EMAIL SUBJECT LINES ANALYSIS FOR HIGH OPEN RATE IN EMAIL MARKETING

Codrin Teiu[1]
[1]*"Alexandru Ioan Cuza" University of Iasi, Romania*
E-mail: codrin.teiu@gmail.com

Abstract
Email marketing is an important channel for online sales. Millions of emails are being sent daily by companies looking to promote and sell their products and services. Out of all these emails only a small part is being opened, but still companies keep on using this channel for its good reach to customers who are most often taking the email with them everywhere because of the widespread use of mobile phones. This paper is looking to study the behavior of email recipients of a large UK telecom company which sent a large volume of marketing related emails in 2019. We are looking to gauge the success of the subject lines of these emails by taking email open rate as an indicator and then determine a series of best practices and common traits by analyzing the best performing subject lines.

Keywords
email marketing, email subject line, email open.

JEL Classification
M3, M37

Introduction
Email marketing is a popular method for companies to contact prospects. This channel is also one of the most cost efficient channels of marketing communications (Dawe, 2015) and this is why many business-to-business (B2B) marketers are relying on email as a foundational piece for their sales strategy. Still, nowadays the quantity of emails that a person receives can be overwhelming and recipients are prioritizing some emails over others, while many of them remain unopened altogether. In prioritizing the emails that they read, besides looking at the sender who should be a known entity that has the right to contact the person for marketing purpose, recipients are looking first at the subject line of the emails. This is the communication that they see first and a big factor in whether that email will be opened or not. Moving forward from here, the personal utility of the message takes over the prioritization process (Wainer , et al., 2011).

Email marketing, in comparison with direct mail, has the ability to capture prospect information and interest via gated landing pages where the prospect can submit a form to request more information, to register for an event or to ask to be contacted, out of many other possibilities. The days of standalone marketing campaigns done via email are long gone, but part of an integrated multi-channel marketing campaign email is still worth the investment. Recent data privacy regulations related to email communications for sales have made the email marketing environment more secure and email subscribers feel protected from being contacted without permission. In business to business communications and sales it is mandatory to obtain permission from prospects to be contacted by email for marketing purpose.

Review of the scientific literature
The number of email communications sent each day is large and growing year after year. The technology research organization The Radicati Group has estimated the number of worldwide email users in 2019 at over 2.9 B (billion), which is over one third of the world population. The number of email communications sent per day in 2019 are calculated at 246.5 B emails, out of which 128.8 B emails per day are accounted for as business emails. The average consumer worldwide received 96 emails per day in the year 2019 (The Radicati Group, 2019). Email has a series of characteristics that make it a very business to business competitive tool for marketing: it is targetable (specific individuals can easily be targeted by doing database segmentation on various criteria), is highly customizable and open for personalization depending on the recipient, has high flexibility as a tool and low commitment needed to run email campaigns (Baggott, 2007).
Other widespread tools like telemarketing or direct mail have high costs to run which meant that large companies had little competition from smaller companies that wanted to enter their market. Email marketing has started to allow this smaller companies to have access to a potent marketing tool to compete with larger companies.
Besides these points an important factor to consider is that nearly everyone uses email as it is a needed tool for both professional and personal use. In comparison to other media, the engagement rates of email are consistently high which makes is a tool with good reach at an affordable cost (Baggott, 2007). It is true that a part of all emails sent are marked as spam (undesirable communication) by recipients, most often in cases of bought email lists where permission to be emailed was obtained through a third party (for example you agree to receive marketing communication from your cable company by they are also sharing contacts with other partner companies) or marketers not doing a proper job in targeting prospects.
Email is often used as a tool in research, many times researchers sending solicitations to participate in a web-based survey (Sappleton & Lourenco, 2015). When this type of communications are being sent, similarly to sales emails, senders are looking to improve the response rate. This process starts with having the contacts read the actual email and the subject line plays a major role in this process. Similarly to phone sales and direct mail sales, the subject line is the opening statement and it has to catch the interest of the person reading it in order for them to process further. This is why there is solid reason to state that a recipient may be sensitive to the message sent via the subject line.
The usual email subject lines have formats that we can sometime expect depending on the content and sender. Sometimes the subject line works toward giving legitimacy to the message, or offer some sort of prize or reward, be an appeal for help for a good cause, ask an interesting or controversial question or even something personal to the recipient like using some of their personal info or past purchasing history (Porter & Whitcomb, 2005).
Even if email is a widely used channel for marketing, there is little academic research looking for the link between email subject line and response rates. A series of experiments looked at survey responses increase or decrease coming from emails with different subject lines. There

was no significant difference from emails with generic subject lines to emails with customized subject lines in matter of survey completion rates (Callegaro, et al., 2009). However, an analysis performed in messages posted in forum newsgroups found that the addition of a question mark increase the likelihood of a response (Burke, et al., 2007).

While some researchers have managed to reveal that the subject line is used a filtering mechanism when reading content, in a case study examining the role of a subject line there could not be made a link between subject line categories and the number of times the message was read (Skogs, 2013). Categories were referring to subject lines reflecting the content of the message, other describing the function of the message while lastly other ones including greetings or salutations.

Within the literature cited on this topic a specific study stands out as one of the most cited and it looked to assess the effect of the email subject line on survey viewing and survey response (Porter & Whitcomb, 2005). A large number of students were asked to participate in a survey related to their university life and were split in four categories of recipients based on the subject line they were going to get. One of the groups received emails with a subject line containing the word "survey", another group received a line containing the name of the university, a third group received a blank subject line while the last one was sent a plea for help. The main observation of this study was that most students opened the email with a blank subject line and the least opened subject line was the one containing the word "survey". Based on this there were a series of takeaway conclusions from this. One is that messages having a subject line that has an information gap, or no subject line at all will increase the likelihood for an open and a read of the message out of human curiosity. Another one was that messages making a plea for help were seen as spam messages.

The main idea coming out of the previously presented case study is that silence in communication can provoke a response. Having an information gap in the message will appeal to human nature to try to close that gap even if it's not following their immediate interest. Many marketers have tried more or less successful to benefit from this human behavior, although while previously presented studies were of academic nature and with small audiences, in real life the ability to utilize a blank subject line to try to precipitate a response has either been hindered by the widespread of spam and totally made an impossible practice by the most recent spam filter regulation of internet service providers. These providers would automatically flag a message without a subject line as junk or spam thus not allowing it to reach an inbox folder.

While the literature of response rates to emails in connection to surveys responses is large, the specific use of the subject line in this process is not addressed in great detail. Research connecting email subject lines to open rates and then to direct responses (whether they are a response to a survey, or a purchase, or a visit to a link in the email) is scarce with live business to business email marketing data, while most of these previous researches were done in either an academic setting or user forum scenarios.

The purpose of this study is to be a road opener and partially address the existing gaps in literature by correlating subject lines to email opens using live marketing emails with real life recipients.

Research methodology

This research is looking to analyze marketing data from marketing communications sent in the United Kingdom by a major telecom company in the European Union during the year of 2019. This marketing communication are primarily sent to the UK market.

The raw data is a database extract containing all email names and subject lines, the numbers of sent and delivered emails per piece, the number of opened emails, the open rate and the unique open rate.

All these communications were business to business communications. For matters or relevance, we will exclude from the data sample email communications being sent to less than 1000 recipients.

We will be calculating an average open rate for all 2019 sends by reporting ourselves to the total number of sent emails that year and the ratio of each individual email sent within that total.

We will be analyzing the resulting data in matters of size from the total numbers of communications sent, in matter of the best overall performing message and it matters of similarity between the different subject lines. As this research paper is a first glimpse at the topic, we will be looking to gauge the average number of words in these subject lines, the average number of characters and to categorize as simply as possible the type of message (question, plain statement, exclamation, prize/sales reduction offering).

Results and discussion

The database resulted for this research consists of 263 different email matching the criteria of being sent each to over 1000 recipients at least. The total number of emails sent using these 263 emails totals 4.1 M (million) emails being sent out of which only 3.54 M emails were received. The total number of email opens is 1.04 M, this bringing us to an open rate of 29.49%. The number of unique opens is 223 K (thousands) emails, meaning we have an unique open rate of 6.3% thus on average each of these opened emails was opened about four to five times by the same recipient.

In order to proceed with our research we have extracted the sample that contains above average email subject lines when reporting ourselves to the open rate. The number of remaining emails that had an open rate higher than 29.49% was 161 from 263.

Out of these 161 email we had to eliminate a few outliers: the first initial highest open rate was belonging to the token authentication email for order details. This emails had a 4265% open rate but accounted for only 2070 delivered emails. We also had to delete one more entry which was a dynamic content subject line for which we could not see the line itself as the reporting tool only provided the programming code for the content.

These 159 resulting emails with an open rate above the 29.49% average had a total number of 858 K emails received out of that 3.54 million total. This means that from the total sent communications, only 24.5% had an open rate above the total average.

The first four ranking emails have open rates of over 300%, the highest having 369.7% and the fourth having 321.83%. These four emails account for about 12 K email sends. Looking into more detail at this 4 remarkable communications we notice:

- Their subject line has between 3 and 6 words;
- Their subject line has between 23 and 36 characters (including spaces);
- All of them are plain statements, none of them being questions of exclamations;
- One of them is using the name of the company and service provided;
- The most successful one uses both the name of the company and a sales incentive of more gigabytes on your offer, doing so in 6 words and 35 characters;
- The second most successful is an out of contract notification.

When comparing these top performers with other studies we can notice that although in the case of the university case study where the university name was used in the subject line yielded poor results, in the case of a business to business communication it can be otherwise. As a matter of fact a number of 51 subject lines out of 159 have the company name part of them.

Analyzing some more into these subject lines we see that the average number of words is 7 with a minimum of 3 and a maximum of 14. The average number of characters within a subject line was 48 with a minimum of 21 and a maximum of 94 (all including spaces).

We will be summarizing the initial sample in the table no. 1.

Table no. 1 Overall research sample

Number of email assets	Number of sent emails	Number of received emails	Total Opens	Unique Opens	Open Rate	Unique Open rate
253	4.1 M	3.54 M	1.04 M	223K	29.49%	6.3%

Source: Personal interpretation

The resulting research sample was significantly smaller after filtered for an above average open rate and the main metrics on it are summarized in the table no. 2.

Table no. 2 Filtered research sample

Number of remaining email assets	Number of sent emails	Number of received emails	Total Opens	Unique Opens
159	962 K	858K	738K	242K

Source: Personal interpretation

An important think to notice is the high deliverability rate in the case of these high performing emails. We can see that 858 K email were received from 962 K being sent. This take the deliverability rate in the 90% area.
The main format taken away in matter of size is that short subject lines with up to 50 characters are the most successful (Table no. 3).

Table no. 3 Subject line size

Average subject line character count	Average subject line word count	Variation in subject line character count	Variation in subject line word count
7	48	From 21 to 94	From 3 to 14

Source: Personal interpretation

Conclusions
We believe this paper sets ground for solid subject line format and open rate dependency research. This initial large sample of over 858k over performing sent emails showed us that common traits of above average subject lines are having a subject line of about 7 words and up to 48 characters. They span from 3 to 14 words and from 21 to 94 characters.
The best performing subject lines from our research have between 3 and 6 words and up to 36 characters. This is all obviously in the context of this research, done on a telecom company. In comparison to previous forum and academic studies, using the name of the company has yielded positive results regarding the open rate of the emails. Almost a third of these high performing subject lines had the name of the company or the name of the service, in this case the name of the service also containing the company name. The company name was also present in two out of the four most successful.
By looking at these results we can see that many times results obtained from practice can be different from results obtained from theoretical studies with case studies taking place under controlled academic circumstances. The use of the company name is one example while another one is the use of a blank line. None of these subject lines involved in the study had a

blank subject line and the reason is practical: they would have never made it to an inbox in real life due to internet service provider filter.

There are many interesting topics to look into for further research on this topic. I find of particular interest the use of personalization in subject lines and its impact open rates. The research could also be extended to clicks within the email or with a research focusing on opened emails but not clicked in comparison to the subject lines of the clicked emails. Other similarly interesting topics would be to create a text processing algorithm that could be "fed" with performant subject lines which could then provide an assessment regarding the open rate of a user input subject line. The mapping of the strength of association between subject line keywords and the close examination of re-occurring keywords can be interesting areas to explore with the co-word analysis technique used many times for social media posts.

Aknowledgements

References

Baggott, C., 2007. *Email Marketing By the Numbers: How to Use the World's Greatest Marketin.* Hoboken. New Jersey: John Wiley & Sons, Inc.

Burke, M., Joyce, E., Kim, T. and Anand, V., 2007. Introductions and requests:Rhetorical strategies that elicit response in online communities. *Proceedings of the thirdcommunities and technologies conference*, London, UK. Michigan State University.

Callegaro, M., Krusse, Y. and Thomas, M., 2009. The effect of e-mail invitationcustomization on survey completion rates in an internet panel: A meta-analysis of 10public affairs surveys. *Proceedings of the AAPOR-JSM conference.* Hollywood, Florida, USA. American Statistical Association.

Dawe, K., 2015. Best practice in business-to-business email. *Journal of Direct, Data and Digital Marketing Practice,* 16, pp.242-247.

Porter, S. and Whitcomb, M., 2005. E-mail subject lines and their effect on web surveyviewing and response. *Social Science Computer Review,* 23, pp.380-387.

Sappleton, N. and Lourenço, F., 2016. Email subject lines and response rates to invitations to participate in a web survey and a face-to-face interview: the sound of silence. *International Journal of Social Research Methodology*, 19(5), pp.611–622.

Skogs, J., 2013. Subject line preferences and other factors contributing to coherence andinteraction in student discussion forums. *Computers & Education,* 60, pp.172-182.

The Radicati Group, 2019. *Reports.* [pdf] Available at: <https://www.radicati.com/wp/wp-content/uploads/2015/02/Email-Statistics-Report-2015-2019-Executive-Summary.pdf> [Accesat 17 April 2020].

Wainer, J., Dabbish, L. and Kraut, R., 2011. Should I open this email?: inbox-level cues, curiosity and attention to email. In: *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*. [online] the 2011 annual conference. Vancouver, BC, Canada: ACM Press.p.3439. Available at: <http://dl.acm.org/citation.cfm?doid=1978942.1979456> [Accessed 20 Apr. 2020].