M. (2014). How entrepreneurs seduce business angels: An impression management approach. *Journal of Business Venturing, 29*(4), 543-564.
- Parhankangas, A., & Renko, M. (2017). Linguistic style and crowdfunding success among social and commercial entrepreneurs. *Journal of Business Venturing, 32*(2), 215-236.
- Parkinson, C., & Howorth, C. (2008). The language of social entrepreneurs. *Entrepreneurship and Regional Development, 20*(3), 285-309.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC 2001*. Mahway: University of Texas at Austin.
- Pitt, M. (1998). A tale of two gladiators: 'Reading' entrepreneurs as texts. *Organization Studies, 19*(3), 387-414.
- Pollach, I. (2003). Communicating corporate ethics on the world wide web: A discourse analysis of selected company web sites. *Business and Society, 42*(2), 277-287.
- Priem, R. L. (2007). A consumer perspective on value creation. *Academy of Management Review, 32*(1), 219-235.
- Priem, R. L., Li, S., & Carr, J. C. (2012). Insights and new directions from demand-side approaches to technology innovation, entrepreneurship, and strategic management research. *Journal of Management, 38*(1), 346-374.
- Putnam, L. L., & Coreen, F. (2004). Alternative perspectives on the role of text and agency in constituting organizations. *Organization, 11*(3), 323-333.
- Rao, D. (2013). Why 99.95% of entrepreneurs should stop wasting time seeking venture capital. Retrieved from <https://www.forbes.com/sites/dileep/2013/07/22/why-99-95-of-entrepreneurs-should-stop-wasting-time-seeking-venture-capital/#6b90e72946eb>
- Ratinho, T., Harms, R., & Walsh, S. (2015). Structuring the technology entrepreneurship publication landscape: Making sense out of chaos. *Technological Forecasting and Social Change, 100*, 168-175.
- Roberts, C.W. (2000). A conceptual framework for quantitative text analysis. *Quality and Quantity, 34*(3): 259-274.
- Rosen, D. E., Schroeder, J. E., & Purinton, E. F. (1998). Marketing high tech products: Lessons in customer focus from the marketplace. *Academy of Marketing Science Review, 1*, 1-19.
- Roundy, P.T. (2014). The stories of social entrepreneurs: Narrative discourse and social venture resource acquisition. *Journal of Research in Marketing and Entrepreneurship, 16*(2), 200-218.

- Roundy, P.T., (2016). Start-up community narratives: The discursive construction of entrepreneurial ecosystems. *The Journal of Entrepreneurship*, 25(2), 232-248.
- Roundy, P. T., Bradshaw, M., & Brockman, B. K. (2018). The emergence of entrepreneurial ecosystems: A complex adaptive systems approach. *Journal of Business Research*, 86(1), 1-10.
- Roundy, P.T., Harrison, D. A., Khavul, S., Perez-Nordtvedt, L., & McGee, J. E. (2018). Entrepreneurial alertness as a pathway to strategic decisions and organizational performance, *Strategic Organization*, 16(2), 192-226.
- Rydin, Y. (1999). Can we talk ourselves into sustainability? The role of discourse in the environmental policy process. *Environmental Values*, 8(4), 467-484.
- Sax, D. *The Revenge of Analog: Real Things and Why They Matter*. Philadelphia, PA: PublicAffairs.
- Say, Jean-Baptiste. 1803[1964]. *Traité d'économie politique: ou, simple exposition de la manière don't se forment, se distribuent et se consomment les richesses*. Translation: *Treatise on Political Economy: On the Production, Distribution and Consumption of Wealth*. New York, NY: Kelley.
- Schank, R. C., & Abelson, R. P. (1995). Knowledge and memory: the real story. In R.S. Wyer, Jr. (Ed.), *Knowledge and Memory: The Real Story* (pp. 1-85). Hillsdale: Lawrence Erlbaum Associates.
- Schumpeter, J. (1934). *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.
- Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review*, 25(1), 217-226.
- Shane, S., & Venkataraman, S. (2003). Guest editors' introduction to the special issue on technology entrepreneurship. *Research Policy*, 32(2), 181-184.
- Short, J. C., Ketchen Jr, D. J., Shook, C. L., & Ireland, R. D. (2010). The concept of "opportunity" in entrepreneurship research: Past accomplishments and future challenges. *Journal of Management*, 36(1), 40-65.
- Snedecor, G. W., & Cochran, W. G. (1989). *Statistical Methods* (8th ed.). Ames, Iowa: Iowa State University Press.
- Sperber, S., & Linder, C. Gender-specifics in start-up strategies and the role of the entrepreneurial ecosystem. *Small Business Economics*, forthcoming.
- Spinuzzi, C., Nelson, S., Thomson, K. S., Lorenzini, F., French, R. A., Pogue, G., ... & Momberger, J. (2015). Remaking the pitch: Reuse strategies in entrepreneurs' pitch decks. *IEEE Transactions on Professional Communication*, 58(1), 45-68.
- Spinuzzi, C. (2016). Introduction to the special issue on entrepreneurship communication. *IEEE Transactions on Professional Communication*, 59(4), 316-322.
- Spinuzzi, C. (2017). Introduction to special issue on the rhetoric of entrepreneurship: Theories, methodologies, and practices. *Journal of Business and Technical Communication*, 31(3), 275-289.

- Steyaert, C. (2007). 'Entrepreneurship' as a conceptual attractor? A review of process theories in 20 years of entrepreneurship studies. *Entrepreneurship and Regional Development*, 19(6), 453-477.
- Subramanian, R., Insley, R. G., & Blackwell, R. D. (1993). Performance and readability: A comparison of annual reports of profitable and unprofitable corporations. *Journal of Business Communication*, 30(1), 49-61.
- Sveningsson, S., & Alvesson, M. (2003). Managing managerial identities: Organizational fragmentation, discourse and identity struggle. *Human Relations*, 56(10), 1163-1193.
- Taylor, J. R., & Robichaud, D. (2004). Finding the organization in the communication: Discourse as action and sensemaking. *Organization*, 11(3), 395-413.
- Wittgenstein, L. (1922). *Tractatus Logico-philosophicus*. New York: Harcourt, Brace, & Company, Inc.
- van Werven, R., Bouwmeester, O., & Cornelissen, J. P. (2015). The power of arguments: How entrepreneurs convince stakeholders of the legitimate distinctiveness of their ventures. *Journal of Business Venturing*, 30(4), 616-631.
- Vaara, E., Sonenshein, S., & Boje, D. (2016). Narratives as sources of stability and change in organizations: Approaches and directions for future research. *The Academy of Management Annals*, 10(1), 495-560.
- Vermeulen, F. (2007). "I shall not remain insignificant": Adding a second loop to matter more. *Academy of Management Journal*, 50(4), 754-761.
- Wong, A., Bhatia, M., & Freeman, Z. (2009). Angel finance: The other venture capital. *Strategic Change: Briefings in Entrepreneurial Finance*, 18(7-8), 221-230.
- Zarei, M. (2017). Entrepreneurial tournaments: Towards disclosing the rivalry process among corporate entrepreneurs. *Journal of Entrepreneurship, Management and Innovation*, 13(2), 33-57.
- Zollo, L., Rialti, R., Ciappei, C., & Boccardi, A. (2018). Bricolage and social entrepreneurship to address emergent social needs: A "deconstructionist" perspective. *Journal of Entrepreneurship, Management and Innovation*, 14(2), 19-47.

### Abstrakt

Uczeni poświęcają dużo uwagi językowi przedsiębiorczości i jego wpływowi na poznanie, zachowanie i wyniki przedsiębiorców oraz ich interesariuszy. Jednak podstawowe tematy, które stanowią język przedsiębiorców, są wciąż niepoznane. W tym częściowo indukcyjnym badaniu identyfikujemy najczęstsze tematy dyskursu na temat przedsiębiorczości i badamy, jak zmieniały się one z czasem. Aby zidentyfikować tematy w języku przedsiębiorców, używamy technik analizy danych połączonych z algorytmami wyszukiwania tekstów i przeprowadzamy długoterminową analizę istoty dyskursu o przedsiębiorczości. Nasze badania ujawniają pięć dominujących i powtarzających się tematów w dyskursie na temat przedsiębiorczości. Są to: działania marketingowe, przedsiębiorczość ukierunkowana na technologię, przedsiębiorczość cyfrowa, inwestycje profesjonalne i przedsiębiorczość z zakresu nowych przedsięwzięć. Wskazując

*kluczowe tematy dyskursu przedsiębiorców i przedstawiając ich transformację w czasie, nasze badanie wnosi teoretyczny i metodologiczny wkład w badania nad przedsiębiorczością. Wyznaczamy obszary, w których literatura akademicka wydaje się być opóźniona w stosunku do dyskusji praktyków i sugerujemy, że uczeni powinni oceniać badania pod kątem tego, jak ściśle tematy są skalibrowane z głównymi tematami w dyskursie przedsiębiorców. Nasze odkrycia przynoszą także praktyczne implikacje dla przedsiębiorców, identyfikując główne tematy, na które zwraca się uwagę, co pozwala przedsiębiorcom ocenić, czy tematy, które składają się na ich codzienny dyskurs, są zgodne z tematami podkreślanymi w szerszym dyskursie na temat przedsiębiorczości.*

**Słowa kluczowe:** *przedsiębiorczość, komunikacja między przedsiębiorcami, dyskurs, analiza tekstu, analityka danych.*

## Biographical notes

**Philip T. Roundy** is the UC Foundation Assistant Professor of Entrepreneurship and Summerfield Johnston Centennial Scholar at the University of Tennessee at Chattanooga. He earned his Ph.D. in strategic management and organization theory at the University of Texas at Austin. His research interests center on social entrepreneurship, entrepreneurial ecosystems, and the role of entrepreneurship in economic development and community revitalization. His work has appeared in *Strategic Organization*, *Journal of Management Studies*, *Journal of Business Venturing Insights*, *Academy of Management Perspectives*, *Journal of Business Research*, *Journal of Entrepreneurship*, and others. He serves on the editorial boards of *Journal of Business and Entrepreneurship* and *Journal of Applied Management and Entrepreneurship*.

**Arben Asllani** is the Marvin E. White Professor of Management at the University of Tennessee at Chattanooga. He earned his Ph.D. in management information systems and operations management at the University of Nebraska. He is a recognized author, scholar, teacher, and consultant in the areas of business analytics, cybersecurity, information systems, and management science. His work has appeared in *Omega*, *Transfusion*, *European Journal of Operational Research*, *Knowledge Management*, *Computers & Industrial Engineering*, *Total Quality Management & Business Excellence*, and others. He is also the author of "Business Analytics with Management Science Models and Methods," published by Pearson/FT.