

テキストマイニングによるブログに対する意識調査の分析

Analysis of Consciousness Investigation of Weblog by Text Mining

高橋 朋子*¹
Tomoko TAKAHASHI

高橋 倫子*²
Michiko TAKAHASHI

*¹ 園田学園女子大学 情報コミュニケーション学科
Department of Information Communication Studies, Sonoda Women's University

*² 関西学院大学 社会学部
School of Sociology, Kwansei Gakuin University

あらまし：ウェブログ(以下、ブログ)が、最近よく利用されているが、ブログは従来のコミュニケーションと異なるのか、また、従来とは異なるコミュニティが形成されるのかなどは興味深い。そこで、ブログに対してどのようなイメージを持っているのか3大学の学生を対象にイメージ調査を行った。イメージ調査の自由記述の内容をテキストマイニング・ソリューションを用いて分析し、ブログをどのように捉えているのか考察する。また、ブログを書いている人と読むだけの人を比較し、コミュニケーション手段としてのブログに対する意識について考察する。

キーワード: ブログ, テキストマイニング, 意識調査, コミュニケーション

1. はじめに

近年、インターネットは急速に普及し、情報通信白書によれば2004年には世帯普及率が88%を超えた。現在では、多くのWebページが情報検索などに利用されている。一方で、大手ネットワーク系企業が次々とブログ作成サービスに参入し、ブログの利用者が一気に増加した。ブログは従来のコミュニケーションとは異なるのか、さらには、従来とは異なるコミュニティが形成されるのかなどは興味深いことである。

本研究では、ブログのコミュニケーションの形成に関して、調査、研究を行うことを目的とする。まず、3大学の学生を対象とし、ブログに対してどのようなイメージを持っているのか意識調査を行った。

本稿では、自由記述に見られる内容をジャストシステム¹⁾のテキストマイニング・ソリューションを用いて分析した結果を報告する。また、ブログを書く人と、ブログを読むだけの人を比較し、コミュニケーション手段としてのブログに対する意識について考察する。

2. インターネット上のコミュニケーション

川浦(2004)は、インターネットにおけるコミュニケーションをWBC(Web-Based Communication)と、ABC(Article-Based Communication)の二つの型に分類している²⁾。Webページに基づくコミュニケーションをWBC、電子掲示板等のコミュニケーションはABCに分類し、ブログはWBCに含まれる。川上(2005)は、WBC

とABCの違いを表1に示す3点であるとしている³⁾。

表1 2種類のインターネット・コミュニケーション(川浦, 2005)

	コミュニケーション		
	型	単位	機能
WBC	呈示	発信主体	発信ないし自己表現
ABC	相互作用	メッセージ	意見や情報の交換

WBCとABCの違いを次のように説明している。

第一に、コミュニケーションの型に違いがある。WBCは「呈示」もしくは「開示」型である。一方、ABCは「相互作用」に基づくコミュニケーションであり、特定、不特定にかかわらず、何らかの相互作用が存在する。

第二に、コミュニケーションの単位に違いがある。WBCは、ページ全体が発信主体の発現としての単位となるのに対し、ABCは一つのメッセージが単位となる。

第三は、コミュニケーション機能の違いがある。WBCは、発信や自己表現が主流となるのに対し、ABCは、主に意見や情報交換という機能が一般的である。

つまり、ABCは書き手と読み手が共通して作り上げている発言の場であり、「公共の場」と捉えられる。一方、WBCは、相対的に「私的な場」ということになる。

以上のことを踏まえると、インターネットコミュニケーション形態は図1のように表現できる。

また、ブログによる自己表現は、トラックバックやコメント機能により、読まれることを意識したコミュニケーション手段となってきたと指摘している。

以上のことから，筆者らは，ブログは図1のWBCからABCへ移行していると考える．

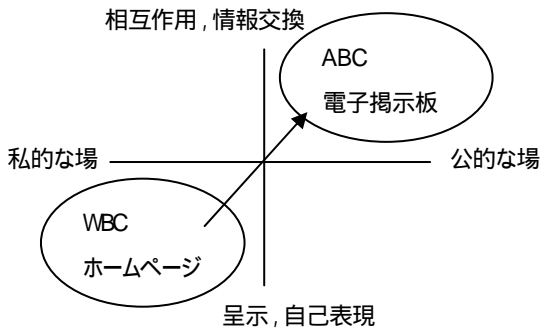


図1 インターネットのコミュニケーション形態

3. イメージ調査分析

2005年11月～12月に，私立の3大学で無記名によるアンケート調査を行った．回答者数，は表2に示すとおりである．なお，A大学は情報系，B大学は文系，C大学は情報系および文系学部の学生である．

表2 調査対象者数

	A大学	B大学	C大学	合計
男	49	13	0	62
女	20	62	69	151
合計	69	75	69	213

同アンケート調査において，「ブログという言葉を知ったことがあるか」の問いに対し，96.7%があると回答している．また，「ブログをどのように利用しているか」の問いに対しては，「他人のブログを読むだけ(以下，読む人)」が55%，「他人のブログも読むし，自分のブログも書く(以下，書く人)」が23%，「ブログを読んだことも書いたこともない(以下，利用しない人)」が22%であった．

以上から，ブログは大学生にはすでに認知されている言葉である．また，ブログを書く人，読む人をあわせて利用している人は，全回答者の約75%であり，ブログは大学生に浸透しているコミュニケーションツールの一つであるといえる．

ブログを知ったことがあると答えた人に対し，どのようなイメージを持っているかを自由記述形式で聞いた．有効回答数および総単語数は，表3に示す．

自由記述に見られる単語数は，読む人は一人あたり平均3.4語，書く人は4.0語，利用しない人は2.8語であった．

表3 有効回答数および総単語数

	全体	読む人	書く人	利用しない人
有効回答数(人)	172	98	42	32
無記入(人)	34	16	5	13
総単語数(語)	587	329	169	89
平均単語数(語)	3.4	3.4	4.0	2.8

次に，総単語の中から意味を持つ単語を抽出(以下，抽出単語という)した．抽出単語数は，書く人は118語，読む人は211語，利用しない人は59語，全体で計388語である．さらに，抽出単語を似た意味をもつ単語同士でグループ化を行い，次の14のカテゴリーに分類した．()の中に代表語句を示す．

- Web (ホームページ, Web, 更新, サイト, ツール)
- 公開 (公開, 開設, 開設者, 開示性)
- 日記 (日記, ブログ)
- 問題点 (悪用, 中傷, 侵害, 出会い系, 匿名性,)
- 自分 (自分, 自己)
- 他人 (人, 他人, 皆, 相手)
- 個人 (個人, 内輪, 興味, 個人的, 趣味)
- 交流の場 (場, 交換, 交流, 友達, 友人, 共有)
- 自己主張 (意見, 表現, 考え, 自己主張, 発言)
- 自己満足 (自己満足, 世界, 嬉しがり, 好き勝手)
- プラスイメージ (自由, 楽しい, 面白い, 良い)
- マイナスイメージ (面白くない, 読みにくい, オタク)
- 作成手軽 (簡単, 気軽, 手軽, 便利, 気楽)
- 作成面倒 (分からない, 大変, 難しい, 面倒くさい)

各カテゴリーにおける抽出単語数を，図2に示す．

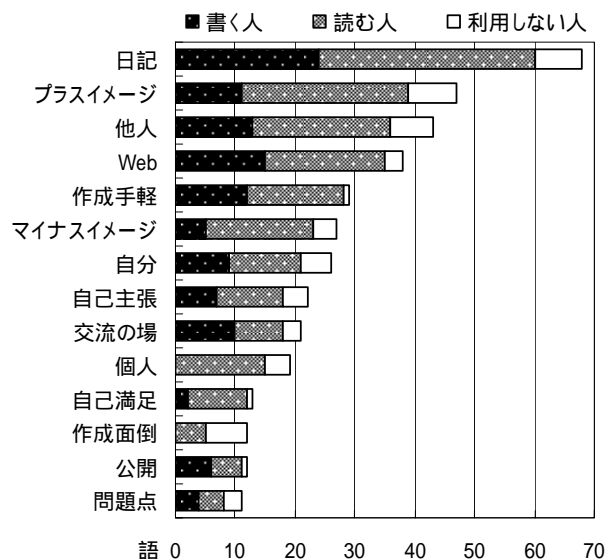


図2 各カテゴリーにおける抽出単語数

4. テキストマイニングによる分析

本分析では，ジャストシステムのテキストマイニング・ソリューションの主題分析を利用する．主題分析

は、テキスト情報から抽出されたキーワードを主題ごとに分類し、主題同士の関係や類似性を明確にすることができる。内容の関連する意見が自動的にグループ化され、クラスタ名として表記される。本文に同じ語句が含まれない場合も、文章の内容が似ていれば同じクラスタに分類される。表4に主題分析一覧を示す。

表4 主題分析一覧

クラスタ名	代表語句
日記01	日記
なし	なし
開設	開設/ホームページ/開設者/日記
意見	意見/参考/他/人/自分
自己満足	自己満足/自己/世界/更新/日記
個人	個人/個人日記/趣味/世界
自分	自分/興味/今日/代/後半
他人	他人/紙一重/行為/物/中傷
人	人/生活/内容/半面/プライベート
自己主張	自己主張/主張/自己/匿名
場	場/表現/交流/反応/方法
web	web/技能/サイト/作成
情報	情報/オタク/友だち/友人/相手
日記02	日記/公開/公開日記/hp/人

分類されたクラスタ数は14項目であり、代表語句は文章の中から抽出した単語である。対象は、無記入を除く172人の文書である。主題分析における各クラスタ文書数を図3に示す。

表4の主題分析一覧にある日記01は単なる日記であり、日記02はwebや公開などインターネットと関連している日記を指している。

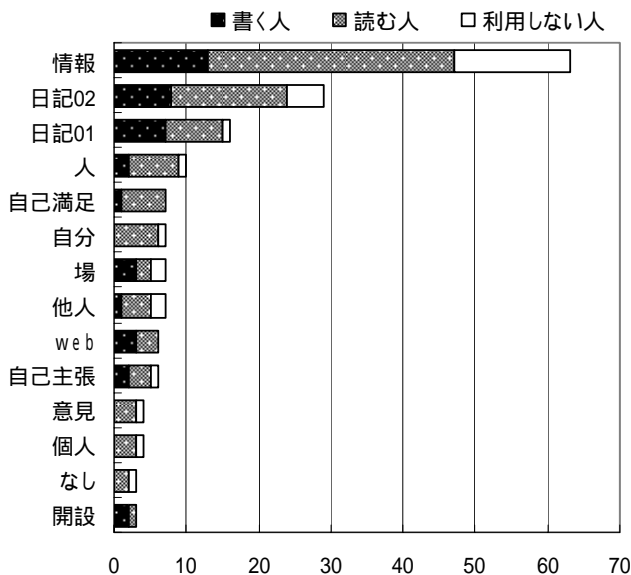


図3 各クラスタ内における文書数

情報に分類される文書数は、面白い、楽しい、陰キャラ、といったブログに対する個人の感性を含むものが分類される。

読む人と、書く人、利用していない人におけるブログのイメージ詳細に分析するため、文書数は少ないが、対比する2つのクラスタを抽出し比較する。

比較する項目は、次の3項目である。

自己満足と自己主張

場と個人

場と自己主張

自己満足と自己主張に分類される文書数を比較したグラフを図4に示す。全体的に見ると、ブログは自己満足と答える人が自己主張と答える人よりも若干多い。しかし、書く人は、自己満足よりも自己主張と答えている人が多く、逆に読む人は自己満足と答える人が多い。このことから、読む人は、書く人よりもブログ世界を客観的に捉えているといえる。

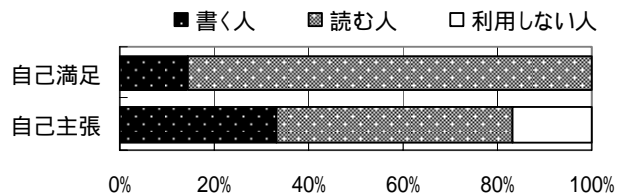


図4 自己満足と自己主張

次に、場と個人に分類される文書数を比較したグラフを図5に示す。全体的に見ると、ブログは交流の場と捉えている人の割合が高い。書く人は、ブログをコミュニケーションを含む交流の場であると捉えている。一方、読む人は、ブログを個人的な世界と捉えている割合が高い。読み手と書き手に若干差がある。

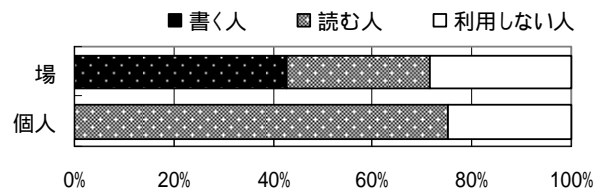


図5 場と個人

次に、コミュニケーションの機能の違いについて、場と自己主張を比較したグラフを図6に示す。読む人は交流の場よりも自己主張と捉えている人が多く、逆に、書く人は、交流の場が主流であると捉えている人が若干多い。

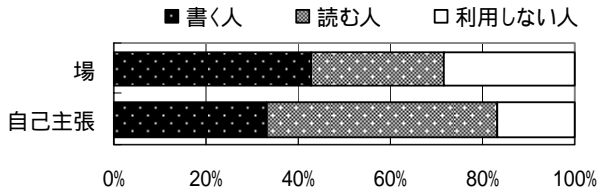


図6 場と自己主張

以上のようにテキストマイニングの分析から、ブログを読む人と書く人のイメージは、以下のような差異があることがわかった。

ブログを読む人は書く人より、ブログの世界を自己満足の世界であると否定的に捉えている割合が高く、客観的に捉えている。

ブログを書く人は、ブログをなんらかのコミュニケーションを含む交流の場、すなわち「公的な空間」であると捉えているが、読む人は、個人的すなわち「私的な空間」と捉えている割合の方が高く、ブログを限られた内輪の狭い世界であると捉えている。

コミュニケーションの機能については、読む人は、発信や情報の呈示といった「自己表現」を意識しているのに対し、書く人は意見や情報の交換、交流といった「コミュニケーション」を意識している。

以上の調査結果は、図7、図8のように表せる。

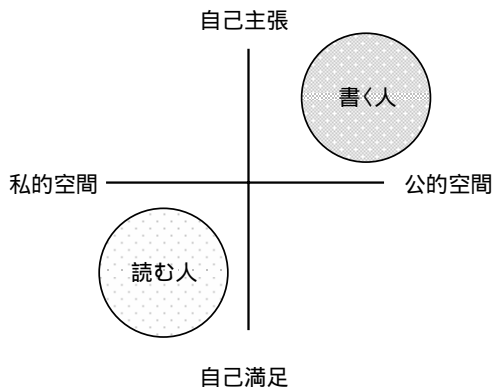


図7 ブログのイメージ調査結果1

図7は、横軸に私的か公的か、縦軸には自己主張の場であるか自己満足の場であるかをとった。ブログのイメージは、書く人には公的空間 - 自己主張型であり、読む人には私的空間 - 自己満足型であるといえる。

図8は、横軸に私的か公的か、縦軸にはコミュニケーションの機能である交流・相互作用型であるか呈示・自己表現型であるかをとった。ブログのイメージは、書く人は、公的空間 - 交流・相互作用型であり、

読む人は私的空間 - 呈示・自己表現型であるといえる。

これらの結果から、図1のコミュニケーションの型に従えば、ブログはWBC型、ABC型双方の特徴を備えたコミュニケーションといえる。

さらに、ブログの書き手はABC型のコミュニケーション機能を意識し、読み手はWBC型のコミュニケーション機能を意識していることが明らかになった。

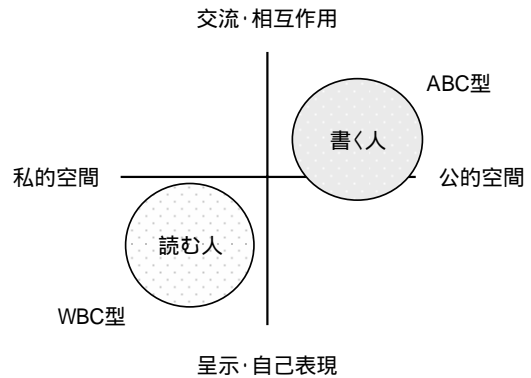


図8 ブログのイメージ調査結果2

5. おわりに

調査結果から、ブログを読む人と書く人では、イメージに差異が見られた。今回の意識調査は限られた大学生に対する調査であるが、ブログが新しいコミュニケーションのツールとして認知されていることは確かである。しかしながら、ブログの書き手と読み手の意識が、異なっていることはコミュニケーションのツールとして、まだ、社会に十分浸透したとはいえない。

今後の課題として、利用者がブログのどのようなことを評価しているのか「評価分析」を用いて分析する。

謝辞

テキストマイニングの分析システム利用に際して、ご配慮いただいた特定非営利活動法人学習開発研究所(代表:西之園晴夫教授)の先生方に感謝します。

参考文献

- 1) ジャストシステム CB Market Intelligence
<http://www.justsystem.co.jp/km/cbmi/>
- 2) 川上善郎:「インターネット時代のコミュニケーション」, 2005 (山下清美, 川浦康至, 川上善郎, 三浦麻子:『ウェブログの心理学』, NTT出版, 2005の1章所収)
- 3) 山下清美, 川浦康至, 川上善郎, 三浦麻子:『ウェブログの心理学』, NTT出版, 2005