

現代社会におけるバイセクシャル

ウェブログのテキストマイニング分析

関川巧真(和光大学)

問題

性的少数者、またはセクシャルマイノリティは近年日本国内でも注目が集まりつつあり、東京の渋谷区で2015年4月に同性パートナーシップ条例が可決され、世田谷区では2015年7月に同性パートナーシップ要綱が登場するなど、法的側面においても目立った動きが見られるようになってきている。現代社会において性の多様性が広まりつつあり、電通ダイバーシティラボの「LGBT調査 2015」によれば「レズビアン(女性同性愛者)」「ゲイ(男性同性愛者)」「バイセクシャル(両性愛者)」「トランスジェンダー(性同一障害など)」のセクシャルマイノリティ(以下 LGBT)は全体の7.6%であるという調査結果があり、徐々にではあるが認知も広がりを見せている。

また、LGBTの人々(以下 LGBT 層)による市場規模は5.94兆円にもなり、経済規模の側面においても無視できないものとなり、NTTドコモ、ソフトバンク、KDDIの電話会社も携帯電話向け家族割引にLGBTを含むようになり、賃貸業界でもLGBT不動産といった同性カップルを家族として契約を行うことが可能な企業が増えている。

しかし多方面の広がりを見せているとはいえマイノリティに違いはなく、いまだ理解は浸透し始めといったところだと言える。LGBTという言葉も日本においてはテレビで最初に使用されたのが2014年5月10日放送、NHK総合「週刊ニュース深読み」という番組にて使用されたのが初めだとされている。

本研究の分析対象にはインターネットで利用可能な「Ameba」にある通称ブログと呼ばれるウェブログ(以下ブログ)のひとつである「アメブロ」を使用する。LGBTは他者からの判別は困難、もしくは不可能であり、自己申告によって判断せざるをえないため、プロフィールの検索機能によりバイセクシャルを明記しているものを対象とした。

LGBTは精神病などの病気ではないものの、世間の認知や思想、宗教によって社会的弱者になる可能性を孕んでいる。しかしながらLGBT当事者の人間関係や自身への関心によって種々様々な人間性が生まれるのではないだろうか。本研究では、ブログをテキスト・マイニングすることにより、「現代社会のバイセクシャルの実態」を探ろうと考えている。

目的

ブログでの書き込みによる実際の LGBT 当事者における恋愛についての考えを分析する。LGBT は病気でもなく、当事者のアイデンティティや社会的立場など複雑に他の領域にも関係しており、自認をしていても他者からの認定などが基本的には難しく、自身が本当に LGBT であるという確証を得難いため、ポジティブな発言とネガティブな発言に分類しがたいものがあるだろう。しかし、自身が LGBT であることをアイデンティティに取り込むことや、LGBT によるコミュニティ、多様な恋愛観によって多くの人間関係を築きあげることも可能であると考え。ブログによる筆記開示やコミュニティの生成が、LGBT の自認とポジティブな交友関係の手助けになるだろう。

方法

1. 分析対象

今回は、「アメーバブログ」でプロフィールにバイセクシャルと明記しているものを分析の対象とした。

2. 分析手順

アメーバブログでプロフィール検索をし、バイセクシャルと明記しているブログから「恋愛」「LGBT」とジャンル分けされたものをタブ区切りテキストにして Excel ファイルにしたものを「Text Mining Studio Ver.5.1」で分析した。

はじめに、Ameba でアメーバブログのウェブページを開き、検索機能によってバイセクシャルを検索し、プロフィールに明記されているものの中から無作為に対象を決める。次に一般公開されている項目の中から恋愛や LGBT と名のついたものを選び、その項目内の最古のブログから検索時最新のものまでを「Word2013」にコピーアンドペーストした。その後添付画像や不要な段落の部分を取り除き、誤字や脱字、文字化け、乱丁の部分修正した。さらに、個人情報となるメールアドレスや電話番号、住所の部分を取り除いた。そして、文章のタイトルと年月日。本文の間に改行を加え、改行によって区切られていた部分をタブに置き換えてタブ区切りテキストにした。それらの文章を書式なしで保存し、「Microsoft Excel 97-2003 ワークシート」でファイルを開き 1 行と 1 列空けて 2 行目 2 列目にデータ表示した。そして、空けた 1 行目の A には「id」、B には「性」、C には「a」、D には「t」、E には「d」、F には「text」と入力した。また、A の列 2 行目からは「1」、「2」、「3」とデータの ID 番号となる数字をすべてに入力し保存した。

「Microsoft Excel 97-2003 ワークシート」の形式で保存した前述のファイルを「Text Mining Studio Ver.5.1」で読み込んで、テキストの基本情報、単語頻度分析、注目語情報分析、特徴語抽出、特徴表現抽出、評判抽出の順に行った。その際、単語頻度分析では上位 30 を抽出、注目語情報分析において、注目語設定タブで「注目することば」の単語を「好き」にし、共起抽出設定タブで出現回数が 5 回以上の共起ルールを抽出した。特徴語抽出では抽出対象とする属性に「性」にし、特徴表現抽出では抽出対象とする属性に「性」にした。なお、これら以外の設定は変更せずに

分析を行った。

結果

1. 基本情報

表 1 はバイセクシャルのブログの基本情報であり、ここでは総行数、平均行数、総文数、平均文長、延べ単語数、単語種別数を示す。まず、総行数は分析対象のブログの件数を表しており、254 項であった。次に、1 項あたりの文字数を表す平均行長(文字数)は268.1 文字であった。このブログの総文数は4746 文、その平均文長(文字数)は14.3 であった。内容語の延べ単語数は24709 個、単語種別数は6201 個だった。

	項目	値
1	総行数	254
2	平均行長(文字数)	268.1
3	総文数	4746
4	平均文長(文字数)	14.3
5	延べ単語数	24709
6	単語種別数	6201

表1 バイセクシャルのブログの基本情報

2. 単語頻度分析

図1はバイセクシャルのブログを単語頻度分析し、上位30の単語を横棒グラフで表したものである。この分析を行うことで、ブログでどの単語が多く用いられているかを明らかにし、バイセクシャルの人の考えを汲み取る。図1を見ると、「人」が最も多く、それに次いで「自分」が多く用いられていた。「自分」以外にも人間関係に関する単語として、「好き」、「一緒」、「付き合う」、「LGBT」、が上位30位内に入っていた。この他に、ポジティブを連想させる「良い」が3番目に多く使われていることが示された。そこでグリッドを表示してより具体的な数値を見たところ、最も多く用いられた「人」は134個であり、次いで「自分」が107個で、「良い」は97個であった。また、恋愛関係におけるポジティブな言葉では「好き」という言葉は女性が68個、男性が7個となっていて、「付き合う」という言葉は女性が35個、男性が4個で、女性が男性よりも大きく上回る回数という数字が出た。恋愛における感情や関係を口にしてポジティブな影響を他者とも共有することで現状の恋愛関係を充実させる効果があると考えられる。

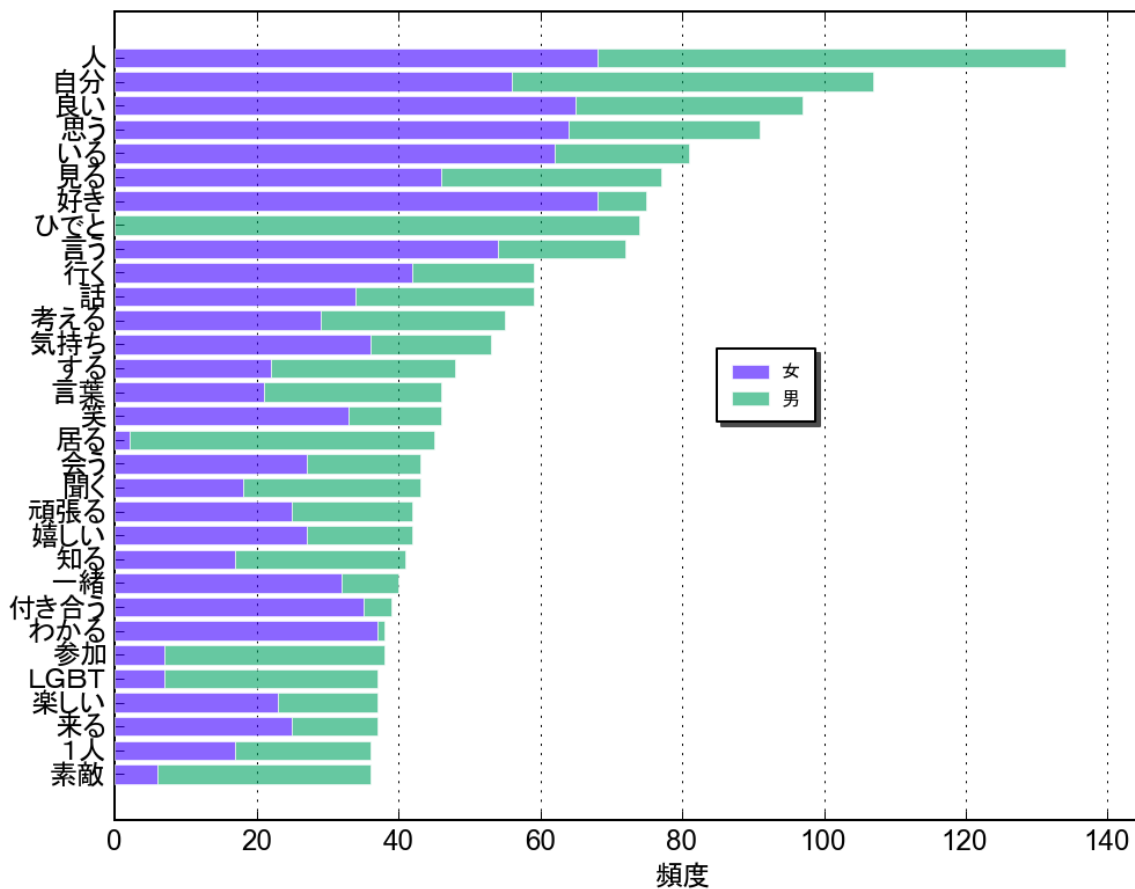


図1 バイセクシャルのブログの単語頻度分析

3. 注目語情報分析

注目語情報分析では、「好き」という単語について分析をする。ブログの中で「好き」という単語が、どのような単語と係り受け関係にあるのかを明らかにし、結果を図2に示した。この分析を行うことで、特定の単語と結ばれている単語群を明らかにし、関係性を探る。図2を見ると、「好き」は「人」、「好き+ない」につながっている。「人」は好きとの関係性から好意をもつ相手のことを示し、「人」、「好き+ない」、「好き+ない」はLGBTの恋愛において交際が始まる前の段階や、交際が終わってしまうだろうと気づく段階でのことを示しているだろう。図2の中にはネガティブな言葉もあり、「こわい」、「好き+ない」、「悩む」、「離れる」、「好き+?」、「淋しい」、「わかる+ない」がこれに当たるだろう。これに対しポジティブな言葉は「幸せ+したい」、「良い」であり、ポジティブな言葉よりもネガティブな言葉が多い。このことから、バイセクシャルの当事者は恋愛に関する発言においてネガティブな表現を使用しやすいと考えられる。

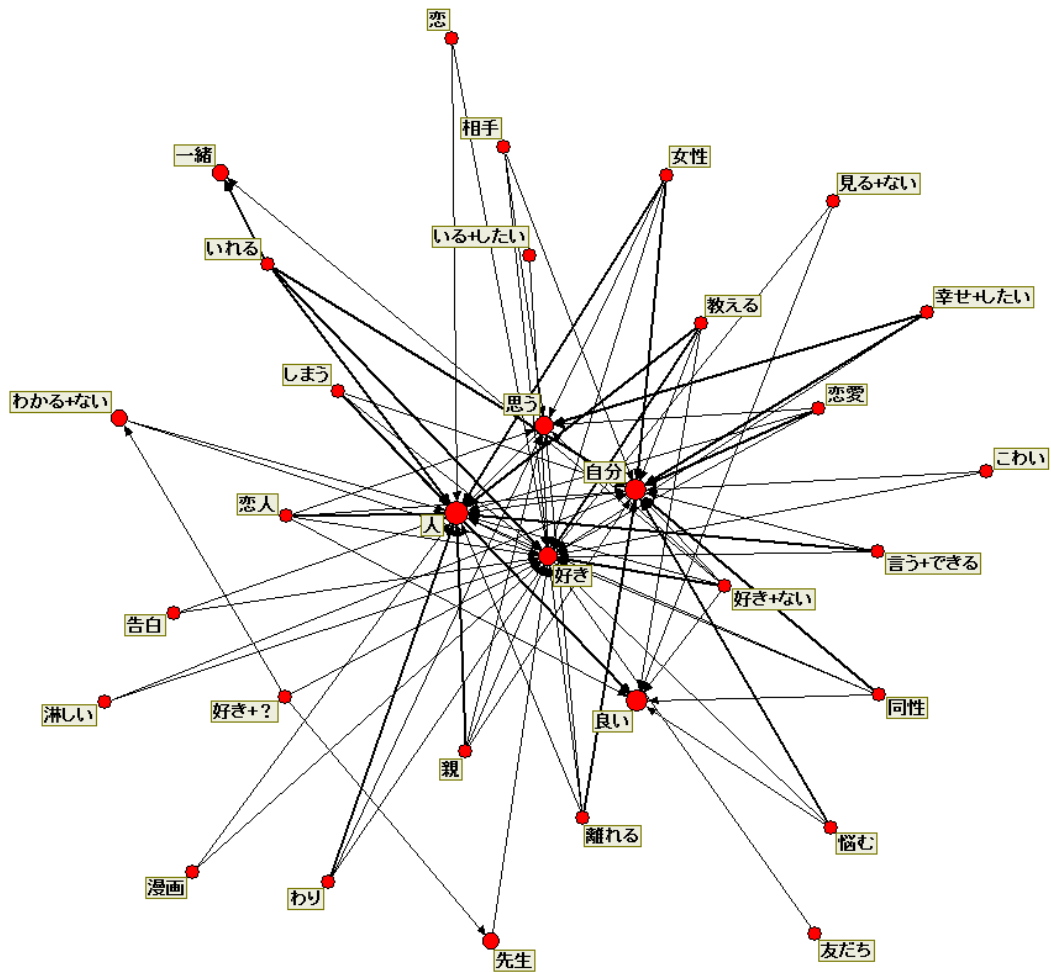


図2 バイセクシャルのブログの注目語情報

4. 特徴語抽出分析

図2と図3はブログ内で使われた単語を性別による属性内で特徴的に出現したものを抽出し、横棒グラフにして表したものである。グラフの横軸の数値は、全体頻度は濃い網のバー、属性頻度は薄い網のバー、指数値を折れ線を表している。また、図2が女性、図3が男性のグラフである。この分析を行うことで、男女別の頻出単語の比較をする。図2の上位3つを見ると、「先生」、「好き」、「おじさん」があり、他者を表す言葉と恋愛関係において重要な「好き」という言葉が出た。図3の上位3つは「ひでと」、「人」、「居る」であり、直接的に関係性が見えるような言葉ではなく、男女で全く違う結果となった。「いる」と「居る」で変換しているかそうでないかの差があるが、その他の語は同じものがなく、性別によってほぼ違う言葉が選ばれている。このことから、同じバイセクシャルであっても性差が存在するのではないかと考えられる。

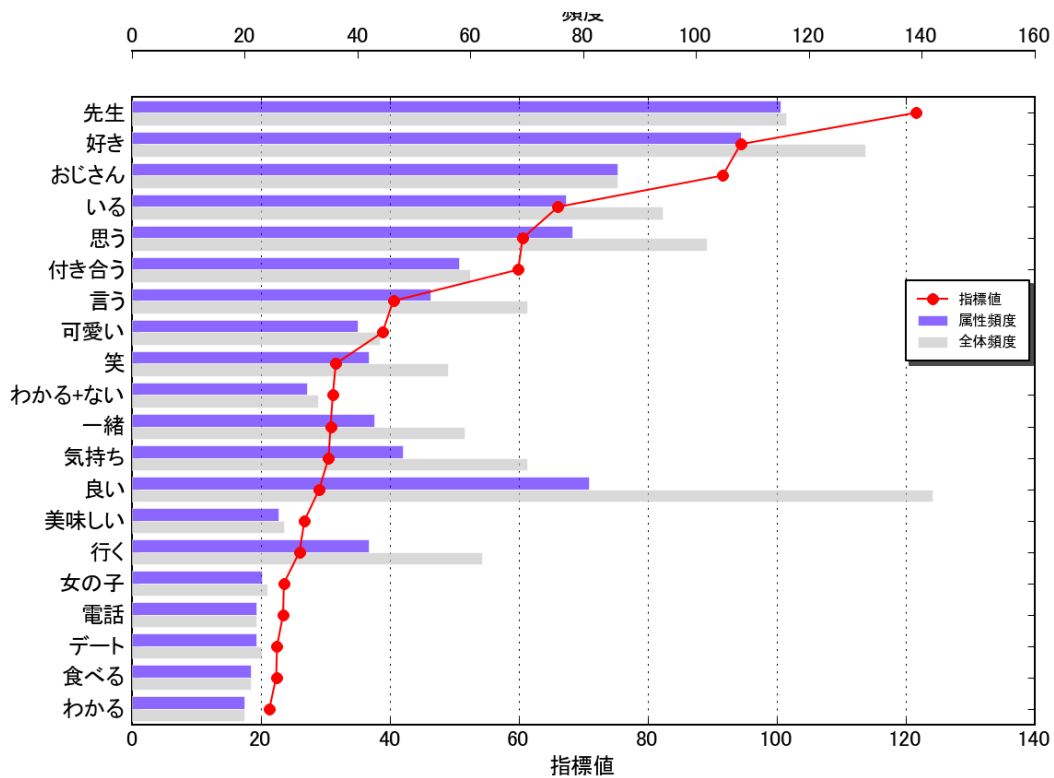


図3 女性のバイセクシャルのブログの特徴語抽出

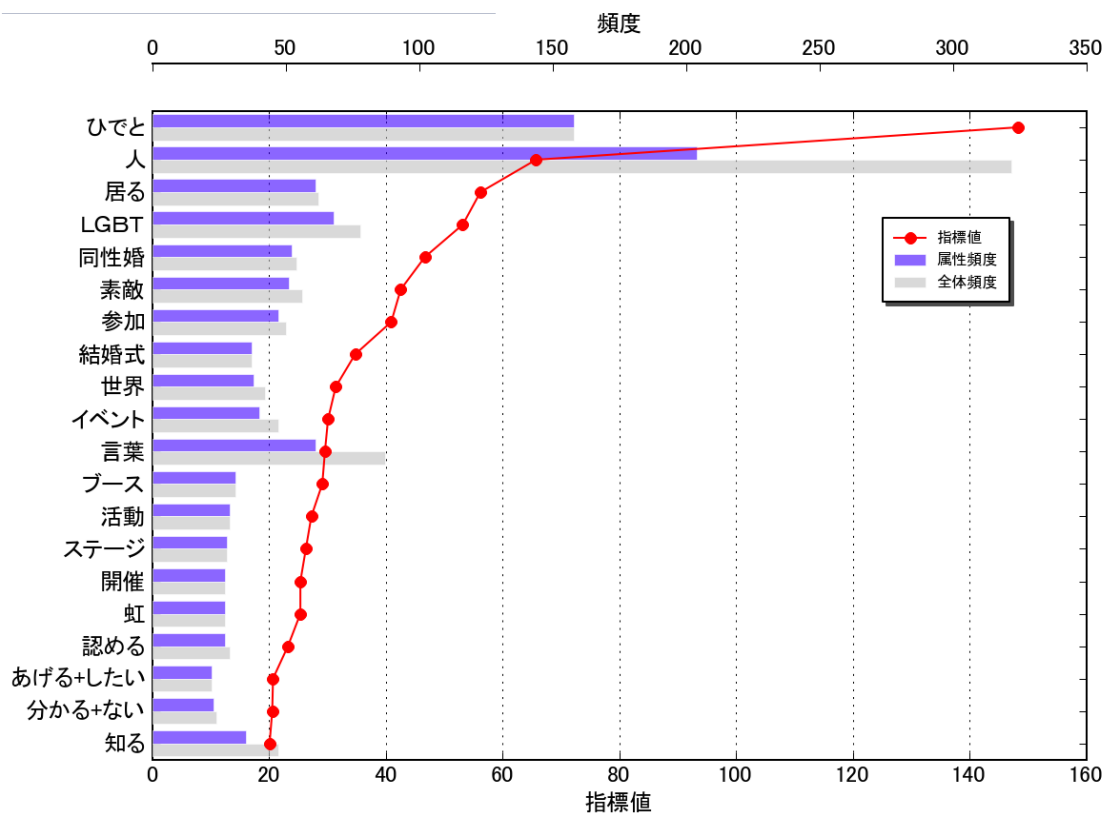


図4 男性のバイセクシャルのブログの特徴語抽出

5. 特徴表現抽出分析

特徴表現抽出では、性別を属性として選択し、特徴的に出現する係り受け表現を抽出し、分析する。グラフの横軸の数值は、全体頻度は濃い網のバー、属性頻度は薄い網のバー、指数値を折れ線を表している。また、図 4 が女性、図 5 が男性のグラフである。図 4 を見ると、「先生－好き」、「好き+?－わかる+ない」、「人－好き」、「先生－大好き」と、「好き」とつながりのある言葉が 4 つあり、直接的な感情表現が多く見られる。しかし、図 5 を見ると「活動－始める」、「学び－癒す」、「参加－いただく+できる」、「実行委員－参加」、「人－集まる」など、活動へのアクションに関する言葉が多い。図 4 では「人－いる」、図 5 では「人－居る」がどちらのグラフでも上位 3 位以内に位置し、この言葉に限って見れば共通して多く使われているが、その他では同じ、または似た言葉が見られなかったことから、同じバイセクシャルであっても性差がするのではないかと考えられる。

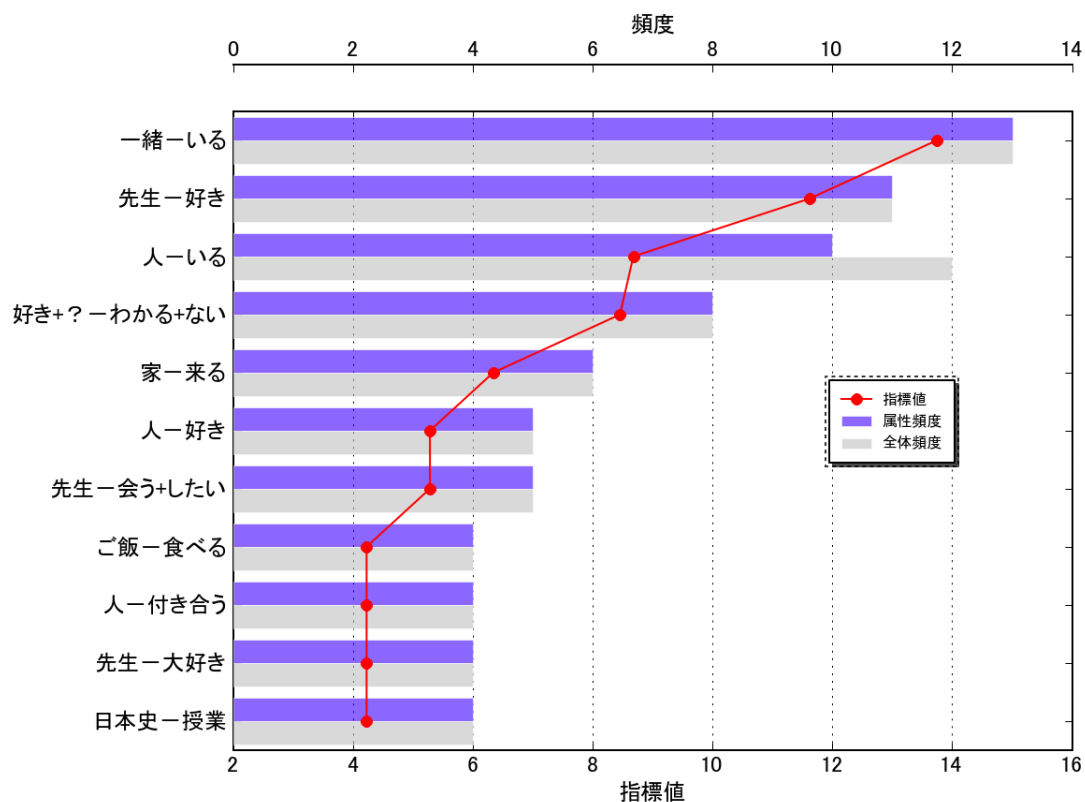


図 5 女性のバイセクシャルのブログの特徴表現抽出

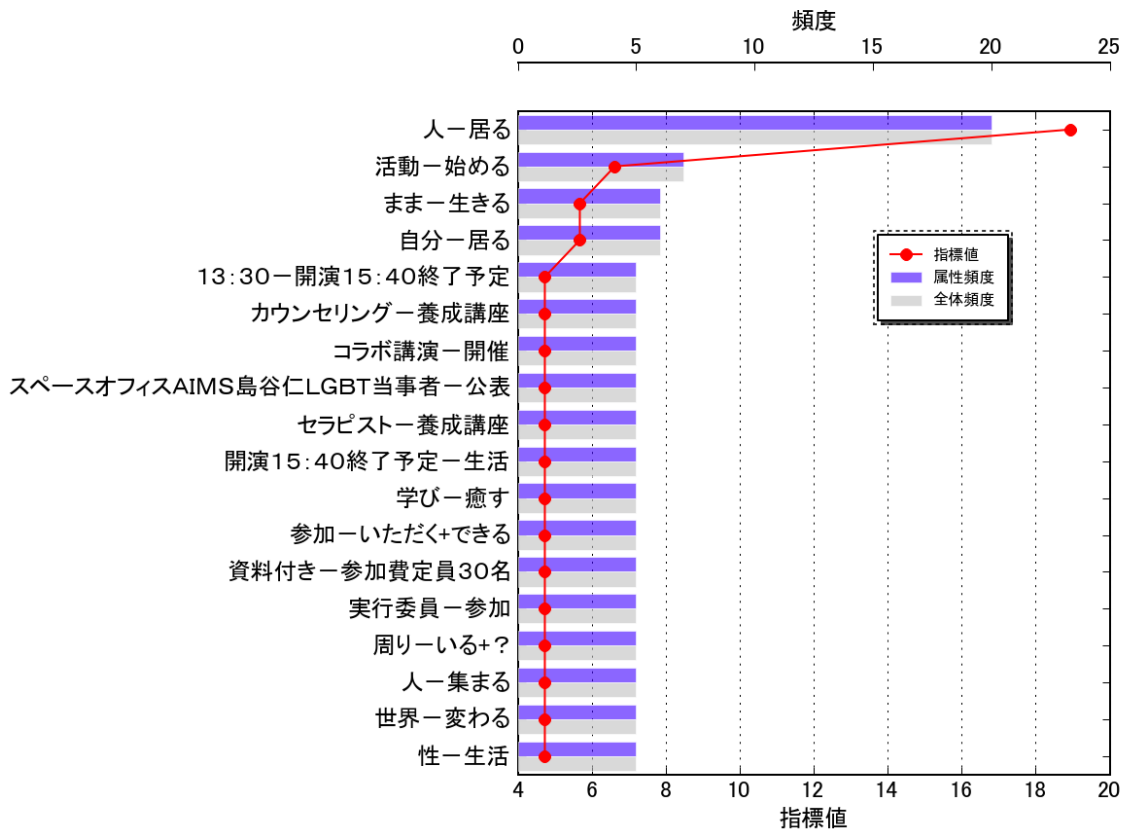


図6 男性のバイセクシャルのブログの特徴表現抽出

6. 評判抽出分析

評判抽出分析では、ブログ内で使われた単語に係り受けする好評語と不評語の頻度を抽出し、単語の評判を分析する。好評語のランキングを示したものが図7であり、不評語のランキングを示したものが図8である。図7を見ると「自分」という言葉が好評と不評の差が少ないが、他はすべて好評の得点が多くなった。図8を見ると、上位5位までは好評の得点と同じか好評の得点が上回ったが、以降は不評語の方が得点が多い言葉がランキングの多くを占めている。しかし、ランキングの多くが不評のみの言葉があり、さらに得点は基本的に少数でとどまっている。このことから、バイセクシャルは不評語よりも好評語で表現することが多く、不評語は意図的に避けているのではないかと考えられる。また、「人」という言葉は図7、図8どちらにおいても好評語として一番になっており、性別に関わりなく共通して好意的に使われている言葉と考えられる。

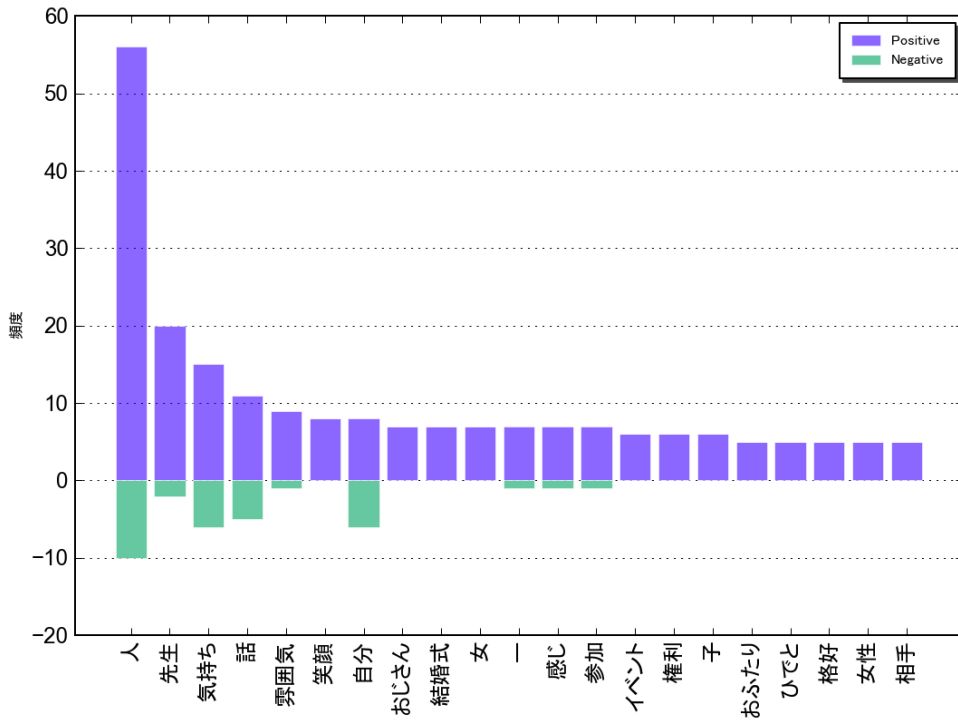


図7 女性のバイセクシャルのブログの評判抽出

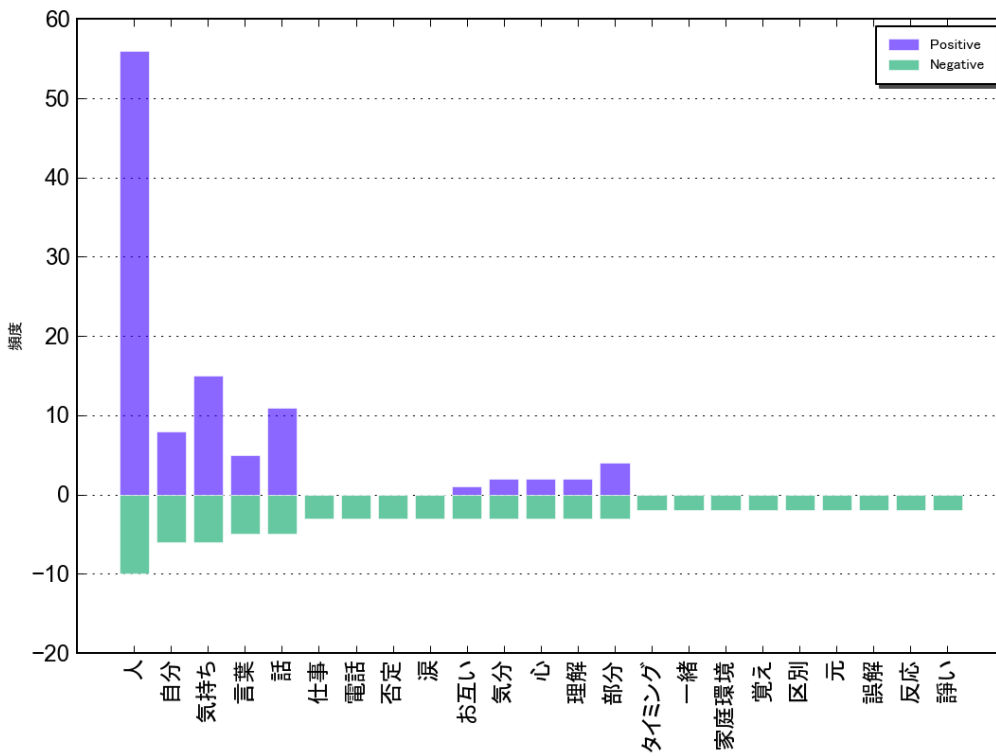


図8 男性のバイセクシャルのブログの評判抽出

考察

1. 『ブログ』での発言に見る恋愛の実態

今回の分析によって、バイセクシャルの当事者たちがブログで発言したものからはポジティブな発言とネガティブな発言の比率が明らかにされた。ブログ中のポジティブな言葉は、単語頻度分析と評判抽出分析によって明らかになった。すなわち、単語頻度分析では「良い」、「好き」が多く、この2つが上位に入っていた。評判抽出分析では好評語が不評語に比べて相当数使用されていた。バイセクシャルの当事者たちはポジティブな発言を多く使用し、自身の環境の良いものを多く捉えようとしていると考えられる。セクシャルマイノリティであっても、ポジティブな視点を持って恋愛関係をもつことで、LGBTは悪いことではないと考えているのではないだろうか。LGBTはいまだに多くの偏見や認識不足があり、LGBTであると自ら他者へ認識させるカミングアウトは簡単にはできないという背景がある。しかし、これらの分析からも分かる通り、バイセクシャルの当事者はネガティブな発言よりも、ポジティブな発言が存在している。LGBTの当事者のポジティブな発言や行動に合わせ、社会が正しい認識を持っていくことで、LGBTが異常ではなく、アイデンティティのひとつとして認知することが実現できるのではないだろうか。

2. 男女のバイセクシャルの表現の特徴

単語頻度分析、特徴語抽出分析、特徴表現抽出分析の結果から、同じバイセクシャルであっても性別によって表現に共通点と差異が見られた。

共通点として単語頻度分析において「人」、「自分」、「考える」、「言葉」という言葉は男女でほぼ同じ頻度で使用されていた。特徴語抽出分析と特徴表現抽出分析では「いる」、「居る」という言葉に共通性が見られた。

性差としては、単語頻度分析において恋愛関係におけるポジティブな言葉である「好き」、「付き合う」という言葉が男性よりも女性が高い頻度で使用していた。また、特徴語抽出分析と特徴表現抽出分析で同じ言葉が出てこないランキングとなり、差異が見られた。これらの分析からバイセクシャルであっても性別の違いによって表現が異なる部分が多いのではないかと考える。

3. 本研究の意義:セクシャルマイノリティであってもすべてが社会的弱者ではない

LGBTである自身を受容し、取り込むことでアイデンティティの一部として強固なものとなるだろう。LGBTは身体的特徴には現れないが、精神病ではなく個人的特徴である。しかしながら日本においてLGBTという存在の認知も広がり始めたところであり、誤解や差別的なものは存在する。柘植(2015)は「セクシャルマイノリティの論文を執筆する研究者においても、正しい理解がされていない。」と述べている。そんな中、LGBT当事者の側からポジティブであり続けることで、現代社会においてセクシャルマイノリティであるLGBTが大手を振って社会に存在することができるだろう。本研究もその手助けのひとつになれると考えられる。

4. 本研究の限界:ブログと実生活との差異

今回の研究では、LGBT 当事者へのインタビューや書籍ではなく、ブログの分析であるため、ブログの筆者によっては誇張表現や故意に秘匿した情報、インターネットという誰にでも観覧可能なものであるために文章への装飾などが含まれ、実際に起きた事象や実際の感想とは違うものである可能性がある。また、LGBT は自身の認知によって成立するものであり、当事者の意識次第では自身が LGBT であることを否定することなどもあり得、全く同じ LGBT の人がいないといえるため、普遍的なデータを得難い点があげられる。しかし、今回の分析によって、ポジティブな言葉や積極的な人間関係を構築する当事者が存在することを明らかにした。また、マイノリティの中の一例として発信することで、LGBT 当事者で社会的弱者に追いやられてしまっている人たちの背中を押すことができれば幸いである。

謝辞

学生研究奨励賞の原稿作成にあたり、「Text Mining Studio Ver.5.1」を使用させて頂きました株式会社 NTT データ数理システム様に感謝いたします。また、本論文を作成するにあたり、指導教員の伊藤武彦教授から丁寧かつ熱心なご指導を賜りましたことに感謝いたします。

文献

柘植道子(2015)『LGBTと心理学』日本における批判心理学の可能性:理論心理学,LGBT心理学,エスニックマイノリティの心理学,平和心理学の立場から 日本心理学会第79回大会発表資料

Wikipedia LGBT <https://ja.wikipedia.org/wiki/LGBT>(2015年10月28日習得)

電通ダイバーシティ・ラボ 2015 LGBT 調査

<http://www.dentsu.co.jp/news/release/2015/0423-004032.html>(2015年10月28日習得)