

平成24年度 卒業論文

テキストマイニングによる
スマートフォンのマーケティング分析
ー売り手の販売戦略と買い手の意識のずれー

専修大学文学部人文学科社会学専攻

指導教員名 金井雅之

LH21-4011D 鈴木 敦史

目次

第1章 研究の目的	1
第2章 先行研究	3
2.1. スマートフォン市場の概況.....	3
2.2. 企業と消費者間のマーケティング関係.....	7
2.3. ビッグデータ解析とテキストマイニング	7
第3章 データと方法.....	11
3.1. 使用データ	11
3.1.1. 売り手のデータ	12
3.1.2. 買い手のデータ	14
3.2. 分析方法	16
第4章 売り手の構成要素比較.....	18
4.1. 上位5形態素の傾向.....	18
4.2. 上位10形態素の対応分析.....	22
第5章 買い手の構成要素比較.....	24
5.1. 上位5形態素の傾向.....	24
5.2. 上位10形態素の対応分析.....	28
第6章 まとめ	30
文献	31

第1章 研究の目的

社会学とは何か。この答えは多様であり、人の数だけ答えがある。ここでは、個人と社会の相互関係に視点を置き、世界的規模で急速に複雑化する現代社会の様々な現象や課題、そして人間の社会的共同生活について研究する学問であると定義する。このように社会学には多種多様な研究が存在し、宗教社会学、家族社会学、都市社会学など様々である。その多種多様な分野の中に、数理社会学¹という分野が存在する。数理社会学が必要である理由には2つあるとされている。一つには、数理に頼らなければ絶対に出来ないことが社会学の中にあるからであり、もう一つには、数理的アプローチが非数理的アプローチに比べて適している科学的営みがあるからである（高坂 1986）。本稿では数理に頼っており、分野を位置づけるのであれば、数理社会学の分野となる。

今も昔もデータ解析の主流は統計解析である。統計解析の基礎にある統計学は長い歴史を持ち、理論体系を構築してきた。そして、この統計学の理論を適用するためには、信頼できる統計データの存在が不可欠であり、例えば標本抽出法²といった技術が用いられてきた。その為、これらの理論や技術を使える人は少数であった。その後、コンピュータの普及につれて統計理論はコンピュータ上に実装され分析は容易になり、ユーザ層が広がってきている。しかし、その反面、インターネット上に溢れる洗練されていない大規模データに対して、統計学の限界も明らかになってきた。インターネットの普及により、大規模データのリアルタイムでの収集、処理、分析が可能になったことで、情報の精度と鮮度が高まっている。例えば、市場や消費者が今まさに求めているニーズにマッチした製品開発、その顧客の潜在ニーズを予測したマーケティング活動が可能となっている。その為、従来から用いられている選択肢型のアンケート調査から、自由な発言内容や記述文によって、消費者の潜在ニーズを把握するテキストマイニングを用いる手法へ移行しつつある（有村 2003, 伊藤 2004, 安藤 2009）。その中でも、特に本稿では、プロダクトライフサイクル³において成長期の段階であるスマートフォンを取り扱うことにした。

¹ 数理社会学とは、計量経済学に対応するもので、一定の数学的仮定から演繹してモデルをつくったり、それを調査や実験から得たデータによって検証したり、あるいは逆に経験的データから数式で表される法則を導いたりする社会学の一分野のことである。

² 主な標本抽出法として、単純無作為抽出法、系統抽出法、多段抽出法、層化抽出法、層別2段階抽出法がある。

³ プロダクトライフサイクルとは、新しい商品が市場に投入されてから、序所に売れ始めて売上高のピークを迎えた後、次第に売れなくなって市場から撤退するまでのプロセスを表したものである。新製品の投入から撤退までのプロセスは一般に導入期、成長期、成熟期、衰退期の四つに分類される（片岡 1997）。

第1章 研究の目的

また、日本の消費者に関する言説をめぐっては、そこに2つの神話が見出される。すなわち、「日本の消費者はタフな消費者である」という神話と、「日本の消費者はイージーな消費者」という神話がある（三浦 1997）。また、タフな面として、製品瑕疵不許容、鮮度信仰、新奇性選考、過剰性選考がある（三浦 1997）。今、それらを満たす商品としてスマートフォンがあり、本稿ではスマートフォンを分析対象とする。

スマートフォン等⁴の普及が消費市場に与える効果を推計すると、直接効果は3.7兆円、産業関連表による波及効果を含めると7.2兆円に達し、半数以上は情報通信産業以外の一般産業に波及する。また雇用創出効果は33.8万人になる（総務省 2012）。このように、スマートフォンには経済波及効果があり、雇用誘発効果もあり、スマートフォンについて分析することは有意義である。

本稿では、これに加えて売り手である企業についても解析している。企業と消費者のニーズ感のズレがどのようになっているかを探り、明らかにすることによって、より消費者のニーズを的確に捉えることが出来るようになると思う。本稿では、インターネット上に溢れる大規模データ、すなわちビッグデータを用いて、売り手である企業が考える「売れる商品」と消費者のニーズは何か、どのような違いがあるのかを探していく。

⁴ スマートフォン等とは、スマートフォン端末だけでなくタブレット端末も含む。

第2章 先行研究

2.1. スマートフォン市場の概況

スマートフォン市場の大まかな流れを把握する為、スマートフォン市場の概況を確認する。総務省の「利用者視点を踏まえた ICT サービスに係る諸問題に関する研究会」における、「スマートフォンを経由した利用者情報の取り扱いに関する WG（第一回）」の会議資料（スマートフォンをめぐる現状と課題）を見ると、スマートフォンの国内における出荷台数は近年急速に増加し、2011 年度上半期には 1000 万台を超えたとされる（総務省 2012c）。また、同報告書によれば「携帯電話端末の国内における年間出荷台数のうち、スマートフォンの占める比率が急速に上昇を続けている」（総務省 2012c）としている（表 1）。

表 1 スマートフォン国内出荷台数と携帯電話出荷台数に占める比率

	台数（万台）	比率（%）
2009 年度上半期	106	5
2009 年度下半期	128	6
2010 年度上半期	223	11
2010 年度下半期	632	30
2011 年度上半期	1004	48

出典：総務省（2012c）

次に、消費者がスマートフォンを選んだ理由はどのようなものかについてみる。総務省はスマートフォンユーザーに対して、携帯電話からスマートフォンに乗り換えた理由について調べている⁵（総務省 2012a）（図 1）。

⁵ スマートフォンユーザーに対して、「携帯電話（フィーチャーフォン）からスマートフォンに乗り換えた理由についてお選びください。スマートフォンが初めて買った携帯電話である場合や、2 台目などで購入した場合は、購入理由についてお選びください。」と質問し、①当てはまるものすべて、②最も決め手になった項目について回答を得たとしている。

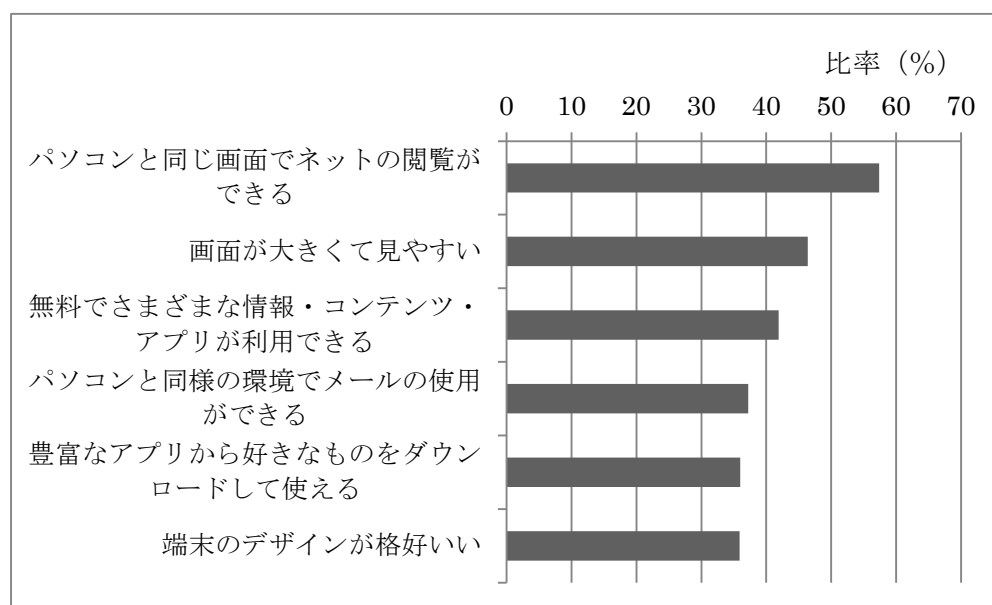


図1 スマートフォン選択理由（当てはまるもの全て）

出典：総務省（2012a）

図1を見ると、「無料でさまざまな情報・コンテンツ・アプリが利用できる」や「豊富なアプリから好きなものをダウンロードして使える」といった機能面より、むしろ、「パソコンと同じ画面でネットの閲覧ができる」や「画面が大きくて見やすい」といった基本的な性能の方に消費者の関心があることが分かる。これは非常に意外なことである。何故なら、日本の製品、サービスにおいて、多機能、多品種、過剰包装などの「過剰性」は至る所で見る事が出来るからである（三浦 1997）。この調査によって消費者が求めるスマートフォン像が浮かび上がってきた。しかしながら、この調査では選択肢型のアンケート調査を用いており、選択肢の設定のない多様な意見を見逃してしまう恐れがある。

一言にスマートフォンといっても様々な機種、様々なメーカーがある。国内スマートフォン市場の契約数のOS別シェアは、2011年3月末時点でグーグル社のAndroidが40%（386万件）、アップル社のiOSが50%（474万件）となっている（総務省 2012c）。AndroidとiOSの契約数を合わせると、スマートフォン市場の9割程を占める（総務省 2012c）（表2）。

表2 契約数のOS別シェア

	台数（万台）	比率（%）
iOS	474	50
Android	386	40
Windows	71	7
その他	25	3

出典：総務省（2012c）

国内スマートフォン市場のおよそ半数を占めるアップル社は、2007 年 7 月に iPhone を発売した。発売当初は、音楽端末 iPod を基にして、ネットワーク機能と通話機能が追加された多機能携帯電話に過ぎなかったが、ユーザインターフェースや UNIX OS、サファリブラウザ等、高い性能と拡張性を持った iPhone として発達し、2011 年に至っては 50%のシェアを占めるに至っている。

アップル社は、OS、ハードウェア、ソフトウェアコンテンツを高レベルで開発し、融合している。そもそもアップルはステップアップ誘導⁶を得意としており、iPod は 2001 に発売し、iPod Touch を 2007 年に投入し、爆発的ヒットした iPhone 3G に繋げた実績がある。Mac、PC 上の iTunes Store を介してモバイル端末をリンクする手法は既存の Mac 販売という市場を維持したまま、新市場の開拓を可能にした。加えて、アップルはメモリ、小型高精度ディスプレイ、バッテリーの世界最大級の需要家という側面を持っており、類似のデバイスの大量投入は部品購入単価の著しい削減をもたらす。このように、アップルは多重らせん構造で進化を続けている。例えば、2009 年度のスマートフォンのエンドユーザ向け販売台数は、アップル社が販売する iOS が 2009 年度の時点で 14%であったが、2011 年になると、19%にシェアを伸ばしている（表 3）。

表 3 スマートフォン市場におけるメーカーとしてのアップル社のシェア

	2009 年		2011 年	
	台数（万台）	比率（%）	台数（万台）	比率（%）
アップル	2413	14%	8963	19%
その他	14824	86%	38211	81%
合計	17237	100%	47174	100%

出典：総務省（2012a）

これに対して、グーグル社は検索からの広告表示が最大の収益源である以上、グーグルが管理できる PC、携帯電話、スマートフォンなどのディスプレイを確保する必要があった。多大な開発費を投入している Android OS⁷を無償で展開することで、多くの通信キャリア・端末メーカーに採用され、結果として世界中の人の手に行き渡ったのである。グーグルは、OS および各種インターフェースを含んだプラットフォーム開発はグーグルが担い、デバイス連携の独自機能などは通信キャリアもしくは端末メーカーが担うという水平分業の特徴を持っている⁸。

⁶ たとえば、iPod と iPhone は両方とも音楽を聴くことができるが、iPhone では携帯としての機能などがある。つまり、従来の商品から、機能の一部が関連する従来の商品より高機能な商品を買わせる戦略のことである。

⁷ 2007 年 11 月に発売された携帯電話用プラットフォームで、オープンソースの Linux をカスタマイズしたものであり、インターネット上において無償で提供されている。

⁸ もともと多くの開発費を必要とする OS 開発は携帯キャリアにとっても端末メーカーにとっても悩みの種であった。際限なく性能向上を繰り返すチップ群や高解像度化が著しいディスプレイ対応、アプリや

Android は 2008 年 10 月に米国の事業者である T-mobile より、台湾 HTC 製の Android 端末 G1 が発売され、日本ではドコモより 2009 年 7 月に発売されている。グーグル社の会社情報によれば、「グーグルの使命は、世界中の情報を整理し、世界中の人々がアクセスできて使えるようにすること」とある。Android の開発目的がグーグルの使命である「世界中の情報を整理すること」、「世界中の人々がアクセスできて使えるようにすること」と考えてみると分かりやすい。

Android を搭載したスマートフォンとグーグル社が提供するサービスが接続されることにより、ユーザは様々な情報を入手できるようになった。多くの発信機が搭載された Android 端末が普及するということは、世界中に発信機を設置したことと同じである。これら発信機から発せられる情報を収集し分析することによって、この「世界中の情報を整理すること」は達成できる。グーグル社の提供するサービスと親和性の高い端末を普及させることにより、グーグル社のサービスにアクセス可能なユーザを増やすことができる。これにより、「世界中の人々がアクセスできて使えるようにすること」が達成できる。

より具体的に述べると、iPhone と Android の違いには大きく二つの違いがある。第一に、アプリケーションの導入環境に違いがある。グーグル社が開発した Android は、アプリケーション実行環境と SDK (Software Development Kit) を無償で提供し、OS、ミドルウェア、アプリケーションフレームワークのソースコードをオープン化するプラットフォームとなっている。その為、Android はアプリケーションを販売しているグーグルプレイ⁹等のマーケット以外からアプリケーションを導入することが出来る。これに対して、iPhone はアップルが認めたアプリしか導入することが出来ない。次に、製造元の違いがある。Android は OS と端末の製造元が異なる為、様々なメーカーが参入し、商品に多様性が出てくる。これに対して、iPhone は OS と端末が同じメーカーである為、多様性が無い。

本稿では、買い手と売り手との求めるスマートフォン意識の違いを見ていくが、これを探っていくにあたって、複数の機種種のデータを用意することが必要である。これに対して、iPhone に限定するとアップル社とユーザとの意識の違いに留まり、アップル内部でしか通用しないということになる。これに対して、Android では販売されている機種が多く、様々なメーカーが参加しており、多様性が高い。

グーグル主導の Android は共存共栄のバランスを上手く保っている。OS、サービス・マーケット、アプリ、デバイスとそれぞれが急速な進化を続けるこの数年ではこの枠組みは固く維持されると考えられる。全てのサービスがクラウドに移行しつつあり、モバイル端

コンテンツの管理技術、GPS やカメラ、各種センサー群とのインターフェース開発、操作性を高める音声認識技術など膨大な技術領域を各社で対応するのは限界であった。その上、ユーザの期待は高く、短期間で実現しなければならなかった。グーグルの水平分業では、通信キャリアにもマーケットプレイスを提供し、グーグル自身が対抗デバイスを作らないという形で端末メーカーにも配慮をしている。

⁹ グーグルプレイとはアプリ、映画、音楽等のコンテンツをオンライン上で管理、ダウンロードすることが出来るサービスのことである。

末が高速かつ安価で接続できることも、グーグルに追い風となるだろう。その為、本稿では Android を分析の対象とする。

2.2. 企業と消費者間のマーケティング関係

今日の消費者行動研究が飛躍的に発展し始めたのは 1960 年代からである¹⁰。その後、激化する企業間のマーケティング競争は、消費者行動研究への期待を増大させてきた（杉本 1999）。1960 年代から 1970 年代にかけての消費生活の実態は、科学技術の著しい進歩により、新しい商品が多く生産され、消費者はこれらの商品に対して今までの経験や知識だけでは対応できなくなってきた。行政はこのような状況から消費者を守るため、法令の新たな制定や改正あるいは行政指導の強化と並行して、消費者に商品の選択に必要な知識などを教えることにより、消費者被害を未然に防ぐ消費者教育を実施することとなった¹¹。

1970 年代までの態度などに見られるように、消費者は企業によるマーケティング・コミュニケーションの受け手として受動的な消費者観が前提としてあったのに対して、1980 年代では消費者は消費目標を達成する為に問題解決すなわち意思決定を行う能動的な存在として捉えられるようになった（杉本 1999）。

一つ目は消費者のニーズを把握することから研究する。二つ目は消費者の反応や選択を通して研究する。すなわち、ニーズを把握する消費者視点の消費者行動研究と反応過程の解明といった企業側視点の消費者行動研究がある（三浦 1993）。本稿では、消費者のニーズを把握する研究法に分類される。

2.3. ビッグデータ解析とテキストマイニング

今、多様な品質のデータが構造化されないまま大量に流通する時代である（岩田 2012）。

消費者の大まかなニーズを把握する為には、大量のデータであるビッグデータ¹²（岩田 2012）が必要となってくる。しかしながら、大規模なデータを一つ一つ丁寧に見ていく方法で有用な情報を取り出すことは難しい。近年、大規模記憶装置としてのデータウェアハ

¹⁰ 初期の消費者行動研究は、企業のマーケティング戦略を構築するために必要とされる消費者行動に関する知識を得ることが主たる目的であり、研究の動機であった。例えば、初期の消費者購買行動に関する概念的統合モデルである Howard & Sheth モデル（1969）は、企業の消費者に対する直接的な働きかけが、購買行動を決定づける情報であることを明確に示している。そして、消費者あるいは顧客の存在を企業活動の中心に据えて、どのような製品を開発し、消費者に流通させるか、あるいは広告や販売促進をいかにして行うかといった政策が立案され、様々なマーケティングが構築されてきた。そのような企業のマーケティング活動を消費者がどのように受け止め、行動するかといった視点から消費者行動が分析されてきた。

¹¹ 行政では、1961 年に経済企画庁の諮問機関として国民生活向上対策審議会が設置され、消費者保護の立場からの取り組み方について検討された。1963 年には同審議会による「消費者保護に関する答申」が出され、「消費者は保護されるべき弱者である」として「一般消費者に対する消費者教育に力を注ぐこと」と述べている。これは国としての消費者教育について、はじめて公的な提言を行ったものである。

¹² ここでのビッグデータの意味は、構造化されないまま大量に流通する多様な品質のデータとしている。

ウスとコンピュータの計算能力の急速な発達に伴って、大量なデータを用いて大量の計算を瞬時に行うことが出来るようになってきている。企業で運用されるシステムには受注、発注、売上、在庫などの多種多様のデータが存在しており、これらは日々の蓄積によって膨大な量のデータとして保存されていく。その中には有益な情報が含まれているが、あまりにも情報が大きすぎることによって有益な情報が引き出されぬまま埋没してしまっている。

これら大量のデータをコンピュータによって解析し、その中に潜んでいるデータ間の相関関係、パターンを発見する技術として発達してきたものがデータマイニングである（岩崎 1999）。データマイニングの「マイニング」とは鉱山から資源を掘り出す採鉱という意味である。大量のデータを鉱山と見立てて、そこから有用な情報資源を採鉱するという意味でデータマイニングという言葉が作られた。データマイニングは、商品の販売データなどを解析して相関関係を可視化し、そこから新しいビジネスチャンスを見つけ出す仕組みとして始まった。一つ事例を挙げると、「おむつを買った人はビールを買う傾向がある」という米国におけるマーケットバスケット分析¹³がある。1990年代前半はデータマイニングという言葉はまだ世間に認知されておらず、技術的にも未熟でビジネスへの適用も理解されていなかった。その中で「おむつとビール」の研究はデータマイニングが注目されるきっかけとなったのである。

データマイニングは1990年代前半に基礎研究が本格化し（鷲尾 2003）¹⁴、データを格納する関係データベース、検索や分析を行うデータウェアハウスなどの技術が用いられてきた。しかし、データは定型化された数値が中心のデータばかりではなく、一律な統計解析では満足な結果が得られなくなってきている。特に、インターネット上に溢れる非定型のテキストデータは、さらに複雑な解析が必要となってくる。こうしたテキストデータには、有益な情報が隠されている。これを得るために、テキストをデータマイニングの手法を応用して解析するのがテキストマイニングである。

テキストマイニングが注目されるようになったのは最近のことである（保田 2003）。売上表などの文書形式や入力数値が定型化されているものが分析対象の中心となるデータマイニングとは異なり、人間が日常で用いる自然文で書かれているテキストをコンピュータに解析することは非常に難しい技術であった。さらに、日本語はアルファベットを用いる欧米の言語とは異なり、単語をスペースで区切る書き方をしないという特徴がある。その為、テキストマイニング初期では欧米の言語を対象にしており（有村 2003）、日本語を分

¹³ マーケットバスケット分析とはデータマイニングの利用法の一つである。主に POS データや EC サイトのトランザクションデータを分析して、一緒に買われる商品の組み合わせを発見する探索的データ分析のことである。尚、流通分野では、小売部門のマーケティングのためのデータマイニング適用が主流であり、POS データを用いた流通全般の業務知識の導出、インストアでの販売用知識の導出、有望顧客に洗い出しなどが行われている。データマイニング技術としては、分類決定木、バスケット分析、重回帰分析、相関解析などが用いられている。

¹⁴ 適用事例として、もっとも多く発表されているのは金融分野である。米国では、1994 年頃から報告されているが、日本でも近年は多くの事例が報告されている。

析対象とした実用的なテキストマイニング技術が確立したのはここ数年のことである（中井 2008）。

現在では、文章から単語抽出する技術も確立し、精度の高いテキストマイニングが行われるようになっていく。今、テキストマイニング技術の活用が進みつつあるのが、Web マーケティングの分野である。例えばインターネット上に存在する掲示板やブログのテキストデータを収集し、そのデータをテキストマイニング処理することによって自社の製品がどのような評価を受けているかが分かる。

テキストマイニングは、金融や流通、製造業や通信サービス業だけでなく様々な分野で用いられ、例えば、美術館の来館者のデータを調べて今後役に立つ分析がある（伊藤 2004）。金融市場についても、和泉ら（2007）が市場を流れる情報をテキストマイニングを通して抽出している。

しかしながら、テキストマイニングという名前が定着し、実用化が進んできたのは、インターネットやコンピュータの普及に伴い電子化テキストが急速に増加した 1990 年代後半になってからである¹⁵。日本語が分析出来るフリーソフトウェアとして、KH coder（樋口 2004）や Tiny Text Miner（松村・三浦 2009）、形態素解析ソフトウェア MeCab¹⁶（工藤 2006）、MeCab を R に組み込んだパッケージ RMeCab（石田 2008）、京都大学黒橋研究室の Juman¹⁷（黒橋・長尾 1998）、ChaSen¹⁸（松本ほか 2007）等がある。形態素解析とは端的に言えば、文章を意味のある単語に区切り、辞書を利用して品詞や内容を判別することである。コンピュータによる自然言語技術の一つである形態素解析は辞書から単語を引いて、与えられた文と照合し、最も自然な単語列を求める。たとえば MeCab では、辞書検索のデータの構造として、TRIE を採用しており、曖昧性の解消として最低コスト法を用いている。

¹⁵ 1980 年代に後半から、自然言語テキストからの情報抽出が自然言語処理と情報検索分野を中心に研究されている。また、1990 年代に入って、機械学習を用いた自動文書分類の研究が、主に情報検索分野で行われている。

¹⁶ 本稿では MeCab0.90 を使用している。MeCab 0.90 は、ユーザが未知語処理を自由に定義することが可能となる。基本的に字種に基づく分かち書きを行い、字種そのものの定義（どの文字コードがどの字種に対応するか）、各字種に対するわかち書きの定義といったことがユーザ自身で定義することが出来る。また、字種を表現するための内部コードに Unicode を用いており、未知語処理の言語非依存性も実現できる。さらに、未知語処理のパラメータは Conditional Random fields(CRF)により推定される為、未知語に対する解析精度の向上が期待できる。

¹⁷ Juman とは OSS である形態素解析器の一つである。Juman 以前の商法的に配布されていた形態素解析器は、辞書や品詞体系、接続規則はほぼ固定されており、ユーザ自身、自由に定義することが出来なかった。Juman はこれらの定義を全て外部に出し、自由な定義が可能になった。

¹⁸ ChaSen とは OSS である形態素解析器の一つである。ChaSen は統計処理によって接続コストや単語生起コストを推定するようになった点にある。これによって、解析ミスを蓄積するだけで自動的にコスト値を推定出来るようになった。また、品詞階層も無制限になり、品詞体系を含めて自由に定義することが出来るようになった。

第2章 先行研究

日本語が分析出来るフリーソフトウェアは様々あるが、**MeCab** は、ユーザ自身が未知語を自由に定義することができる¹⁹。本稿ではスマートフォンという比較的新しいものを対象にしており、形態素解析器が未知語を上手く判別できない可能性があり、ユーザ自身が未知語を自由に定義できる **MeCab** を使用する。**MeCab** には他の分析ソフトウェアにデータを渡す役割も持っている。例えば、**MeCab** を用いて独自の因果ネットワークを構築している分析がある（佐藤・堀田 2006）。本稿では **RMeCab** を用いるが、**RMeCab** を用いた研究として特許情報を可視化した研究がある（安藤 2009）。

¹⁹ ユーザ自身が未知語を自由に定義することができるのは、**MeCab** のバージョンアップ版である **MeCab0.90** 以降であるが、**MeCab** と表記する。

第3章 データと方法

ここでは、売り手と買い手のそれぞれのデータについて見ていく。

3.1. 使用データ

本稿の分析目標は、売り手と買い手の求めるスマートフォンの違いについて明らかにすることである。データは、前者は携帯電話会社が新製品を発表する際のプレスリリース、後者は評価サイトにおけるレビューを用い、分析手法としては対応分析を行うことにする。

分析対象はスマートフォン 23 機種である（表 4）。23 機種を分析の対象とするが、1 つの製造会社が複数の機種を開発していることもある。これは機種名の頭文字のアルファベットで判別することができる。たとえば L-04C は LG Electronics Japan 株式会社, N-04C は NEC（日本電気株式会社）などである。尚、表 4 の公開日とは売り手のデータであるプレスリリースの公開日のことである。また、発売日とはその機種が発売された日のことである。プレスリリースはいずれも発売日の数日前に発表されていることが分かる。

表4 機種別データ一覧

機種名	製造会社名	公開日	発売日
L-04C	LG Electronics Japan	2011/3/4	2011/3/6
N-04C	NEC	2011/3/8	2011/3/15
SO-01C	ソニー ²⁰	2011/3/22	2011/3/24
L-07C	LG Electronics Japan	2011/6/15	2011/6/18
SC-02C	サムスン電子ジャパン	2011/6/16	2011/6/23
N-06C	NEC	2011/6/16	2011/6/24
Bold9780	Research in motion Limited	2011/6/22	2011/6/29
SO-02C	ソニー	2011/6/27	2011/7/9
SH-13C	シャープ	2011/7/28	2011/8/6
F-12C	富士通	2011/8/3	2011/8/7
P-07C	パナソニック	2011/8/10	2011/8/13
SO-03C	ソニー	2011/8/10	2011/8/27
SO-01D	ソニー	2011/10/19	2011/10/26
P-01D	パナソニック	2011/10/31	2011/11/10
T-01D	富士通	2011/11/10	2011/11/18
SC-03D	シャープ	2011/11/21	2011/11/24
F-03D	富士通	2011/11/21	2011/11/25
SH-01D	シャープ	2011/11/25	2011/12/2
SC-04D	サムスン電子ジャパン	2011/11/29	2011/12/2
SH-02D	シャープ	2011/12/1	2011/12/7
L-01D	LG Electronics Japan	2011/12/9	2011/12/15
F-05D	富士通	2011/12/14	2011/12/17
P-02D	パナソニック	2011/12/19	2011/12/23

3.1.1. 売り手のデータ

携帯電話市場にはエヌ・ティ・ティ・ドコモ（以下ドコモと呼ぶ）、KDDI株式会社のブランドである au、ソフトバンクモバイル株式会社のブランドであるソフトバンクがある。au とソフトバンクは Android と iOS の両方を取り扱っている。これに対して、現時点でドコモは Android のみを取り扱っているという違いがある。すなわち、ドコモが Android と競合する iOS 搭載のスマートフォンを取り扱っていないことから Android 搭載機同士の競争について分析することに適している。その為、本稿ではドコモのスマートフォンを分析の対象とする。

²⁰ 「ソニー」は「ソニーモバイルコミュニケーションズ」を省略して記したものである（以下同様）。

また、メーカー側の広告情報には、プレスリリース、カタログ、CM 等がある。プレスリリース以外の資料は写真、映像などテキストマイニングで取り扱うことの出来ない情報を含んでいる。前述で述べたとおり本稿ではテキストマイニングを使う方針である。したがって、本稿ではプレスリリースを使用する。

分析対象期間は、プレスリリース公表日が 2011 年 1 月 1 日から 2011 年 12 月 31 日までのプレスリリースである。また、統一性を図る為、一つの機種につき一つのプレスリリースとしている。また、分析対象は Android 搭載のスマートフォンに関する記事のみとし、それ以外のプレスリリースは分析の対象外としている。

表 5 は各機種のプレスリリースに登場する形態素数をまとめたものである。ここで出現数とは、MeCab が与えられた文章を判断し、分割した形態素の数のことである。尚、表 5 では、「総数」「名詞・形容詞のみ」「不要文字除去後」の 3 つの段階を示している。1 つ目の「総数」とは MeCab によって区切られた形態素の総数のことである。例えば、「すもももももものうち」という語であれば「すもも」「も」「もも」「も」「もも」「の」「うち」と区切られ、出現数は 7 となる。2 つ目の「名詞・形容詞のみ」とは「総数」から名詞・形容詞のみに絞り込んだものである。全ての品詞を用いないのは、助詞や助動詞等の語は文章中に多く用いられており、内容とは無関係であると考えた為である。先程の、「すももももももものうち」であれば、「も」「の」は助詞と判断し除去され、「すもも」「もも」「もも」「うち」といった名詞が残り、出現数は 4 となる。

3 つ目の「不要文字除去後」とは「名詞・形容詞のみ」から更に句点や句読点、「-」「/」といった記号などを除去したものである。「名詞・形容詞のみ」に絞り込んでも、Mecab の性質上、句点や句読点、記号を名詞あるいは形容詞として判断してしまう為、ここで除去する。尚、種類数は出現数の重複を除いた数のことである。出現数は平均して 192 形態素であり、種類数は平均して 131 形態素とおおよそ出現数の 7 割となっている。

表5 売り手の形態素の出現数

機種	総数		名詞・形容詞のみ		不要文字除去後	
	出現数	種類数	出現数	種類数	出現数	種類数
L-04C	323	158	167	107	135	95
N-04C	360	185	188	131	155	121
SO-01C	555	256	311	201	269	183
L-07C	264	138	131	88	107	77
SC-02C	352	177	184	134	163	124
N-06C	411	212	213	159	188	146
Bold9780	307	156	158	115	151	111
SO-02C	352	177	410	236	368	218
SH-13C	293	152	139	101	129	94
F-12C	301	150	162	106	145	96
P-07C	242	124	121	87	105	78
SO-03C	597	271	328	219	297	205
SO-01D	400	173	205	129	193	121
P-01D	309	151	143	105	133	99
T-01D	589	270	342	221	292	193
SC-03D	405	189	213	144	187	131
F-03D	401	190	195	130	177	120
SH-01D	604	242	311	183	284	173
SC-04D	256	129	155	93	127	81
SH-02D	408	187	198	136	189	129
L-01D	401	186	222	146	203	139
F-05D	626	270	341	211	296	190
P-02D	274	140	134	97	116	87
平均	393	186	216	143	192	131
合計	9,030	4,283	4,971	3,279	4,409	3,011

3.1.2. 買い手のデータ

インターネットなどの情報コミュニケーション技術の発達は、消費者の購買環境に大きな変化をもたらした。新商品やサービスの普及の地域格差は劇的に縮小し、新たな購入方法や消費者と企業との新しい関係が次々に生み出されている（挾間ほか2004）。

また、消費者が製品についてのクチコミをやり取りしているインターネットコミュニティが現在、注目されつつある。インターネットコミュニティの中でも、消費者の嗜好を分

析しやすいクチコミサイト²¹に今回、注目する。買い手の意識を探る為のデータとして、クチコミが非常に多く集積する評価サイトの株式会社カカコム²²（以下、価格コムと呼ぶ）のレビューを用いることにする。それというのも、こうした評価サイトは商品認知の情報を提供する情報源として有効であることが判明しているからである（佐々木 2004）。

しかしながら、形態素解析器 **MeCab** の性能には限度があり、カカコム全てのレビューを分析対象とすることは出来ない。その為、発売日直後の文字数 3000 字分のレビューを分析することにした。また、売り手のデータ同様に名詞と形容詞への限定、及び、不要文字の除去を行い、最終的におよそ三分の一の形態素が分析対象となる（表 6）。不要文字除去後の出現数は平均して 556 形態素であり、種類数は 318 形態素と出現数の約 3 分の 2 程度である。買い手の出現数に対する種類数の割合が売り手の場合よりも高い。その理由として、売り手側は買い手側と異なり、販売戦略の一環として意図して同じ単語を用いているということが考えられる。

²¹ ここでのクチコミサイトとは、「一般消費者によって発信された特定の製品やサービスに関する使用体験や感想といったクチコミ情報が大量に蓄積されているサイト」と定義する。

²² 株式会社カカコムでは、様々なサービスを提供しているが、ここでは購買支援サイト「価格.com」の部分についてのみ言及する。

表 6 買い手の形態素出現数

機種名	総数		名詞・形容詞のみ		不要文字除去後	
	出現数	種類数	出現数	種類数	出現数	種類数
L-04C	1,725	514	664	354	577	329
N-04C	1,766	507	635	326	532	301
SO-01C	1,764	496	622	323	543	301
L-07C	1,753	508	665	356	569	329
SC-02C	1,775	499	653	337	541	312
N-06C	1,764	592	703	400	617	376
Bold9780	1,806	552	660	367	562	346
SO-02C	1,750	495	637	329	555	305
SH-13C	1,717	503	642	338	565	315
F-12C	1,823	540	664	351	551	321
P-07C	1,795	511	639	343	496	308
SO-03C	1,735	541	671	359	579	340
SO-01D	1,711	487	647	329	560	308
P-01D	1,769	529	690	357	559	330
T-01D	1,824	548	679	351	532	321
SC-03D	1,840	523	673	341	537	313
F-03D	1,759	471	591	283	509	265
SH-01D	1,782	555	723	364	631	341
SC-04D	1,797	533	680	360	550	334
SH-02D	1,738	518	618	321	518	296
L-01D	1,759	519	707	360	622	336
F-05D	1,835	488	624	304	514	278
P-02D	1,789	531	690	334	563	305
平均	1,773	520	660	343	556	318
合計	40,776	11,960	15,177	7,877	12,782	7,310

3.2. 分析方法

売り手と買い手のメッセージから、それぞれの関連を探索的に明らかにしていく。ソフトウェアとして、統計ソフトウェア R、形態素解析器 MeCab、および MeCab を R から使用するためのインターフェースである RMeCab を用いる。分析対象は名詞と形容詞に限定し、記号等の不要と判断した形態素も除いたデータである。

売り手データと買い手データともに、2 段階の分析を行う。まず初めに上位 5 形態素を取り出し、おおよその傾向を見つけていく。次に上位 10 形態素を対象とする対応分析を行い、形態素同士の関連や形態素と機種との対応関係を見る。

最後に、まとめとして、売り手データと買い手データの対応している箇所について確認していく。ここでは売り手と買い手との間で共通する点と異なる点を見つけ出し、論じる。

第4章 売り手の構成要素比較

まず，売り手側の販売戦略を見るためにプレスリリースに出現する形態素を分析する．分析対象は MeCab で抽出した 9,427 個の形態素から句読点，助詞，特殊文字を除いた残りの 4,385 形態素である．種類数としては 1,572 種類となる．

4.1. 上位 5 形態素の傾向

表7は売り手側の上位5形態素と出現数に占める比率を機種別に示したものである．尚，ここでの出現率は名詞と形容詞に限定し，不要文字を除去した出現率である．

5%を超える形態素が多く存在し，上位 5 形態素の比率の合計も 13%から 34%と多くの割合を占めている．

表 7 機種別売り手の上位 5 形態素

機種	上位五形態素 の比率合計	1 位	2 位	3 位	4 位	5 位	出現数
L-04C		搭載	キー	やすい	対応	モード	
	34%	14%	7%	5%	4%	4%	95
N-04C		性	スマート	アプリ	薄	世界	
	13%	3%	3%	3%	2%	2%	121
SO-01C		撮影	対応	実現	機能	搭載	
	18%	5%	4%	3%	3%	3%	183
L-07C		ディスプレイ	明るい	薄い	色	輝度	
	21%	5%	4%	4%	4%	4%	77
SC-02C		搭載	対応	再生	ディスプレイ	録画	
	16%	5%	3%	3%	3%	2%	124
N-06C		入力	amadana	アプリ	防水	対応	
	16%	5%	4%	3%	2%	2%	146
Bold9780		可能	搭載	メール	SNS	機能	
	24%	5%	5%	5%	5%	4%	111
SO-02C		搭載	撮影	機能	可能	対応	
	22%	6%	4%	4%	4%	4%	218
SH-13C		やすい	機能	シンプル	優しい	搭載	
	21%	9%	5%	3%	2%	2%	94
F-12C		搭載	やすい	機能	音声	アプリ	
	27%	7%	7%	5%	4%	4%	96
P-07C		画面	キー	情報	自分	やすい	
	23%	6%	5%	4%	4%	4%	78
SO-03C		搭載	入力	対応	機能	撮影	
	17%	4%	4%	3%	3%	3%	205
SO-01D		ゲーム	搭載	対応	快適	可能	
	28%	10%	7%	4%	4%	3%	121
P-01D		搭載	自分	コンパクト	Fit	便利	
	19%	4%	4%	4%	4%	3%	99
T-01D		搭載	撮影	対応	技術	映像	
	16%	4%	3%	3%	3%	3%	193

第4章 売り手の構成要素比較

機種	上位五形態素 の比率合計	1 位	2 位	3 位	4 位	5 位	出現 数
SC-03D		搭載	対応	可能	GHz	動画	
	18%	5%	5%	3%	3%	2%	131
F-03D		アプリ	充電	手書き	デザイン	ジュエリー	
	21%	5%	4%	4%	4%	4%	120
SH-01D		AQUOS	表示	搭載	対応	機能	
	16%	4%	3%	3%	3%	3%	173
SC-04D		搭載	HD	快適	最新	ディスプレイ	
	34%	9%	9%	6%	5%	5%	81
SH-02D		操作	便利	入力	搭載	機能	
	22%	5%	5%	4%	4%	4%	129
L-01D		操作	搭載	機能	画面	ディスプレイ	
	17%	4%	4%	3%	3%	3%	139
F-05D		搭載	対応	辞書	機能	可能	
	19%	7%	3%	3%	3%	3%	190
P-02D		搭載	写真	キレイ	ピクチャ	くさい	
	23%	6%	6%	5%	3%	3%	87

次に、これら上位5形態素をいくつかのカテゴリーに分類してみよう。

「搭載」という形態素は「QWERTY キーも搭載しているので、長い文章でもらくらく文字入力ができます」といった今までにない新しい機能、「対応」は「快適な文字入力を実現する、ソニー・エリクソン独自の日本語・英語予測変換・入力アシスト機能「POBox Touch」を搭載。そのほかFM ラジオ機能対応。」といった新しい機能に対応するという意味合いで用いられている。また、「最新」は例えば「Android 4.0 ならではの最新 GoogleTM サービス体験ができる。」といった新しいサービスを提供するという意味で用いられる。さらに、「HD」は「フル HD 録画 (30fps) ・フル HD 再生に対応。HD 画質の YouTubeTM も快適に楽しめる。」といった新しい機能を表現するのに用いられる。本稿では、そういった今までにない新しい機能、新しいサービスを総称して「新機能」と呼ぶ。

「新機能」とは反対にスマートフォンを使用する上で最低限必要な機能を「基本機能」と呼ぶ。例えば、「メール」の「@docomo.ne.jp のメールもプッシュ方式で届くので、外出先でもメールを受信できます。」のように既存の機能については、「基本機能」として分類しよう。

一方、「コンパクトなジュエリーデザイン！ きらめくエレガントスマートフォンジュエリーカットキーが輝くコンパクトデザイン&ジュエリーイルミネーション.」という文脈で用いられる形態素を本稿では「デザイン」と呼ぶ。ジュエリーのように装飾を意味する形態素だけでなく、装飾を意味する形態素と一緒に用いられる傾向にある軽量、小型を意味する形態素も「デザイン」と呼ぶことにする。

「その他（機能）」は「新機能」と「基本機能」のどちらかに分類できない形態素を分類し、「その他」は、三つの分類に振り分けられない形態素を分類している。例えば、「優しい」は「手に優しいフレンドリーなデザイン アシンメトリーに大きく配置されたミラーパネルが、エレガントでありながら新しいシンプルさを演出。操作する手に優しい使い勝手を実現した、大きくて押しやすいキーボタン.」といった基本機能がデザインに区分けするかの判断が難しい為、「その他」に分類している。また、「AQUOS」「amadana」のように固有名詞もその他としている。

表8は、これら「新機能」「基本機能」「デザイン」「その他（機能）」「その他」の5つに分類した形態素をまとめたものである。「新機能」に分類される形態素が多く、売り手側は新機能重視のスマートフォンを販売しようとしていることが伺える。

表8 形態素の分類

新機能		基本機能	デザイン	その他（機能）	その他
搭載	便利	ディスプレイ	薄 ²³	明るい	やすい
動画	映像	メール	色	再生	モード
対応	辞書	画面	シンプル	入力	性
撮影	写真	キー	コンパクト	情報	スマート
実現	ピクチャ	充電	デザイン	快適	世界
輝度	アプリ	音声	ジュエリー	技術	amadana
録画	手書き	操作	キレイ	機能	優しい
防水	HD				Fit
可能	最新				AQUOS
SNS	GHz				表示
ゲーム					自分

²³ 尚、「薄い」は「薄」として含まれている。

4.2. 上位 10 形態素の対応分析

次に 23 機種の上位 10 形態素を対応分析した (図 2)。

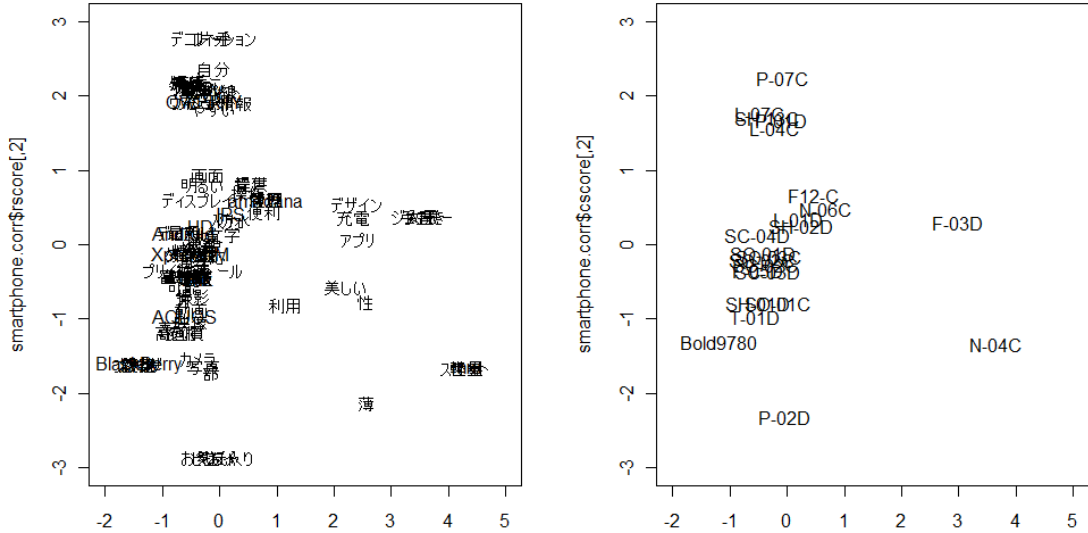


図 2 売り手の上位 10 形態素とスマートフォン 23 機種との対応分析

図 2 の対応分析表には 23 機種×10 形態素で 230 の形態素が表されている．その為，必然的に似た意味の形態素は集中し重なる構造になっている．図 2 を見ると，形態素と機種がおおよそ対応しており，大きく 2 つの固まりに分けることができる．つまり「搭載」「防水」といった新機能に関する形態素 (図 3) と「デザイン」「薄」といったデザインを重視する形態素の固まり (図 4) と分けられる．

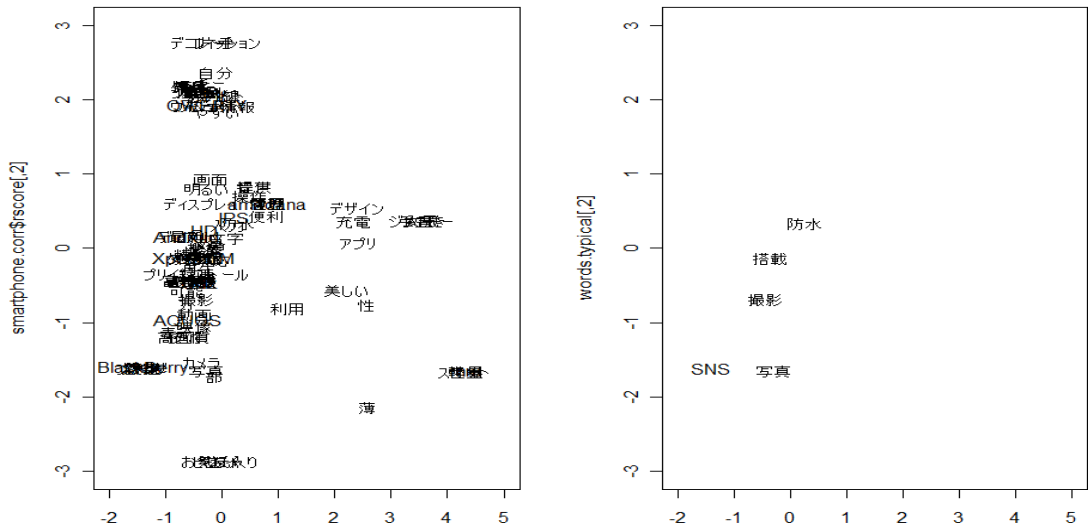


図 3 新機能に関する典型的な要素の対応分析表

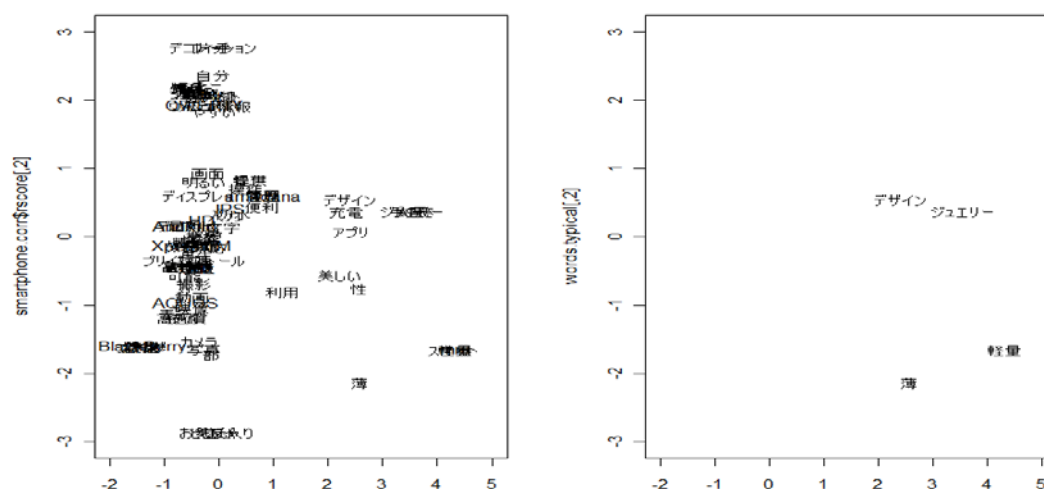


図4 デザインに関する典型的な要素の対応分析表

F-03D²⁴, N-04C²⁵は「デザイン」の固まりに位置するが、機種が2機種と少ない(図2). 反対に、「搭載」や「防水」といった新機能に形態素、機種が集中しており(図2)(図3), 形態素自体の量も非常に多い(表8). 以上の対応分析からも、売り手が考える販売戦略は新機能を重視したものであると考えられる.

²⁴ F-03D は、ジュエリーデザインが施されたコンパクトなスマートフォンである.

²⁵ N-04C は、薄さ 7.7mm (最厚部は 8.7mm) と薄型で、重量 105g と軽量さが特徴である.

第5章 買い手の構成要素比較

買い手側の意識を見るために、カカクコムレビューに出現する形態素を分析する。分析対象は、MeCabで抽出した40,776個の形態素から句読点、助詞、特殊記号を除いた残りの形態素12,782形態素である。種類数としては、7,310種類になる。

5.1. 上位5形態素の傾向

表9は買い手の上位5形態素と出現数に占める比率を機種別に示したものである。「機種」、「ボタン」、「画面」といった形態素が上位にきていることが分かる。

表 9 買い手の上位 5 形態素一覧²⁶

機種	上位五形態素 の比率の合計	1 位	2 位	3 位	4 位	5 位	出現 数
L-04C	9%	機種	キーボード	画面	ボタン	携帯	577
N-04C	10%	ボタン	画面	バッテリー	操作	不具合	532
SO-01C	12%	ボタン	デザイン	良い	使用	操作	543
L-07C	7%	良い	問題	機能	携帯	機種	569
SC-02C	8%	機種	デザイン	通話	人	ボタン	541
N-06C	7%	問題	ボタン	無し	カード	薄い	617
Bold9780	10%	問題	携帯	変換	ガラケー	普通	562
SO-02C	11%	機種	問題	画面	操作	携帯	555
SH-13C	9%	携帯	充電	変換	やすい	良い	565
F-12C	7%	良い	画面	ボタン	iPhone	購入	543
P-07C	9%	やすい	操作	機種	画面	良い	496
SO-03C	9%	ボタン	画面	音質	バッテリー	良い	579
SO-01D	11%	ゲーム	操作	機種	良い	普通	553
P-01D	7%	機種	文字	良い	小さい	画面	559

²⁶ 「よい」の形態素は「良い」の形態素と同義とみなし、「良い」として表記している。

第6章 まとめ

機種	上位五形態素 の比率の合計	1 位	2 位	3 位	4 位	5 位	出現 数
T-01D		充電	良い	機種	普通	使用	
	8%	2%	2%	2%	1%	1%	532
SC-03D		機種	良い	大きい	画面	バッテリー	
	10%	3%	2%	2%	2%	1%	537
F-03D		いい	アプリ	機種	やすい	充電	
	11%	3%	2%	2%	2%	2%	520
SH-01D		画面	音	良い	ボタン	変換	
	7%	2%	2%	1%	1%	1%	619
SC-04D		ボタン	良い	操作	変換	自分	
	8%	2%	2%	2%	1%	1%	546
SH-02D		携帯	バッテリー	キー	いい	画面	
	11%	3%	2%	2%	2%	2%	520
L-01D		良い	バッテリー	音	悪い	機種	
	10%	3%	2%	2%	2%	1%	620
F-05D		バッテリー	普通	良い	変換	画面	
	10%	2%	2%	2%	2%	2%	530
P-02D		充電	普通	音	購入	安い	
	8%	2%	2%	2%	1%	1%	560

表 9 には様々な形態素があるが、いくつかに分類できる。そこで買い手においても、売り手同様に「新機能」「基本機能」「デザイン」「その他（機能）」「その他」の 5 つに分類した（表 10）。

表 10 買い手の形態素の分類

新機能	基本機能	デザイン	その他（機能）	その他	
ゲーム	キーボード	デザイン	機種	携帯	iPhone
アプリ	画面	薄い	機能	不具合	購入
	ボタン	小さい	交換	良い	大きい
	バッテリー		音質	使用	いい
	通話		音	問題	悪い
	画面			人	安い
	充電			無し	
	キー			カード	
	テンキー			ガラケー	
	操作			普通	
	文字			やすい	

ここから、消費者が基本機能に対して関心が非常に高いことが分かる。売り手側が新機能を重要視している為、消費者も同様ではないかとも考えられるが、2 章におけるスマートフォンの選択理由（総務省 2012b）で確認した。「パソコンと同じ画面でネットの閲覧ができる」、「画面が大きくてみやすい」が上位にきていることを考えると消費者が基本機能を重視するのも自然なことである。

5.2. 上位 10 形態素の対応分析

以上のことを確認するために，上位 10 形態素の対応分析をしてみよう．

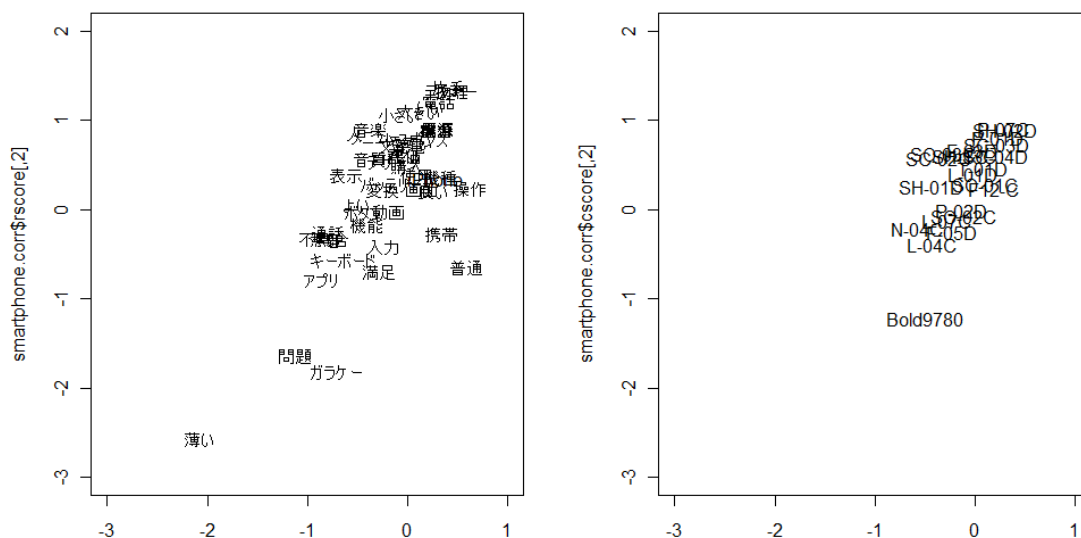


図5 買い手の上位 10 形態素とスマートフォン 23 機種との対応分析

図 5²⁷を見ると，一箇所に集中して固まっているということが分かる．機種も形態素に対応し，一箇所に集中している．

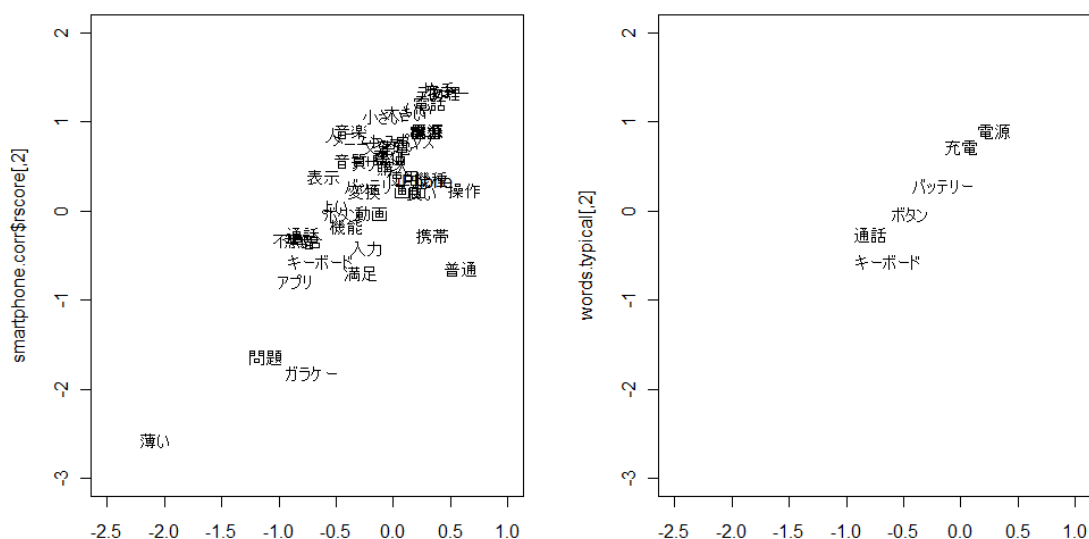


図6 基本的な機能を意味する形態素に絞った対応分析

図 6 では，右上の形態素が集中している箇所に，「電源」「バッテリー」などの「基本機能」の形態素が集中して位置していることが分かる．買い手のレビューを参照した所，例

²⁷ 図 5，図 6 において「ゲーム」，「無い」の 2 つの形態素は枠外に位置している．

例えば「バッテリー」は、「バッテリーの容量が明らかに足りません。通勤で1時間くらい音楽を聴いたりウェブを見たりしていると、30～35%バッテリーが減ります。」のように買い手にとってバッテリーは充実している必要性があるということが分かる。また、「ボタン」「キーボード」についても「ボタン操作 QWERTY フルキーボードがついているので便利。特に長文を打つ時は楽。」といったように、必要性が高いということが分かる。

以上の分析から、売り手は「防水」などの機能を重視していたが、買い手は、そういった「新機能」の面よりも「基本機能」の方を重視しているということが分かる。

尚、「薄い」は表9を見ると、N-06C のたった1箇所にのみ存在し、比率も1%しかない。また、図5においてN-06C と「薄い」の対応関係を見ても、遠く離れており関係性は認められない。さらに、一番近い機種である Bold9780 においても「薄い」という形態素は存在しない為、図6の「薄い」という形態素に関しては意識しないことにする。

第6章 まとめ

データは、携帯電話会社が新製品を発表する際のプレスリリース、評価サイトにおけるレビューを用い、分析手法としては対応分析を行った。

これまでの結果から、売り手は新機能を重視した販売戦略を取り、買い手は基本機能を重視しているということが分かった。このように買い手側は基本性能の方に関心が高いようであるが、以前から売り手が買い手の声を重要視しておらず、改善される傾向が見られない。かなり前に発売されたスマートフォンを購入した消費者の声として、「バッテリーこれは以前の SH-12C²⁸もそうですが、今のスマホならしょうがないと諦めています。」という声がある。尚、SH-12C と文章元である SH-01D の間には7ヶ月もの差がある。

本稿では、売り手の販売戦略を見る為にプレスリリースを用いたが、そもそもプレスリリースは自社の製品やサービスなどをマスコミや買い手に広く伝えるものである。その為、プレスリリースに買い手の声を取り上げられていないということは、買い手の声を正しく認識していないということになる。勿論、売り手は基本機能が付いていることが当たり前と考え、重要視していない可能性も考えられる。例えば、プレスリリースの文章において基本機能よりも新機能をピックアップすることによって他機種と差別化を図るということも考えられる。しかしながら、買い手が重要視する肝心の基本機能について、おざなりにしているようでは問題である。

本稿では対応分析を行い、売り手と買い手の意識及び意識のずれを探索的に分析した。本稿ではテキストマイニングの手法を用いることによって、言語データを量的に分析したことに一つの特徴があり、質的研究と量的研究の溝を埋めるという点においても意味があると考えられる。ただ、いくつかの課題があり、大きく3つの課題がある。1つ目は、本稿では、分析対象が Android のスマートフォンであったことから、得られた結果には、Android のスマートフォンに対してのみ当てはまるという偏りがある。事実、「赤外線」「防水」の形態素は iPhone には無い Android 独自の機能である。2つ目は、MeCab の性能面の限界により文字数が限定されてしまったことである。3つ目は、対応分析にとどまっており、クラスター分析など高度な分析まで到達していないということがある。今後、本稿をより良いものにするにあたって、更なる発展が望まれる。

²⁸ SH-12C は 2011 年 5 月 20 日に発売されたスマートフォンである。ユーザの声は SH-01D のカカクコム
のレビューを引用したものであり、発売日は約7カ月の差がある。

文献

- 安藤俊幸, 2009, 「テキストマイニングと統計解析言語 R による特許情報の可視化」『情報管理』, 52(1): 20-31.
- 有村博樹, 2003, 「テキストマイニング——ウェブデータからの知識発見を目指して」『日本化学会情報化学部会誌』, 21(2): 28.
- 挾間雅義・笠雄二・赤木文男, 2007, 「企業—消費者間取引の現状と問題点について」『生産管理』 11(1): 73-78.
- 保田明夫, 2003, 「テキストマイニングの技術と適用性」『薬学図書館』, 48(4): 247-252.
- Howard, J.A. and J. N. Sheth, 1969, *The Theory of Buyer Behavior*, John Wiley and Sons.
- 樋口耕一, 2004, 「テキスト型データの計量的分析——2つのアプローチの峻別と統合」『理論と方法』, 19(1): 101-115.
- 石田基弘, 2008, 『R によるテキストマイニング入門』森北出版株式会社.
- 伊藤大介, 2004, 「テキストマイニング手法を用いて分析した美術館来館者の生活における美術館の存在意義」『文化経済学』, 5(3): 101-110.
- 岩崎学, 1999, 「データマイニングと知識発見——統計学の観点から」『行動計量学』, 26(1): 46-58.
- 岩田修一, 2012, 「ビッグデータの時代」『情報管理』, 55(8): 543-551.
- 和泉潔・松井宏樹・松尾豊, 2007, 「人工市場とテキストマイニングの融合による市場分析」『人工知能学会論文誌』, 22: 397-404.
- 片岡寛, 1997, 「商品の寿命」『マテリアルライフ』, 9(3): 99-111.
- 工藤拓, 2006, *MeCab : Yet Another Part-of-Speech and Morphological Analyzer*, Kyoto, Japan, <http://mecab.sourceforge.net/>.
- 黒橋禎夫・長尾真, 1998, 『日本語形態素解析システム JUMAN version 3.6 使用説明書』, 京都大学大学院 情報学研究科.
- 松本裕治・高岡一馬・浅原正幸, 2007, 『形態素解析システム「茶筌」 version2.4.0 使用説明書第二版』, 奈良先端科学技術大学院大学松本研究室, (2012 年 12 月 17 日取得, <http://sourceforge.jp/projects/chasen-legacy/docs/chasen-2.4.0-manual-j.pdf/ja/2/chasen-2.4.0-manual-j.pdf.pdf>).
- 松村真宏・三浦麻子, 2009, 『人文・社会科学のためのテキストマイニング』誠信書房.
- 三浦俊彦, 1993, 「消費者情報処理とマーケティング戦略」『消費者行動研究』, 1(1): 101-120.

- , 1997, 「日本の消費者はタフな消費者か? ——消費者行動の国際比較研究へ向けての一試論」『消費者行動研究』, 5(2): 59-76.
- 百瀬章, 2009, 「携帯電話機メーカーに求められる新しいサービスモデル」『開発工学』, 29(1):29-34.
- 中井隆, 2008, 「テキストマイニングによる知財ポートフォリオ分析」『情報管理』, 51(3): 194-206.
- 佐々木裕一, 2004, 「商品購買における評価サイトの有効度——評価サイトユーザーにおける評価サイト／雑誌／ロコミの有効度比較」『情報メディア研究』, 3(1): 29-42.
- 佐藤岳文・堀田昌英, 2006, 「Web マイニングを用いた因果ネットワークの自動構築手法の開発」『社会技術研究論文集』, 44(0): 66-74.
- 杉本徹雄, 1999, 「これからの消費者行動研究」『繊維製品消費科学会誌』, 40(1): 25-30.
- 高坂健次, 1986, 「数理社会学の意義と必要性」『理論と方法』, 1(1): 1-14.
- 鷺尾隆, 2003, 「データマイニング——その発展と今後の展望」『日本化学会情報科学部会誌』, 21(2): 37.
- 総務省, 2012a, 『平成 24 年版 情報通信白書』, 株式会社ぎょうせい.
- , 2012b, 「情報通信産業・サービスの動向・国際比較に関する調査研究」, (2012 年 12 月 1 日取得, http://www.soumu.go.jp/johotsusintokei/linkdata/h24_05_houkoku.pdf).
- , 2012c, 「スマートフォンをめぐる現状と課題」, 総務省通信基盤局電気通信事業部消費者行政課 利用者視点を踏まえた ICT サービスに係る諸問題に関する研究会, (2012 年 12 月 1 日取得, http://www.soumu.go.jp/main_content/000143085.pdf).