View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Carolina Digital Repository

Jimi A. Radabaugh. Informal Language and Its Relationship to the Acquisition of Twitter Followers. A Master's Paper for the M.S. in I.S. degree. July, 2012. 33 pages. Advisor: Richard Marciano

This paper argues that increased use of Informal Language in an organization's Tweets will increase their number of Followers on Twitter. An unobtrusive content analysis was conducted on approximately 12,000 Tweets authored by 18 academic institutions, all members of the iSchools organization, in order to extract instances of Informal Language.

The findings suggest that Informal Language is an effective means of communication in the context of Twitter, and imply more broadly that conforming to the conventions established by any social media platform is an effective means of increasing presence and influence in a given medium.

Headings:

Content analysis (Communication)

Microblogs

Virtual communities

INFORMAL LANGUAGE AND ITS RELATIONSHIP TO THE ACQUISITION OF TWITTER FOLLOWERS

by
Jimi A. Radabaugh

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

July 2012

Approved by

____

Richard Marciano

TABLE OF CONTENTS

INTRODUCTION

In less than six years, the social media platform Twitter has seen exponential growth in the use of its service. More importantly, Twitter has undergone a fascinating transformation from just another corner of the internet where users can engage in meaningless babble to a platform with the capacity for real social impact (as shown by its role in the Arab Spring demonstrations and protests) and the ability to disseminate news at an unprecedented rate, both in terms of speed and the immediate reach of its global audience. While meaningless babble remains a component of Twitter, an increasing number of organizations are taking Twitter seriously as a means of communicating with stakeholders.

In contrast to the similarly ubiquitous social media platform Facebook, Twitter operates more as a broadcast medium, in which interested parties choose to follow the content (thereby becoming Followers) of a Twitter feed without the need for reciprocated approval (i.e. approving a friend request in the Facebook model). In this sense, Twitter promotes the idea of social networks based around common interests regardless of personal associations outside the "Twittersphere," rather than a strict requirement that members of a particular network know each other in any kind of direct, real-world manner. The implications of this are relevant to the following research in that an opportunity exists for organizations to connect with individuals or other organizations outside of a fixed circle of associates, beyond even the associate's circle of associates, and so on.

Of course, it is one thing to simply recognize this potential and quite another thing to effectively capitalize on the potential to reach outside of a known circle of associates in order to produce tangible results. In this case, results can reasonably be measured in terms of the number of Followers a Twitter account has, and particularly in the number of Followers not directly tied to the organization in question. Significantly, despite the fact that the targeted audience is a group with no obvious personal connections to the organization, it is my strong contention that a sense of familiarity should be fostered through the use of various forms of Informal Language (perhaps counter-intuitively among a group of people unfamiliar with one another), and that this is the key to increasing Followers on Twitter.

In the paper that follows, I will argue in favor of the following hypothesis: Increased use of Informal Language in an organization's Tweets will increase their number of Followers on Twitter. This hypothesis should not be taken as a suggestion for organizations to shift from providing information-rich, authoritative content to peddling meaningless babble in order to successfully acquire Twitter Followers. Rather, it should be taken as a provocation for organizations to incorporate Informal Language into Tweets in a purposeful way, in order to facilitate a sense of intimacy and familiarity to increase the likelihood that an unfamiliar audience will willingly engage with its content.

LITERATURE REVIEW

While an online search query for “more Twitter Followers” or some similar query would likely yield a number of articles written by bloggers and supposed experts proposing strategies to quickly add followers, it seems that very little academic research has been conducted on the specific topic of determining what factors have the most impact on increasing the number of Followers on Twitter. This is not to suggest that little research has been done on the subject of Twitter in general; quite the contrary, especially given Twitter’s relatively short history as a company. Some of the ideas discussed in other research are relevant to the present topic and will be discussed below.

A number of studies deal with identification of types of Twitter users and assessments of the types of interaction, ranging from the personal to the impersonal. Wu, Hofman, Mason and Watts (2011) noted a high correlation between categories of users and categories of content those users tended to follow, describing Twitter as highly “homophilous,” in the sense that celebrities tend to follow other celebrities, bloggers tend to follow other bloggers, etc. At the other end of the humanity spectrum, Chu, Gianvecchio, Wang and Jajodia (2010) have explored the distinction between human content producers versus automated “bots,” finding that a disproportionate external URL ratio is generally associated with “bot”-produced content. For the purposes of the present study, it is important to note that new Followers are likely to come from the same user category as the organization being followed and that relying too heavily on external links will, if not necessarily give the impression that the content was produced by a robot, perhaps suggest an undesirable degree of impersonal communication.

In terms of the far-reaching broadcast potential of Twitter, Kwak, Lee, Park and Moon (2010) have made the compelling assertion that “any re-tweeted tweet is to reach an average of 1,000 users no matter what the number of followers is of the original tweet.” Clearly, this suggests that re-tweeting (i.e. re-posting tweets originally posted by others) carries an impressively high potential for exposure to a larger audience. However, there is no evidence to suggest that greater exposure automatically ensures an increase in Followers.

Other studies have explored the concept of user influence, which is notoriously difficult to quantify, but nonetheless an important area to investigate. Cha, Haddadi, Benevenuto and Gummadi (2010) highlighted the fact that influence “is not gained spontaneously or accidentally, but through concerted effort.” Cha et al. (2010) discovered that in the context of Twitter, focusing on a single topic or a narrow range of topics and maintaining the focus over a long period of time is a highly effective way to increase one’s influence.

A similar study by Anger and Kittl (2011) argued that mentions by other Twitter users could be used to gauge influence, and suggested attaching a sentiment rating (i.e. whether the context of the mention was positive or negative) in order to further increase the utility of this measure. Both of these studies offer intriguing approaches to the problem of determining how much influence a given Twitter user has over those within that user’s associated network. Although again, there is no evidence to suggest that a high degree of influence translates directly into an increase in Followers.

Another approach that is relevant to the present research is the area of user classification. Choudhury, Diakopoulos and Naaman (2012) separated the Twitter universe into three broad categories: organizations, journalists/media bloggers, and ordinary individuals. Choudhury et al. (2012) found that organizations as a whole pose fewer questions than the other categories, and that individuals tend to use more "sentiment words (e.g. excited, awesome, bad)." Based on this information, increasing the number of questions and sentiment words in the Tweets of organizations would mirror the behavior of individuals and contribute to a more familiar tone, which fits the model of using Informal Language to increase the number of Followers.

A similar study by Rao, Yarowsky, Shreevats and Gupta (2010) examined user classification in the context of predicting certain attributes based on language usage variations. The study by Rao et al. (2010) was primarily concerned with teasing out distinctions along gender, age, regional origin and political orientation lines, but the list of socio-linguistic features they used to classify these distinctions will form the basis of the operational definition for Informal Language used in this paper. Although the conclusions drawn by Rao et al. (2010) regarding the varied usage of these features are interesting in their own right, their overall function as indicators of Informal Language are particularly useful in the present context and will be discussed in further detail below.

METHODOLOGY

Twitter is an ideal platform for content analysis, as all of its content is public and easily accessible and the raw numbers for Followers and Tweets are recorded and made available by the system itself. Until this point, the units of analysis have only been generally defined as "organizations." In order to narrow this definition and establish a sample appropriate in both quality and quantity to the task, I will focus on the iSchools organization (www.ischools.org), and more specifically the Twitter accounts of all 36 institutions involved. This systematic sample of organizations is familiar to the researcher and likely to be familiar to the present audience as well.

The two key variables in this study will be number of Followers and instances of Informal Language. We will first tackle the outcome variable, which is generally defined as number of Followers. On its own, number of Followers is a useful metric, but the main issue with using this number in an unmodified form is the potential for unreliable data owing to the substantial variation among iSchools in the frequency and duration of their respective activity on Twitter.

A potential solution to this problem is to use a ratio based on other relevant information. Other researchers, such as Anger and Kittl (2011), have suggested the Follower/Following ratio as a useful ratio. However, I believe that the ratio of Followers to Tweets will be even more useful, as it will reflect the overall impact of the total amount of content being produced in the context of how many Followers that content reaches. Essentially, the Followers to Tweets ratio will be a measure of how much "bang for your buck" each Tweet has produced, theoretically removing the bias that would result if iSchools that generated a small amount of content were expected to attract the same number of Followers as iSchools that generated a large amount of content.

The second variable will be instances of Informal Language. As mentioned previously, I will use the list of Socio-Linguistic features produced by Rao et al. (2010) as a basis for the operational definition of this variable. Their list includes fifteen features, namely Emoticons, OMG, Ellipses, Possessive Bigrams, Repeated Alphabets, Self-References, Laugh, Shout, Exasperation, Agreement, Honorifics, Affection, Excitement, Single Exclaim, and Puzzled Punctuation. Taken together with the findings suggested by Choudhury et al. (2012), which were an increase in questions and sentiment words, I will re-combine these features into the following seven categories: Emoticons, Abbreviated Text-Speak, Shout/CAPS, Repeated Alphabets, Repeated Punctuation, Exclamation, and Sentiment Words.

The first category, Emoticons, will be defined as any combination of punctuation that attempts to mimic a facial expression, including :-), :-D, and so forth. Abbreviated Text-Speak will include both the category outlined by Rao et al. (2010) as OMG (Oh my God), as well as expressions such as LOL (Laugh out loud), ROTFL (Rolling on the floor laughing) and so forth. The Shout/CAPS category will include any instance of a word or string of words formatted in all caps for emphasis. Examples of Repeated Alphabets include “niceeeee,” “noooo waaaay,” etc.

Punctuation is another important form of informal expression. The Exclamation category will be comprised of single exclamation points. Repeated strings of exclamation points, as well as any combination of question marks and other punctuation (typically exclamation marks) will be included in the Repeated Punctuation category.

Finally, the Sentiment Words category figures to be the least clear-cut, but will include words that indicate an emotionally colored emphasis, such as “awesome” or “super.”

Once the data has been coded according to the above categories, the total numbers for each category will be compared to the Followers to Tweets ratio for each of the iSchools. It is expected that institutions with high Followers to Tweets ratios will have the largest numbers in each category of Informal Language, but the categories are broken down in the event that some categories do not support this hypothesis. The primary advantage of an unobtrusive content analysis, however, is that the data can always be re-examined and re-coded should other significant patterns emerge during the course of the research. This is not to suggest a lack of confidence in the hypothesis, but rather an acknowledgement of the possibility that unanticipated, yet stronger arguments may present themselves.

The inherent limitation of any content analysis is that the opportunity for deeper analysis beyond the published content of the Tweets does not exist. Additionally, a variety of factors contribute to the number of followers a given organization has, and the described methodology only hopes to effectively isolate one of those possible factors, at the risk of appearing to ignore all others. Furthermore, the relative value of the content itself in terms of relevance or importance to potential Followers is virtually impossible to quantify, but nonetheless likely to factor heavily into one’s decision to follow or not to follow.

As mentioned previously, the expected outcome of this research is the discovery that instances of various forms of Informal Language are positively correlated with a strong Followers to Tweets ratio. This conclusion stands to immediately affect the social media strategies of the iSchools included in the sample, but the ultimate goal is that this research will benefit any organization with an interest in expanding its Twitter presence beyond known associates.

By selecting a sample composed of iSchools, this research will be grounded in a familiar context that should strengthen the perceived reliability of this study and it may even be the case that the audience is generally familiar with the majority of the content found in the Tweets under analysis. Of course, it is expected that the present findings will have implications for other types of institutions as well.

RESULTS

Of the 36 institutions involved in the iSchools organization, 18, or exactly half, were studied. As it turns out, 14 of the institutions were ineligible on the basis of two factors. The first factor was simply the lack of a dedicated Twitter account for the institution in question. Every institution in the iSchools organization is affiliated with a university or academic organization, and while all of these organizations maintain some kind of Twitter presence, this study is focused on Information Schools (and associated disciplines where applicable, such as Computer Science, Informatics, etc.). In other words, while the University of Washington maintains a Twitter account, the Information School at the University of Washington does not, and therefore that institution was determined to be ineligible. The second factor was ineligibility due to Tweets having been produced in a foreign language.

Additionally, four other institutions were not included in the study on the grounds of constituting either too small or too large of a sample. The Twitter accounts for the University College London and the University of California, Los Angeles contained fewer than 100 Tweets, which was considered too small of a sample to produce reliable results. On the other end of the spectrum, Carnegie Mellon University (nearly 3,000 Tweets) and Syracuse University (more than 10,000 Tweets and counting) represented an unreasonably large amount of Tweets for individual coding and analysis.

In the end, 18 institutions met the criteria of having a dedicated Twitter account within the department, which had produced between 100 and 2,000 Tweets in English. These institutions, and their associated Twitter handles, are as follows:

1. University of British Columbia – School of Library, Archival & Information Studies (@slaisubc)
2. University of California, Berkeley – School of Information (@BerkeleyISchool)
3. University of California, Irvine – The Donald Bren School of Information and Computer Sciences (@UCIbrenICS)
4. Drexel University – College of Information Science and Technology (@iSchoolatDrexel)
5. Georgia Institute of Technology – College of Computing (@gtcomputing)
6. University of Illinois – Graduate School of Library and Information Science (@gslis)
7. Indiana University – School of Informatics and Computing (@iusoic)
8. University of Kentucky – College of Communication & Information Studies (@uk_ci)
9. University of Maryland – College of Information Studies (@I_UMD)
10. University of Michigan – School of Information (@umsi)
11. University of North Carolina – School of Information and Library Science (@uncsils)
12. University of North Texas – College of Information (@UNTCOI)
13. The Pennsylvania State University – College of Information Sciences and Technology (@ISTatPENNSTATE)
14. Rutgers, the State University of New Jersey – School of Communication and Information (@RutgersCommInfo)
15. University of Sheffield – Information School (@Shef_iSchool)

16. University of Texas, Austin – School of Information (@UTiSchool)

17. University of Toronto – Faculty of Information (@ischool_TO)

18. University of Wisconsin-Milwaukee – School of Information Studies (@uwmsois)

Results were tabulated by reviewing each Tweet from the institution in question and marking each instance of informal language in the appropriate category. The data for all institutions were then grouped into a single table for purposes of comparison, as shown in the figure below.

Institution	**IL**	**!**	**!?!**	**:-)**	**Aaa**	**CAPS**	**LOL**	**Sentiment**
Georgia Tech	383	182	4	34	7	20	22	114
Penn State	248	156	9	10	1	25	10	37
Michigan	213	136	14	8	4	10	11	30
Toronto	205	127	0	0	0	2	21	55
Drexel	191	121	11	10	1	16	10	22
Indiana	172	132	7	3	0	14	0	16
Rutgers	167	66	0	13	1	6	39	42
Illinois	123	91	6	0	4	5	1	16
Berkeley	119	81	4	0	3	3	6	22
Texas	101	52	5	0	4	3	9	28
North Carolina	98	54	8	1	0	9	7	19
Maryland	91	41	0	0	0	35	4	11
Irvine	69	49	2	0	0	7	5	6
Milwaukee	63	40	6	2	0	2	2	11
Kentucky	51	32	1	0	0	6	5	7
Vancouver	33	24	1	0	0	0	1	7
North Texas	31	24	3	1	0	1	0	2
Sheffield	29	13	0	0	0	0	1	15

Figure 1: Table of Informal Language instances by category

In Figure 1, institutions are listed in descending order according to the total number of Informal Language instances. However, this arrangement fails to take into account the context for those numbers, as institutions with a higher number of Tweets would surely have a greater chance at accumulating more instances of Informal Language. In the next figure, the numbers of Tweets and Followers for each institution as of June 29, 2012 have been included alongside the Informal Language instances and associated ratios, which will be discussed below.

Institution	**F/T**	**IL/T**	**Followers**	**Tweets**	**IL**
Toronto	4.49	0.68	1346	300	205
Texas	3.30	0.58	574	174	101
Rutgers	2.31	0.36	1078	467	167
Georgia Tech	1.99	0.30	2541	1276	383
Vancouver	1.53	0.31	164	107	33
Berkeley	1.48	0.23	757	512	119
Penn State	1.37	0.25	1347	981	248
Sheffield	1.33	0.21	181	136	29
North Carolina	1.20	0.16	745	623	98
Milwaukee	1.04	0.23	290	279	63
Maryland	1.02	0.24	392	383	91
Drexel	0.95	0.30	596	629	191
Kentucky	0.87	0.16	284	328	51
Illinois	0.76	0.07	1319	1737	123
Michigan	0.70	0.12	1196	1720	213
Indiana	0.69	0.13	883	1277	172
North Texas	0.63	0.14	143	228	31
Irvine	0.52	0.09	396	759	69

Figure 2: Table of Informal Language instances and Twitter statistics

In Figure 2, institutions are listed in descending order according to the Followers to Tweets ratio, abbreviated here as F/T. The next column shows the Informal Language instance to Tweets ratio, abbreviated here as IL/T. Even a brief scan of these two columns appears to confirm the hypothesis that these two ratios are positively correlated, as the institution with the largest F/T also has the largest IL/T, while the institution with the smallest F/T has the second smallest IL/T, with a similar relationship occurring in all points in between. In terms of statistical analysis, the ratios F/T and IL/T have a Correlation of 0.95573 and the probability of the null hypothesis (Prob $> |t|$) is $< .0001$. In other words, the data supports the hypothesis that more Informal Language instances increases the likelihood of more Followers relative to the number of Tweets.

ANALYSIS

Once again, the seven categories of Informal Language and their corresponding abbreviations in the column headings of Figure 1 are: Exclamation (!), Repeated Punctuation (!?!), Emoticons (:-)), Repeated Alphabets (Aaa), Shout/CAPS (CAPS), Abbreviated Text-Speak (LOL), and Sentiment Words (Sentiment). It is important to note that acronyms composed of capital letters, which were seen frequently, were not counted as belonging to the Shout/CAPS category. Full lists of specific examples of Abbreviated Text-Speak and Sentiment Words used in Tweets can be found in the Appendices section.

It quickly became apparent while coding the data that the Exclamation category, defined by this study as a single exclamation point, was the most commonly occurring category. In fact, the Exclamation category comprised a significant portion of the Informal Language instances (approximately 60%), representing anywhere from 40% to 77% of the total for each institution. This is problematic because it could be argued that the use of an exclamation point is a weak expression of Informal Language in that it is not as radically different from any standard definition of Formal Language as examples such as Emoticons or Abbreviated Text-Speak.

Considering this potential over-emphasis on the Exclamation category, the following figure shows how the data changes if the Exclamation numbers are removed from the total number of Informal Language instances.

Institution	F/T	IL -!/T	IL -!	Tweets
Toronto	4.49	0.26	78	300
Texas	3.30	0.28	49	174
Rutgers	2.31	0.22	101	467
Georgia Tech	1.99	0.16	201	1276
Vancouver	1.53	0.08	9	107
Berkeley	1.48	0.07	38	512
Penn State	1.37	0.09	92	981
Sheffield	1.33	0.12	16	136
North Carolina	1.20	0.07	44	623
Milwaukee	1.04	0.08	23	279
Maryland	1.02	0.13	50	383
Drexel	0.95	0.11	70	629
Kentucky	0.87	0.06	19	328
Illinois	0.76	0.02	32	1737
Michigan	0.70	0.04	77	1720
Indiana	0.69	0.03	40	1277
North Texas	0.63	0.03	7	228
Irvine	0.52	0.03	20	759

Figure 3: Table of Informal Language instances minus Exclamation

While perhaps not as convincing of a comparison as F/T vs. IL/T seen in Figure 2, a clear relationship between the two columns remains even after the Exclamation category has been removed, expressed here as IL -!/T. In terms of statistical analysis, the ratios F/T and IL -!/T have a Correlation of 0.90922. Thus, it may be argued that removing the most prominent, yet potentially also the least informal of the seven categories of Informal Language (Exclamation) does little to alter the overall impression that more Informal Language instances increases the likelihood of more Followers relative to the number of Tweets.

Continuing our closer examination of the data, let's compare the numbers for Georgia Tech and Drexel. As seen in Figure 2, both institutions have an IL/T of 0.30, but Georgia Tech ends up with a much higher F/T – 1.99 compared to Drexel's 0.95. In other words, both institutions used Informal Language in the same proportion (and in similar proportion minus the Exclamation category – 0.16 vs. 0.11), but for some reason, Georgia Tech has attracted roughly twice as many Followers per Tweet. These numbers are shown in the figure below.

Institution	**F/T**	**IL/T**	**IL -!/T**
Georgia Tech	1.99	0.30	0.16
Drexel	0.95	0.30	0.11

Figure 4: Table comparing Georgia Tech and Drexel

There are two ways to look at this example. The first would be to argue that Drexel is simply an outlier, as it is ranked #12 according to the F/T ratio while it is tied for #4 according to the IL/T ratio, which is the largest ranking discrepancy among all of the institutions. The second way to look at this example would be as a suggestion that the real difference between these two sets of numbers lies in the details. A breakdown of all seven categories of Informal Language is shown in the figure below.

Institution	**IL**	**!**	**!?!**	**:-)**	**Aaa**	**CAPS**	**LOL**	**Sentiment**
Georgia Tech	383	182	4	34	7	20	22	114
Drexel	191	121	11	10	1	16	10	22

Figure 5: Table comparing Georgia Tech and Drexel by category

Conveniently for the sake of comparison, the overall number of Informal Language instances (IL) for Georgia Tech is almost exactly double that of Drexel (383 vs. 191). Thus, we could reasonably expect the numbers in each category to be approximately twice as large for Georgia Tech when compared to Drexel's numbers. Looking back over Figure 5, the largest discrepancy is apparent, as Georgia Tech has more than five times the number of Sentiment Words as Drexel (114 vs. 22). This suggests that the relatively high frequency of Sentiment Words may be a factor in Georgia Tech attracting a significantly higher ratio of Followers relative to its number of Tweets. A table isolating only the Sentiment Words category for all of the institutions is shown in the figure below.

Institution	**F/T**	**S/T**	**Sentiment**	**Tweets**
Toronto	4.49	0.18	55	300
Texas	3.30	0.16	28	174
Rutgers	2.31	0.09	42	467
Georgia Tech	1.99	0.09	114	1276
Vancouver	1.53	0.07	7	107
Berkeley	1.48	0.04	22	512
Penn State	1.37	0.04	37	981
Sheffield	1.33	0.11	15	136
North Carolina	1.20	0.03	19	623
Milwaukee	1.04	0.04	11	279
Maryland	1.02	0.03	11	383
Drexel	0.95	0.03	22	629
Kentucky	0.87	0.02	7	328
Illinois	0.76	0.01	16	1737
Michigan	0.70	0.02	30	1720
Indiana	0.69	0.01	16	1277
North Texas	0.63	0.01	2	228
Irvine	0.52	0.01	6	759

Figure 6: Table of Sentiment Words

By isolating the Sentiment Words category from the other categories of Informal Language, we have discovered another way of viewing the data that corresponds with the initial hypothesis, but which narrows the focus considerably. In terms of statistical analysis, the ratios F/T and S/T have a Correlation of 0.94078. This is nearly as strong as the Correlation between the ratios F/T and IL/T, which is 0.95573. In other words, this isolated chunk of data (Sentiment Words), which represents roughly 20% of the total Informal Language instances, yields statistical results that are practically indistinguishable from those yielded by 100% of the data.

In summary, Figure 2 shows the comparison between IL/T and F/T, which takes into account 100% of the data. Figure 3 compares IL -!/T and F/T, which removes the Exclamation category, leaving 40% of the data to produce similar results. Figure 6 compares S/T and F/T, which removes an additional 20% (the sum of the remaining five categories), leaving 20% of the data to again produce similar results. Admittedly, the nearly 12,000 Tweets that were coded for the purpose of this study are a virtual drop in the bucket in the context of the larger "Twittersphere," but these numbers do suggest that ratios for Sentiment Words could closely parallel ratios for Informal Language instances as a whole. It stands to reason that this notion could be practically applied to research into the Tweets of other types of organizations as well.

DISCUSSION

In the tabulation of Sentiment Words from the Tweets under analysis, an interesting trend emerged regarding their consistent tone. Broadly defined, a Sentiment Word may include anything with an emotionally colored emphasis, ranging from positive words like “great” or “awesome” to negative words like “awful” or “terrible.” However, within the scope of this study, virtually all of the Sentiment Words encountered from every institution were positive (A full list of Sentiment Words can be found in the Appendices section). This trend is understandable given the kind of relationship each institution has with its intended audience.

In this context, Tweets have been used primarily to inform, encourage and congratulate. The implied audience in nearly every case appears to be students or prospective students and quite a few announcements have to do with wishing them luck or congratulating them on relevant achievements. As this paper has argued, the language being used in the Tweets is an important factor in the acquisition of Followers, but given the lack of negatively expressed Sentiment Words, it is impossible to tell from this sample how much of an effect negativity has versus the effect of positivity within this category of Informal Language. In the application of the present research to other types of organizations, the appearance of negative Sentiment Words and the implications on attracting or not attracting Followers would have to be considered.

Another unanticipated outcome was that the numbers were generally consistent across the board in every category, meaning that for the most part, institutions tended to utilize or avoid categories of Informal Language at very similar rates. Certainly, some institutions relied more heavily on Emoticons for example, whereas others may have favored CAPS a bit more. Overall though, whatever stylistic distinctions there may have been, the general impression is that each category of Informal Language represented a fairly consistent percentage of the total Informal Language instances for each institution. Once again, this would potentially not be true if the research was broadened to include other types of organizations, as opposed to focusing on institutions within the iSchools organization, as the present research has done.

In addition to the data collected on Informal Language and the role it appears to play in determining the relative success of a Twitter account, a few general observations emerged regarding the kinds of things potential Followers look for (or dismiss) in Tweets. The first point to consider is that the majority of Followers tend to be individuals rather than other institutions, which further reinforces the idea that a familiar tone is more appropriate, as the implied interaction is likely to more closely resemble a person-to-person dialogue. It is important to remind ourselves here that Twitter is a unique social media platform in that it functions both as a broadcast medium and a conversational medium. In the conscious development of the former, one should not overlook the latter.

Similarly, Tweets that appear to be automatically generated from the institution's website are likely to give a potential Follower the impression that he or she is the recipient of a broadcast, rather than a participant in a conversation. This is less likely to engage an individual who not only wishes to be informed, but more importantly wishes to be involved. Simply put, Tweets that come across as formal or authoritative are less likely to make an impression on potential Followers than Tweets that come across as familiar and friendly. Informal Language is a significant part of this equation, and incorporating such elements into the content appears to be an appropriate way to attract potential Followers. But it is important to recognize that Informal Language is simply the most visible manifestation of a larger goal, which is promoting the interests of an institution or an organization by establishing and maintaining connections with individuals.

CONCLUSION

The overall goal of this research is to contribute positively to the ability of organizations to attract followers on Twitter outside of a fixed circle of associates. It is my contention that the deliberate incorporation of Informal Language into otherwise formal content contributes to a familiar atmosphere in which potential Followers will be more receptive to unfamiliar content and by extension, more receptive to the organization itself. The research outlined above indicates that there is a significant correlation between Informal Language instances and the number of Followers relative to the number of Tweets. This suggests that increasing the use of Informal Language in all of its various forms is an effective strategy towards the goal of increasing Followers on Twitter.

In terms of an organizational best-practice statement, the results of this study suggest a reasonably clear guideline to include more of the outlined forms of Informal Language in the content of an organization's Tweets. The only additional recommendation for the best-practice statement would be to exercise some restraint in the deliberate incorporation of Informal Language. This research should not be taken as a provocation to conclude every Tweet with fifteen exclamation points or to add Abbreviated Text-Speak until Tweets become exaggerated strings of nonsense. Rather, organizations should strive to achieve a balance between the familiar, informal forms of expression highlighted in this study and the information-rich content already being provided.

Of course, as the Analysis section of this paper has revealed, not all Informal Language instances are created equal. Specifically, Sentiment Words appear to be the one category among the seven (Emoticons, Abbreviated Text-Speak, Shout/CAPS, Repeated Alphabets, Repeated Punctuation, Exclamation, and Sentiment Words) that bears the most resemblance in isolation to the total numbers for Informal Language instances. This category, which represents 20% of the data, appears to be a strong predictor of the overall effect that Informal Language has on the acquisition of Twitter Followers. However, without the benefit of additional research using the same categories of Informal Language, it is impossible to tell whether this category is as significant as it appears or simply an anomaly of this particular sample.

It is also important to note that the conventions of Twitter were derived from the conventions of SMS, or text-messaging. Shortening words to fit into the space allotted was one reason for abbreviating, but there is also the component of text messages being composed quickly in order to communicate in an off-the-cuff manner. Twitter represents a different kind of environment in the sense that Tweets remain in the "Twittersphere" and anyone can review the entire backlog at any time, as I have done here. However, the atmosphere of communicating off-the-cuff, which gives the impression of information communicated quickly and directly without over-editing or over-scrutinizing the information being presented remains an important aspect of Twitter.

This leads us to the larger point that success on any social media platform hinges on the ability of the participant to conform to the conventions of the medium at hand. Twitter is a good example of this, but the general concept can be applied to any social media platform in which an organization wishes to increase its presence and influence.

The important thing to stress here is that organizations should conform to the conventions that already exist within Twitter, as opposed to viewing Tweets as merely an extension of the tone adopted for communication on the organization's website or in printed formats. All of the forms of Informal Language outlined above are suggested here as ways of achieving this goal.

REFERENCES

Anger, I. & Kittl, C. (2011). Measuring influence on Twitter. *i-KNOW '11: Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies.*

Cha, M., Haddadi, H., Benevenuto, F., & Gummad, K.P. (2010). Measuring user influence on twitter: The million follower fallacy. *4th Int'l AAAI Conference on Weblogs and Social Media*, Washington DC.

Choudhury, M., Diakopoulos, N., & Naaman, M. (2012). Unfolding the event landscape on twitter: classification and exploration of user categories. *CSCW '12: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work.*

Chu, Z., Gianvecchio, S., Wang, H., & Jajodia, S. (2010). Who is tweeting on Twitter: human, bot, or cyborg? *ACSAC '10: Proceedings of the 26th Annual Computer Security Applications Conference*.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? *WWW '10: Proceedings of the 19th international conference on World wide web.*

Rao, D., Yarowsky, D., Shreevats, A., & Gupta, M. (2010). Classifying latent user attributes in twitter. *SMUC '10: Proceedings of the 2nd international workshop on Search and mining user-generated contents.*

Wu, S., Hofman, J.M., Mason, W.A., & Watts, D.J. (2011). Who says what to whom on twitter. *WWW '11: Proceedings of the 20th international conference on World wide web.*

APPENDICES

Appendix 1: Table of Abbreviated Text-Speak

Abbreviation	**Meaning**
2	To
4	For
2nite	Tonight
4get	Forget
4S	Force
4ward	Forward
asap	As soon as possible
attn	Attention
b/c	Because
b4	Before
b-cuz	Because
blv	Believe
bsmnt	Basement
c	See
cld	Could
cuz	Because
DM	Direct Message
eng	English
esp	Especially
folo	Follow
foloing	Following
FTW	For The Win
fwd	Forward
grt	Great

int'l	International
IRL	In Real Life
LOL	Laugh Out Loud
mgmt	Management
mgrs	Managers
mkt	Market
msg	Message
nite	Night
OMG	Oh My God
OMGBBQ	Oh My God, Barbecue
OWS	Occupy Wall Street
pls	Please
plz	Please
ppl	People
r	Are
S/O	Shout Out
sez	Says
thks	Thanks
tho	Though
thx	Thanks
tix	Tickets
tomo	Tomorrow
ur	You’re
w/	With
w/o	Without
wk	Week
wknd	Weekend
wks	Weeks
yrs	Years

Appendix 2: List of Sentiment Words

Absolutely, Amazing, Astounding, Awesome, Beautiful, Best, Better, Brilliance, Brilliant, Completely, Cool, Coolest, Cute, Delicious, Embrace, Enjoy, Enjoyable, Enjoying, Epic, Excellent, Excited, Excitement, Exciting, Eye-popping, Fantastic, Favorite, Friendly, Fun, Funny, Glad, Good, Gorgeous, Gratitude, Great, Happy, Heartfelt, High-5, Hooray, Huge, Important, Impressive, Incredible, Inspiration, Inspirational, Inspired, Interesting, Keen, Laugh, Like, Love, Loved, Lovely, Loving, Major, Marvelous, Neat, Nice, Nicely, Outstanding, Overwhelmed, Phenomenal, Pleased, Positive, Powerful, Praise, Psyched, Rad, Rocks, Scrumptious, Shucks, Sigh, Special, Strong, Super, Sweet, Tasty, Terrific, Thrilled, Thrive, Totally, Useful, Very, Voila, Whoa, Wonderful, Woot, Wow, Yay!