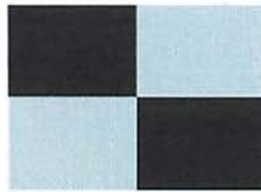


平成 27 年度修士論文

ユーザプロフィール獲得に向けた  
高齢者と音声対話ロボットの雑談履歴の解析

A Dialogue Analysis of Elderly Person

with a Chat Robot to Acquire User Profiles



TOKYO METROPOLITAN UNIVERSITY

首都大学東京

学籍番号：14890515

氏名：篠田 遥子

指導教員：山口 亨

首都大学東京

システムデザイン研究科

情報通信システム学域

知能情報処理研究室

## 概要

近年、世界各国で高齢化が進んでいる。日本国内は他国と比較すると高齢化・少子化が進み深刻な超高齢社会を迎えている。それに伴い、独居高齢者の割合も増加している。独居高齢者の問題として、家でのコミュニケーションの頻度が低く、社会との接点が減少していることがあげられる。また日本でのうつ傾向の割合は中高年がもっとも頻度が高いと報告されている。会話の減少はうつ傾向を引き起こしやすくなる。予防のためにも会話を行うことが重要な要素であると考えられる。

一方で、先行研究のロボットは短期間の会話を想定しており、長期間の利用を想定していない。また長期間のコミュニケーションロボットであっても同じ会話の繰り返しになってしまい、雑談時に個人に適した返答や情報を与えることができない。

この問題を解決するために音声対話ロボットと雑談を行い、その結果である雑談履歴を解析することで、高齢者と音声対話ロボットの『雑談履歴』から特徴キーワードやユーザプロファイルの獲得を目標とする。

ロボットがユーザの話しているカテゴリを認識し、会話の流れをコントロールするために複数の特徴キーワードから成り立っているカテゴリが必要である。高齢者との会話は特定のテーマを定めたものではなく、人同士でよく行われている雑談という形を用いて会話を行う。

本研究の解析として、音声対話ロボットが自動的におこなったカテゴリ付け結果を検証する。また、ユーザプロファイルの獲得に向け特徴キーワードが獲得できるか検証する。

# Abstract

Recently, the number of elderly people increases all over the world. Japan has large number of elderly people and small number of children compare to other countries, so it is a serious super ageing society. The percentage of elderly people living alone is also increasing with that. "Decrease of communication at the house" and "Decrease of contact with society" are considered as a problem of elderly people living alone. "The middle and old age are highest." is reported the percentage of the depressed tendency in Japan. Decrease of conversation tends to cause the depressed tendency. The conversation is an important element for prevention.

On the other hand, a robot of the preceding study assumes short term conversation instead of long term conversation. Even a long communication can be done by a robot by repeating the same conversation. And in a chat, it can't give suitable reply and information to an individual.

The research proposes a chat with a chat robot to solve this problem. The dialogue of the result is analyzed. And the research aims to get the feature keyword and the user profile from a dialogue of elderly person with a chat robot. A robot recognizes a category of the user's chat contents, and a category which consists of the feature keyword is needed to control a flow of conversation. The research has decided to do "conversation" using form of chat which does not have any theme as communication with a robot.

As an analysis of this research, a category assignment result which a chat robot did automatically is inspected. And this research inspects whether it gets the feature keyword to acquire the user profile.

# 目次

概要 .....	I
Abstract.....	II
目次 .....	III
1. はじめに .....	1
2. ユーザプロフィールによる共起関係獲得 .....	5
2.1 カテゴリ作成 <sup>[19]</sup> .....	5
2.2 共起関係の獲得に向けて .....	6
3. 音声対話ロボットと雑談履歴獲得実験 .....	8
3.1 雑談システム <sup>[19]</sup> .....	9
3.2 対話内容の更新 .....	10
4. 高齢者と音声対話ロボットの雑談履歴のカテゴリ付け検証 .....	13
4.1 音声対話ロボットと人のカテゴリ付け検証.....	14
4.2 人同士のカテゴリ付け結果比較 .....	16
5. ユーザプロフィール獲得に向けた 特徴キーワード .....	18
5.1 ユーザプロフィール .....	18
5.2 カテゴリごとの特徴キーワード .....	19
5.3 ユーザプロフィール獲得に向けた特徴キーワードマップ .....	23
6. おわりに .....	27
参考文献 .....	28
発表文献 .....	30
謝辞 .....	31
付録 .....	32
カテゴリごとの特徴キーワードマップ .....	32
18 カテゴリの特徴キーワードマップ .....	41
高齢者ごとの特徴キーワードマップ .....	44



## 1. はじめに

近年、日本の総人口に対して 65 歳以上の占める割合は増加傾向にある。1956 年、国際連合により高齢化率による分類が制定された。日本では 1970 年以降の国勢調査の結果、高齢化率が年々上昇し、2007 年ついに最高基準の 21% を超え超高齢社会となった。事実、図 1 で示すように 2011 年に内閣府が発表したデータによると総人口の 12,780 万人に対して 65 歳以上の高齢者の数は 2,975 万人であり、総人口に対し 23.3% が高齢者となっている。また、近年日本では少子化が進んでいることは周知の事実である。この事からも分かるように、今後日本の総人口は減少傾向になっていく。しかしながら高齢化率は上昇していき、2060 年頃には 2.5 人に 1 人が 65 歳以上、かつ 4 人に 1 人が 75 歳以上になっていく。その一方で、高齢化は進んでいるが在宅高齢者のうち 8 割は介護や支援を必要としない高齢者(以下、元気高齢者)である<sup>[1]</sup>。また一方で、独居高齢者の割合も年々増加傾向にある。独居高齢者の問題として、『家でのコミュニケーションの頻度が低い』、『社会との接点が減少している』ことがあげられる。また日本でのうつ傾向の割合は中高年がもっとも頻度が高いと報告されている。会話の減少はうつ傾向を引き起こしやすくなるため予防のためにも会話を行うことが重要な要素であると考えられる。

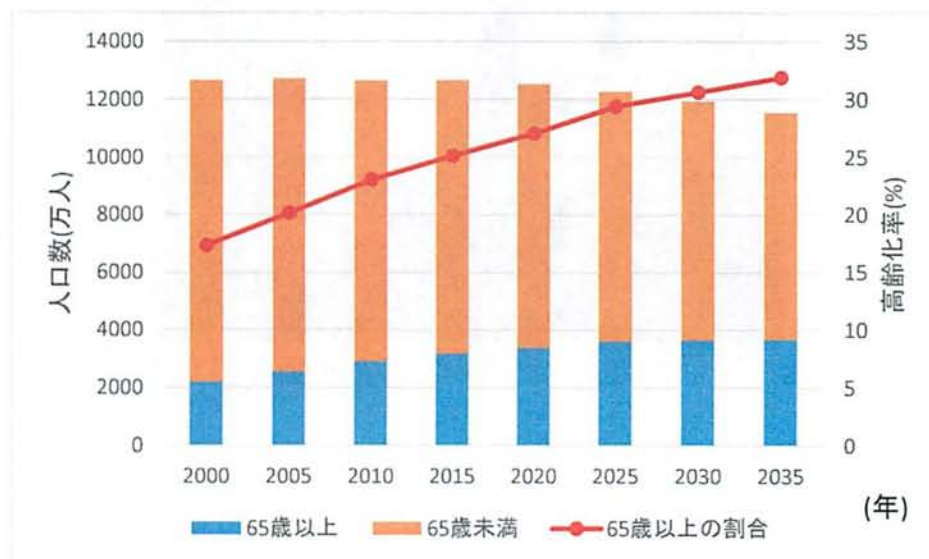


図 1 高齢者の割合推移<sup>[1]</sup>

また近年、美術館や病院の受付など様々な場所に音声対話ロボットが置いてあるが、それらはあらかじめユーザに対して発話内容が決められている。ロボットを用いた先行研究では、人と共存する生活支援ロボットとしての聞き分けロボット、お供ロボット(図 2)を用いた研究<sup>[2][3]</sup>や、インターフェースロボットである東芝製の ApriPoco を用いたコミュニケーションによる高齢者向けの対話インターフェースの提案<sup>[4][5]</sup>(図 3)などが行われている。



図 2 聞き分けロボット(左)・お供ロボット(右)<sup>[3]</sup> (TOSHIBA)

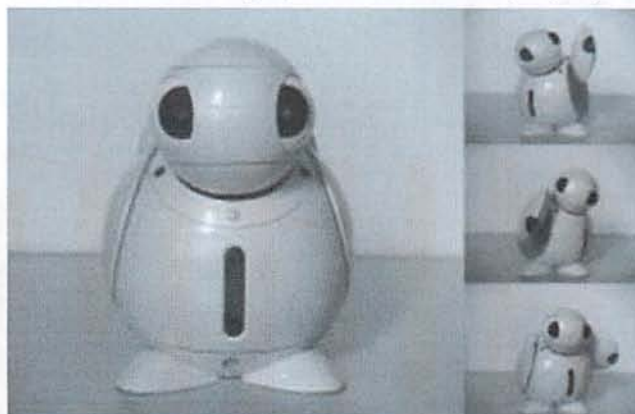


図 3 インターフェースロボット ApriPoco (TOSHIBA)<sup>[4]</sup>

また、言語による違いが会話に与える影響について調査した研究<sup>[6]</sup>や会話ロボットの先行研究として、来談者中心療法用のロボットの ELIZA<sup>[7]</sup>や ELIZA をもとにして作られた ALICE<sup>[8]</sup>がある。また、Sing らは、ほかの会話システムの調査・比較を行っている<sup>[9]</sup>。

しかしこれらのロボットは短期間の会話を想定しており、長期間の利用を想定していない。また従来は、長期間のコミュニケーションロボットであっても同じ会話の繰り返しになってしまい、雑談時に個人に適した返答を行うことや情報を提供することができない。これらの問題を解決するために Lee らは、事例ベースの会話システムを提案し、天気やテレビの番組紹介などの様な様々な状況に応じて対話を行うモデルの作成・開発を行い、対話を行った<sup>[10][11]</sup>。また、Yang らは、協調フィルタリングを対話システムに用いてシステムの構築を行った<sup>[12]</sup>。このシステム構築の結果、ユーザの知らない情報推薦に有効であることが獲得でき、更に事例ベースの対話システムを用いることで、雑談に対して有効であることが獲得できている<sup>[13]</sup>。

ユーザプロフィール・特徴の獲得に関する先行研究には、Twitter を用いて個人の特徴抽出を行う研究<sup>[14][15]</sup>、新聞記事からの記事ごとの特徴キーワードを自動的に獲得する研究<sup>[16]</sup>が行われている。しかし Twitter を用いて行われている研究は、ユーザの「今」の興味や特徴を知るために有益であるが、長期間の利用ではユーザの発言内容によりユーザの興味や特徴が変化し、長期間の利用に適していないことが挙げられている。また、新聞記事からの特徴キーワードを自動的に獲得する研究では、新聞の記事のトピック単位で獲得できる文章・キーワード自体がある程度まとまっているため、人同士の会話のようなキーワード発見が難しくない。また、ユーザ間の対話からユーザプロフィールを行うも研究<sup>[17]</sup>では、ユーザ間の対話はチャット形式で行われている。ユーザ同士の対話からユーザのプロフィールを抽出することが出来ているが、チャット形式で行っているため、長期間の利用を自宅で行うことが難しい。また、対話からの話者のプロフィール自動獲得<sup>[18]</sup>なども行われている。



本研究では、独居高齢者の問題である「家でのコミュニケーションの頻度が低い」に着目し、またうつ傾向の予防である「会話」を行うことでこの問題を解決することにした。そして本研究での「会話」はテーマを定めたものではなく、人同士が日常的に行う「雑談」の形式を用いることにした。この問題を解決するために、図 4 に示すように音声対話ロボットと雑談を行い、その結果である雑談履歴を解析することで高齢者個人の雑談に出てくる趣味・趣向などの特徴的なキーワードやその雑談からどのような内容の会話が良くされているのかといったユーザプロフィールの獲得を行うことにした。

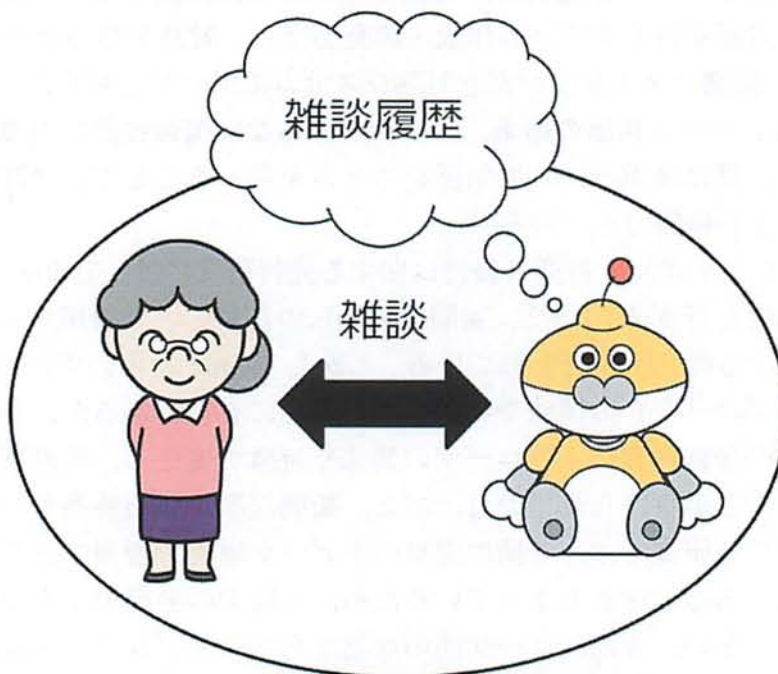


図 4 高齢者と音声対話ロボットの雑談・雑談履歴イメージ図

第 2 章では、本研究で用いているカテゴリの有効性の検証を含め、ユーザプロフィールによる共起関係の獲得の有効性を示す。さらに第 3 章では、音声対話ロボットとの雑談履歴獲得実験を行う。第 4 章では、本実験で用いた音声対話ロボットのカテゴリ付けの検証を行う。第 5 章では、本実験結果から実際にユーザプロフィールを行うために個人の特徴となる特徴キーワードの獲得を行い、その有効性を示す。



## 2. ユーザプロフィールによる共起関係獲得

### 2.1 カテゴリ作成<sup>[19]</sup>

本研究の音声対話ロボットは雑談をするためにその場で獲得した高齢者の雑談文にカテゴリ付けを行っている。このカテゴリは昨年度本研究室の野村が提案したものを用いている<sup>[19]</sup>。ロボットの雑談に用いるための定型文として学生 10 名により 313 文を作成した。その 313 文を学生 10 名でグルーピングし、18 種類作成した。また、作成段階で 18 種類に含まれなかったものを 19 番目として「その他」のカテゴリを用意し分類を行った。表 1 に作成された 18 カテゴリを示す。

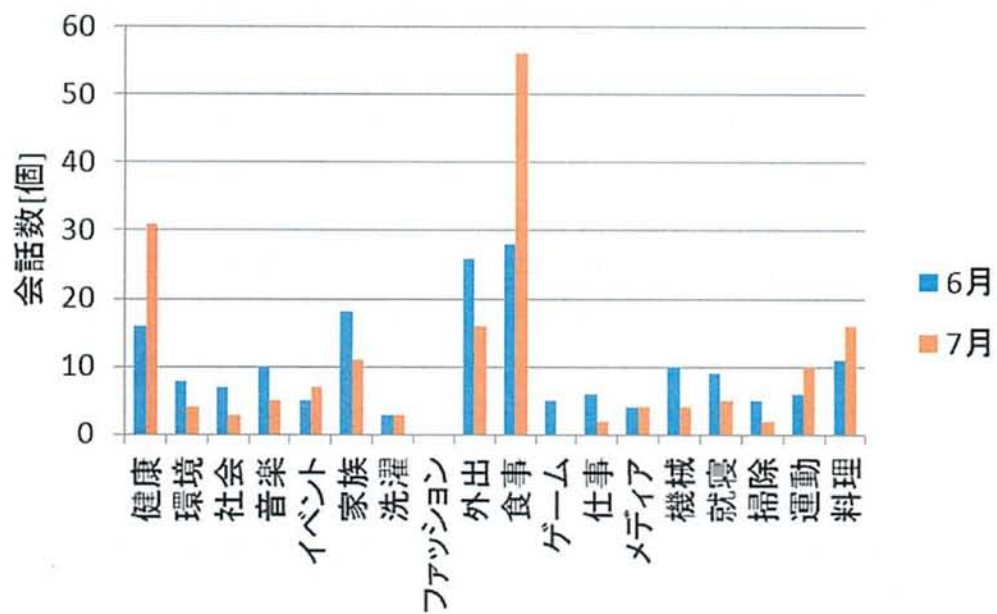
表 1 作成したカテゴリおよびカテゴリ番号

1 健康	11 ゲーム
2 環境	12 仕事
3 社会	13 メディア
4 音楽	14 機械
5 イベント	15 就寝
6 家族	16 掃除
7 洗濯	17 運動
8 ファッション	18 料理
9 外出	19 その他
10 食事	

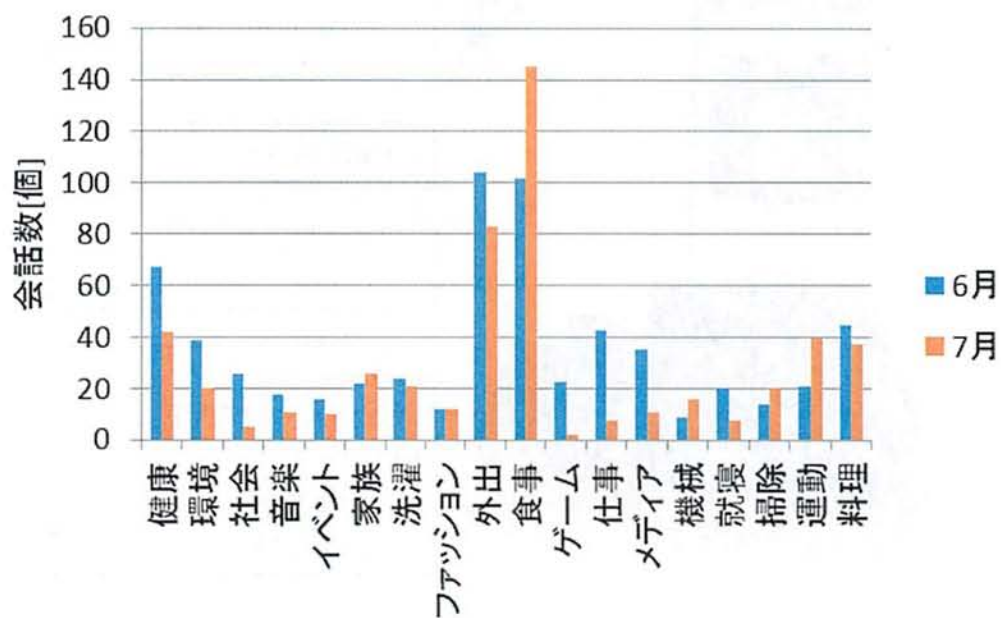
## 2.2 共起関係の獲得に向けて

音声対話ロボットとの雑談を通し、ユーザプロファイルを行うことでユーザ同士の共起関係を見つけることが出来るのではないかと考える。また作成した18カテゴリを用いて雑談履歴にカテゴリ分けを行うことで、カテゴリを用いた分析が有効であることの確認を行う。

2014年の夏季(2014年6月~2014年7月)に行った実験の雑談文にカテゴリ付けを行った結果からユーザプロファイルの獲得を行った。学生2名で実験期間中の雑談文にカテゴリ付けを行った結果の代表として、2名の高齢者の結果を図5に示す。この結果から、「高齢者毎の話題の特徴が得られる」ことが獲得できた。また、2名の結果を見ると用意したカテゴリの1つである「食事」に関する話題が多いことが分かる。この事からこの2名の被験者が「食事」の話題に関して共起の関係にあることが考えられる。この事から、高齢者間で同カテゴリに対し興味があるか等、他者との共通点の発見を行うことができ、共起関係に向けて雑談文にカテゴリ分けを行うことが有効であることが獲得出来た。更にこの事から、新たな特徴キーワード集合間のマッチングを利用する手法により新たな共起関係の獲得を行うことが可能である。



高齢者1



高齢者2

図5 カテゴリ分け結果



### 3. 音声対話ロボットと雑談履歴獲得実験

本実験で実際に用いた音声対話ロボット及び使用機能を図 6 に示す。本研究での「会話」はテーマを定めたものではなく、人同士が日常的に行う「雑談」の形式を用いている。

本雑談に関する実験の被験者は 65 歳以上の高齢者 10 名であり、実験は 2014 年度の冬季期間(2014 年 12 月~2015 年 3 月)に行い、実験期間は高齢者 1 人当たり 3~4 週間程度行った。本研究の音声対話ロボットのシステムは昨年度本研究室の野村が作成したものを用いて行っている<sup>[19]</sup>。実験の開始・終了は被験者の意思で行う形式を用いており、音声対話ロボットの挨拶から始まり、ロボットが聞き取った雑談の文章内の単語をキーワードとして会話が継続されていく。対話内容は「事前に準備しておいた定型文・対話データベース(以下、対話 DB)内の文・Wikipedia から生成した文」の 3 種類である。本実験の音声認識ソフトは「Google Speech Recognition API β 版」である。

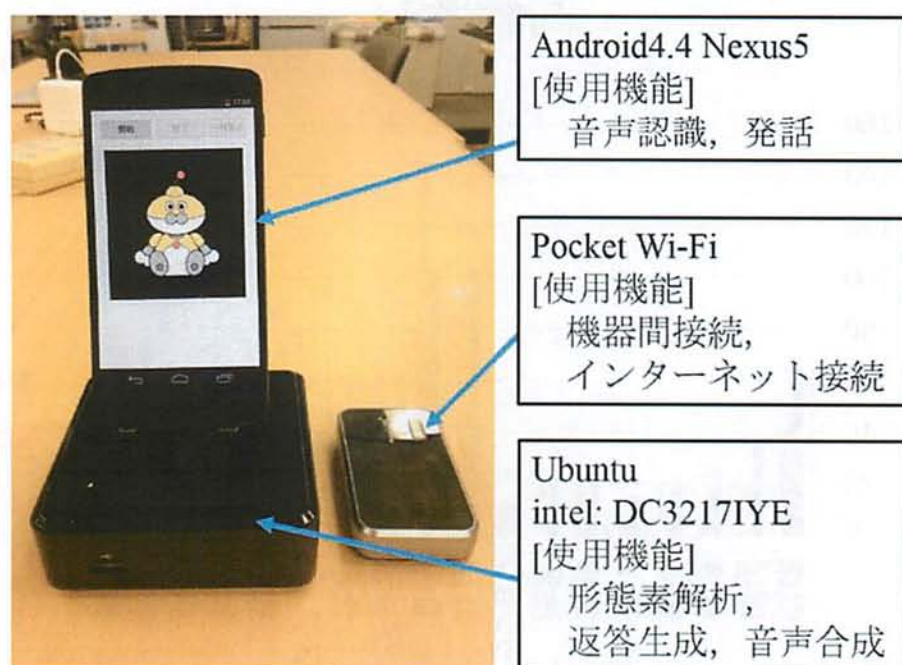


図 6 音声対話ロボット及び使用機能<sup>[19]</sup>

### 3.1 雑談システム<sup>[19]</sup>

本実験で用いた雑談システムのフローチャートを図 7 に示す。本実験の音声対話ロボットは被験者である高齢者が実際に発話した内容を聞き取ったものに形態素解析を行い、その結果から返答を生成している。形態素解析を行った結果の名詞に着目し、対話 DB 内にその名詞に対応する発話文をランダムに応答する。しかし対話 DB 内に名詞が含まれていない場合は、事前に準備してある定型文もしくは Wikipedia から生成した文章を用いて返答する。

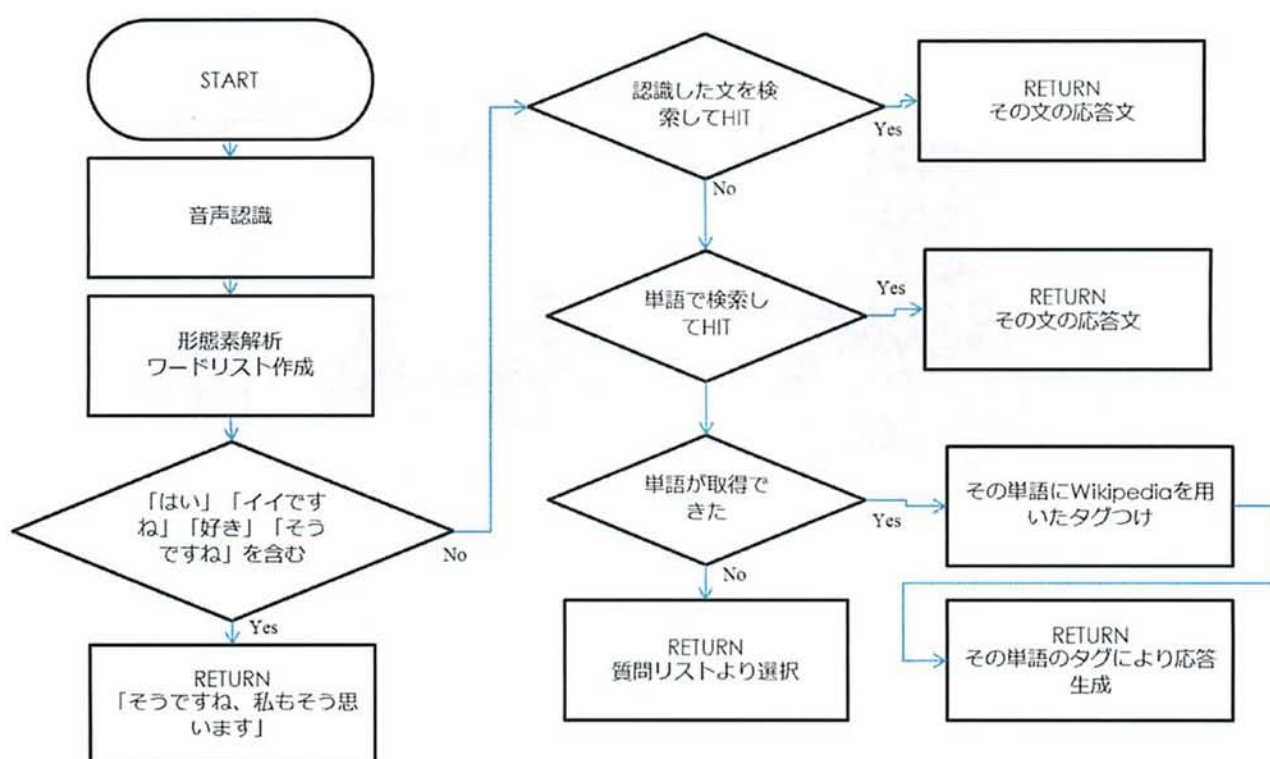


図 7 雑談システムのフローチャート<sup>[19]</sup>

### 3.2 対話内容の更新

音声対話ロボットの対話 DB 内には各ユーザの雑談内容が毎日保存されている。音声対話ロボットはユーザの発話した雑談内容を毎日クラウド上にアップデイトし、サーバ側でチェックを行う。その後、整形・選択された文がクラウド上の対話 DB に保存され、毎日各ユーザ側の対話 DB が自動的に更新されていく。更新後、ユーザ自身や他のユーザに対する音声対話ロボットの返答として利用している。図 8 にこのシステム概要図を示す。



図 8 システム概要図



2014 年の夏季(2014 年 6 月~2014 年 7 月)実験において，実際に音声対話ロボットの対話 DB 内の文章を日々更新した結果の累計文数を図 9 に示す．

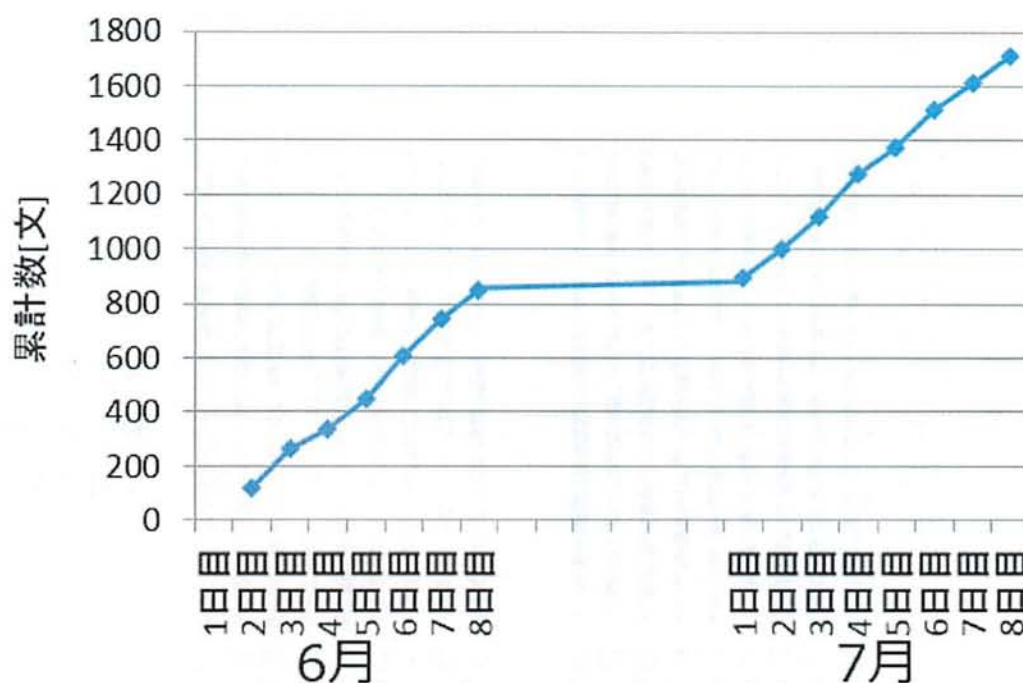


図 9 対話 DB 内の文章数の変移

このように対話 DB 内の文をクラウドサーバ上で更新を行うことで，音声対話ロボットの発話する文のレパートリィが増えていく．また 2014 年夏季は被験者数が 6 名だったので，2014 年冬季の実験では対話 DB 内の文章数はより多くなっている．

また音声対話ロボットは「事前に準備しておいた定型文・対話 DB 内の文・Wikipedia から生成した文」の 3 種類発話内容である。2014 年の夏季実験で対話 DB を更新することで、音声対話ロボットの発話内容出現率の推移を求めたものを図 10 に示す。

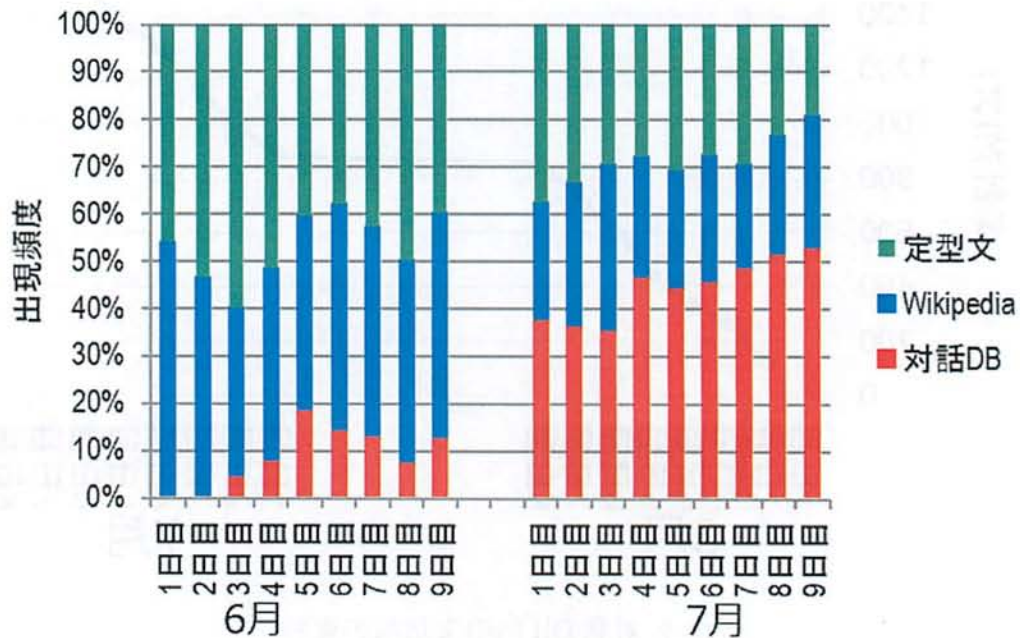


図 10 雑談内容推移

図 10 より、図 9 の結果を踏まえても対話 DB 内の文が増加すると共に、出現率も増加していることより、雑談で同じような応答をしてしまう確率が減少し、会話のマンネリ化が減っていると考えられた。

このように、図 9・10 のような結果を得ることができた図 8 に示したシステムを用いて行った雑談からユーザごとのプロフィールの獲得を目指していく。

## 4. 高齢者と音声対話ロボットの雑談履歴の カテゴリ付け検証

本研究の音声対話ロボットは、雑談をするためにその場で獲得した高齢者の雑談文にカテゴリ付けを行っている。ユーザプロフィールに向け、この自動的なカテゴリ分けがどの程度行えているのか確認を行った。表 2 にカテゴリ付け方法を示す。カテゴリ付けは、被験者の発話・その前のロボットの発話・更に前の被験者の発話の 3 箇所を見て行う。音声対話ロボット・学生共に作成したカテゴリに属さないものは作成したものとは別に「その他」のカテゴリを用意し分類を行った。

表 2 カテゴリ付け方法

ロボット発話内容	被験者発話内容	カテゴリ番号
宝くじは買ったことありますけど、当たったことはありません。	宝くじはそうですねーないですねだから最近会わせる	5
昔は海外で暮らしたいと思っていましたんですけど、最近はあまり思わないですね。	そうですね テロと金 色んな事件が起きて 日本安全かな と思いますね私も恐れに行きたいなあと思ったことがありますよ 英語 できたらいいなって	5
何をするにもお金がかかりますね。	金かかりますねえ 習い事するのに胸下の痛みるんですけどね	19
今日は習い事の日だったので、色々な方と色々な話をしました。楽しかったですよ。	そのですね 私も良行お話ししますよね w でさらにきましたね 甲子園飲みながらね]	10



## 4.1 音声対話ロボットと人のカテゴリ付け検証

音声対話ロボットの雑談履歴のカテゴリ付け検証に向け、まず被験者 1 名に対して学生 1 名でカテゴリ付けを行った。その結果を表 3 に示す。

表 3 学生(1 名)ー音声対話ロボット間のカテゴリ付け一致率

	「その他」含む	「その他」含まない	雑談文総数(文)
被験者 1	31.8%	53.3%	1277
被験者 2	50.3%	64.7%	725
被験者 3	39.9%	32.6%	1999
被験者 4	56.0%	33.3%	125
被験者 5	45.7%	37.4%	278
被験者 6	45.7%	49.0%	1581
被験者 7	36.5%	35.0%	2656
被験者 8	39.3%	39.6%	698
被験者 9	42.5%	38.4%	3355
被験者 10	51.9%	49.6%	1033

表 3 の結果より、音声対話ロボットと学生のカテゴリ付けの一致率は 30～56% であることがわかる。また、学生が「その他」に分類した雑談文を除いた結果の一致率を確認してみたところ、表 3 より比較的どの被験者も一致率にさほど変化が無いことが分かった。しかし、被験者 1 と 4 は他の被験者に比べ 20% 近くの差が出ている事が分かる。被験者 1 は「その他」を含まない方の一致率が高くなっている事より、ロボット・学生共に比較的作成した 18 カテゴリに分類できる内容であったことが考えられる。一方で被験者 4 は「その他」を含まない方の一致率が低くなっていることより、被験者 1 とは逆の 18 カテゴリに分類しきれない内容であったことが考えられる。これらより、被験者ごとにロボットが被験者の発話した内容が聞き取れているかが問題として考えられた。

しかし学生 1 名だけのカテゴリ付け結果を見てしまうと偏りも生じてしまうのではないかと考えられたため、被験者 1 名に対して学生 3 名でカテゴリ付けを行いその結果とロボットの結果を比較することにした。またカテゴリ付けを行う学生はランダムで選択したため、被験者ごとに対する学生は全員異なる。また学生にカテゴリ付けを行う順番は指定していない。学生 3 名のカテゴリ付けをまず「全員一致・2 人一致・全員不一致」で分け、「全員不一致」を除いた結果とロボットの結果を比較している。学生 3 名のカテゴリ付け結果は表 4 に示すように、「全員一致」の場合はそのカテゴリを、「2 人一致」の場合は一致し

た方のカテゴリを使用した。また、この手法で決定した学生 3 人のカテゴリとロボットのカテゴリ付けの一致率結果を表 5 に示す。

表 4 学生のカテゴリ付け決定方法

学生 1	学生 2	学生 3	使用するカテゴリ番号
9	9	9	9
19	10	10	10
10	18	10	10
2	9	19	×

表 5 学生(3 名)ー音声対話ロボット間のカテゴリ付け一致率

	「その他」含む	「その他」含まない	雑談文総数(文)
被験者 1	36.1%	64.9%	1184
被験者 2	46.3%	68.7%	681
被験者 3	52.3%	52.7%	1888
被験者 4	56.6%	35.3%	122
被験者 5	49.6%	40.5%	250
被験者 6	43.3%	54.9%	1488
被験者 7	43.6%	46.3%	2336
被験者 8	44.7%	53.0%	623
被験者 9	44.5%	40.8%	3246
被験者 10	53.4%	53.7%	965

表 4 の結果より、音声対話ロボットと学生(3 名)のカテゴリ付けの一致率は 35～56%であること、また学生が「その他」に分類した雑談文を除いた結果の一致率は 35～68%であることが分かる。しかし被験者 1 と 4 だけではなく被験者 2 の「その他」の有無による結果も他の被験者に比べ 20%近くの差が出ている事が分かる。この事から、複数の学生で見る事でより正確な一致率が求められることがまず考えられる。また全体を通し、現状の音声対話ロボットが行えるカテゴリ付けは「その他」を除いて平均 51.1%あるので、比較的カテゴリ付けが行えていると考えられる。



## 4.2 人同士のカテゴリ付け結果比較

次に音声対話ロボットが確実にカテゴリ付けを行うためには、まず人同士のカテゴリ付け結果を比較することが必要であると考えられた。そこで表 4 で示したカテゴリ付けの決定方法を基に学生同士のカテゴリ付け結果の比較を行う。その結果を表 6 に示す。

表 6 人同士のカテゴリ付け比較

	全員一致	2 人一致	全員不一致
被験者 1	48.6%	44.2%	7.2%
被験者 2	48.3%	45.7%	6.0%
被験者 3	42.7%	51.8%	5.5%
被験者 4	66.4%	31.2%	2.4%
被験者 5	35.6%	54.3%	10.1%
被験者 6	44.3%	49.8%	5.9%
被験者 7	29.9%	58.1%	12.0%
被験者 8	41.7%	47.6%	10.7%
被験者 9	57.8%	38.9%	3.3%
被験者 10	37.2%	56.2%	6.6%

表 6 より、本研究で使用した「全員一致・2 人一致」の割合は全高齢者を比較しても 80%後半は出ていることが分かった。しかし、その内訳を確認すると「全員一致」の割合が 60%を超えている高齢者もいれば、30%近くしかない高齢者もいることが分かる。

このように高齢者ごとに一致率が異なり、差がある事から、現在使用している表 1 に示した 18 カテゴリの中で学生ごとに異なったカテゴリ付けを行っているカテゴリの確認を行う。そして主な誤ったカテゴリ付けの組み合わせを表 7 に示す、



表 7 主に誤ってつけるカテゴリの組み合わせ(例)

ロボット発話内容	高齢者の発話内容	学生 1	学生 2	学生 3
今日はお弁当です。	ダメだ作ります]	10	10	18
魚を捌くのは、難しいですね。	さばけないですよ やっぱり	18	10	18
春が好きです。桜が咲いてきれいですよね。	そうですねー 桜川もう綺麗ですね 佐倉のちょっと残念ですけども	9	2	2
新聞は大好きで毎日全部読んでます	そうですかー 偉いですね マジですけど あまり全部は難しいの 良い汗と国政治経済とかは飲みません	12	13	2
高齢者って映画の料金がかかり安いんですね。だから最近よく映画見に行きます。	今料金なー制作円でいつも 1800 円ですけど かなりを安くなっています]	12	13	9

「食事(:10)ー料理(:18)」や「環境(:2)ー外出(:9)」のようなよく見られる組み合わせが比較的多く見られた。その一方で、表 7 の下に示すような「環境(:2)ー仕事(:12)ーメディア(:13)」や「外出(:9)ー仕事(:12)ーメディア(:13)」のようなあまり見られない組み合わせもあることが分かった。

カテゴリの付け方のみ統一し、その他の判断は各学生に任せているため異なった結果が出てしまうことが分かった。またこれは、1 つのカテゴリは複数の様々なキーワードから成り立っているため、その特徴となるキーワードをどのカテゴリと判断するか人それぞれによって異なるために生じた結果である。この結果から、「複数のテキストでカテゴリを表現することが重要である」と考えられ、この問題を解決するためにカテゴリごとの特徴キーワードマップの作成を行うことにしました。

## 5. ユーザプロフィール獲得に向けた 特徴キーワード

### 5.1 ユーザプロフィール

長期間のロボットがユーザの話しているカテゴリを認識し、会話の流れをコントロールするために複数の特徴キーワードから成り立っているカテゴリが必要である。本研究では、図 11 に示すようなユーザプロフィールとして雑談文におけるカテゴリによるプロフィールとそのユーザプロフィールによる各ユーザの特徴キーワードの獲得を行う。

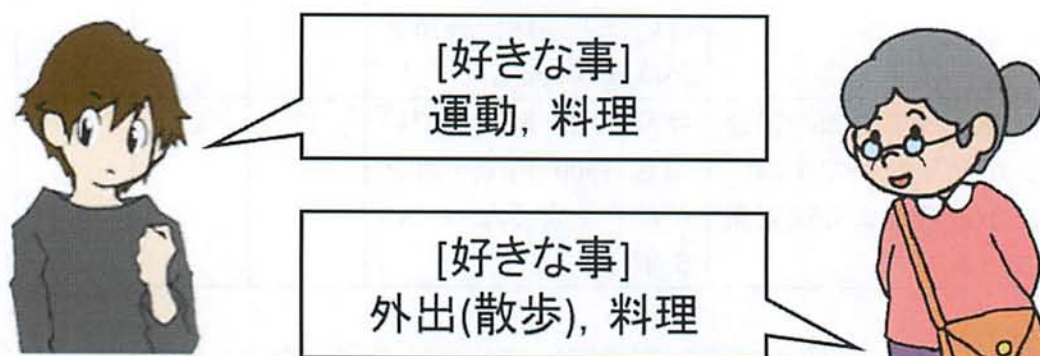


図 11 ユーザプロフィールのイメージ図

## 5.2 カテゴリごとの特徴キーワード

本研究の音声対話ロボットは雑談文を獲得した後，ユーザの発話にどのカテゴリの単語の出現率が高いかを確認し，発話がどのカテゴリに属するかの分類している．そのため，カテゴリごとの特徴キーワードを獲得することでより正確に獲得した雑談文のカテゴリ分類が行えるようになると思う．そこで，カテゴリごとの特徴キーワードマップを 4.2 節と同様に求める．図 12 にカテゴリ「外出」の結果を示す．

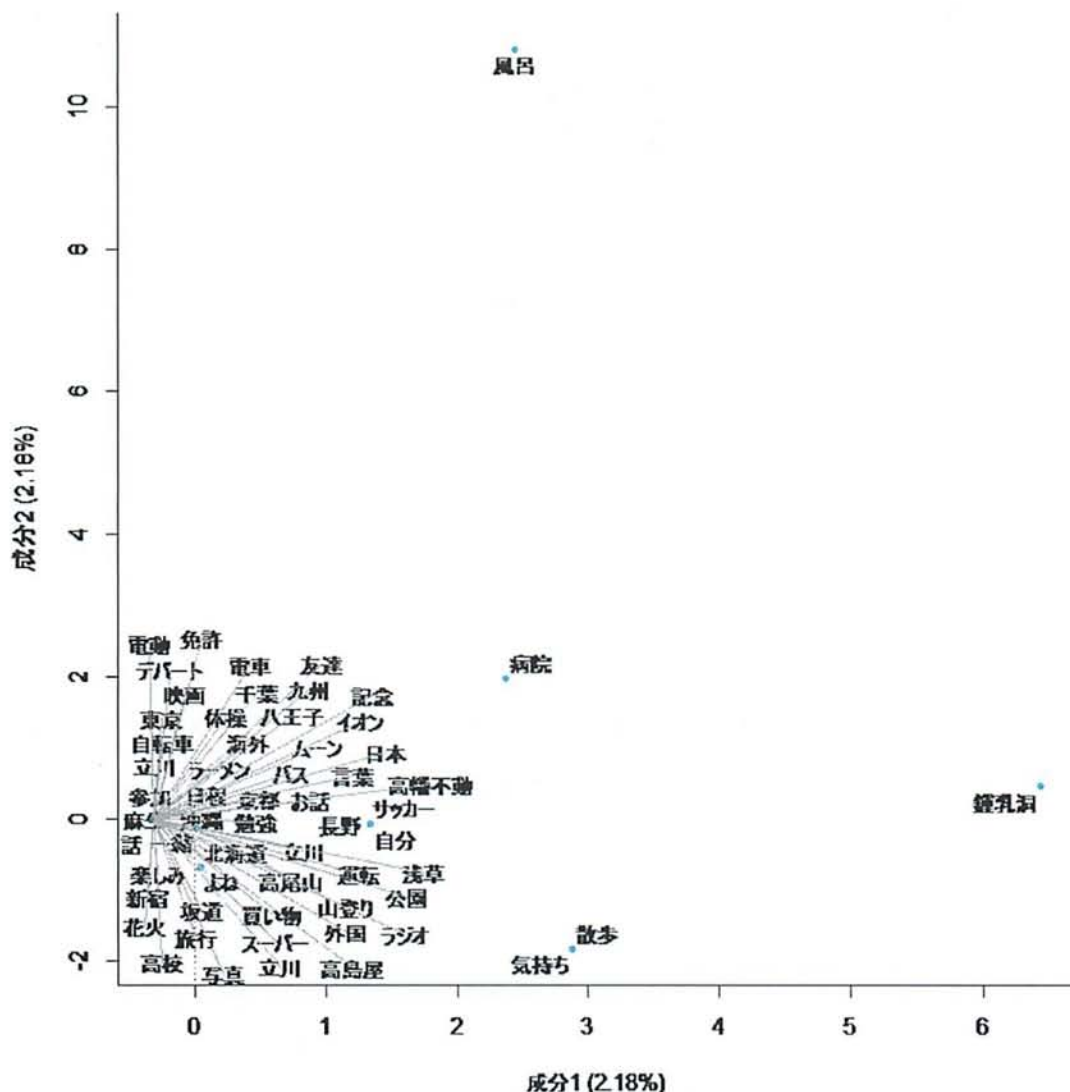


図 12 「外出」の対応分析結果

図 12 より，「外出」の特徴キーワードとして「散歩」「病院」「鍾乳洞」があげられる．また 3.2 節で学生がつけた結果から「環境(:2)－外出(:9)」が間違いやすい組み合わせとして分かった．そこで「環境」の特徴キーワードマップを作

成し、違いが見られるか検証することにした。図 13 に「環境」の特徴キーワードマップを示す。

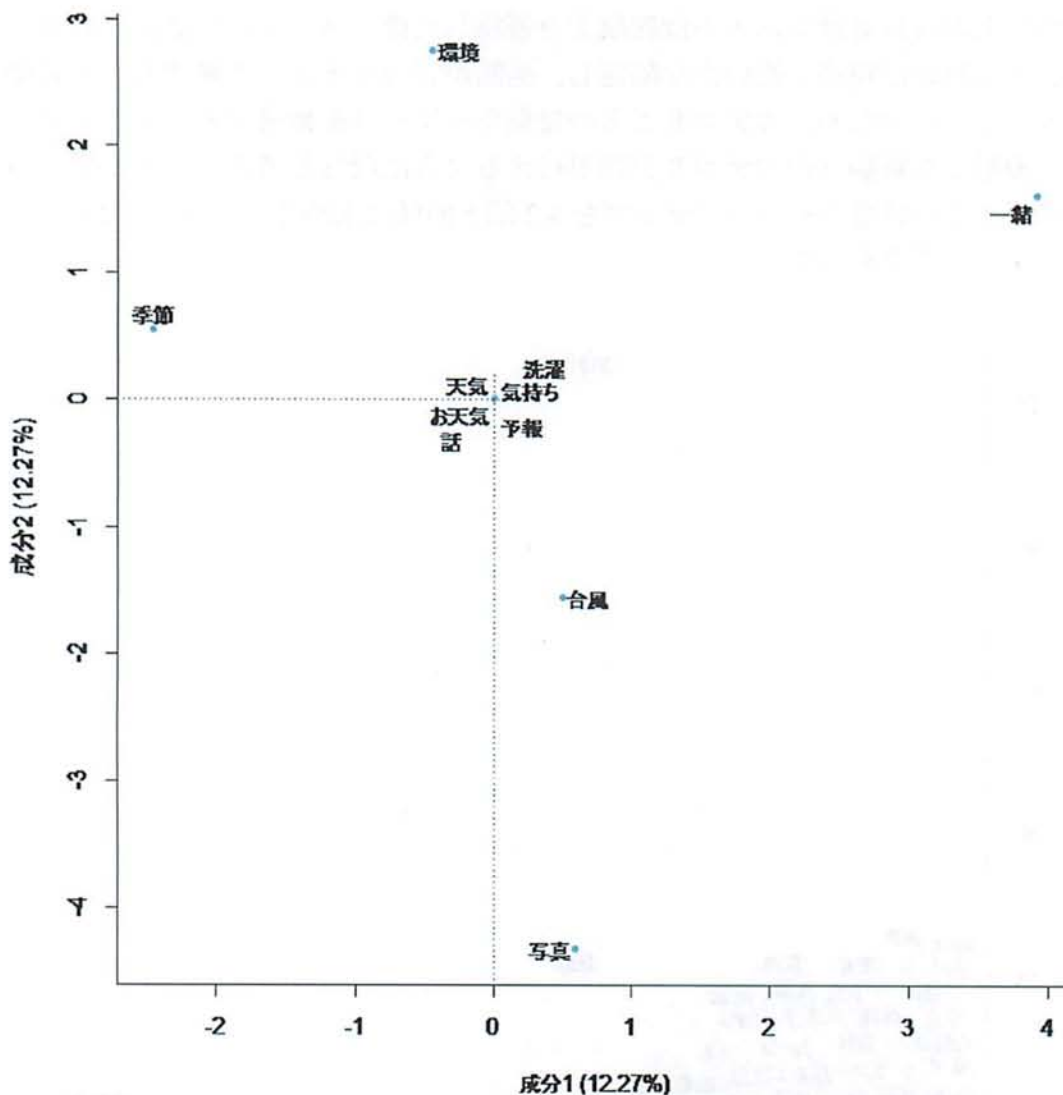


図 13 「環境」の対応分析結果

図 13 より、「環境」の特徴キーワードとして「環境」「季節」「台風」があげられる。このことから、「外出」の特徴キーワードとして「散歩」「病院」「鍾乳洞」が「環境」の特徴キーワードとして「環境」「季節」「台風」を新たに獲得することで互いを判別出来ると考えられる。「外出」の特徴キーワードとして獲得できた「病院」は「健康」の特徴キーワードとも考えられる。このことからカテゴリごとの新たな特徴キーワードを獲得する際に、すべてのカテゴリについて特徴キーワードマップを獲得することで、より分かりやすくカテゴリごとの特徴キーワードを判断することが可能になると考えられる。



また、全カテゴリに属する雑談分を全て対応分析にかけた結果を図 14 に示す。

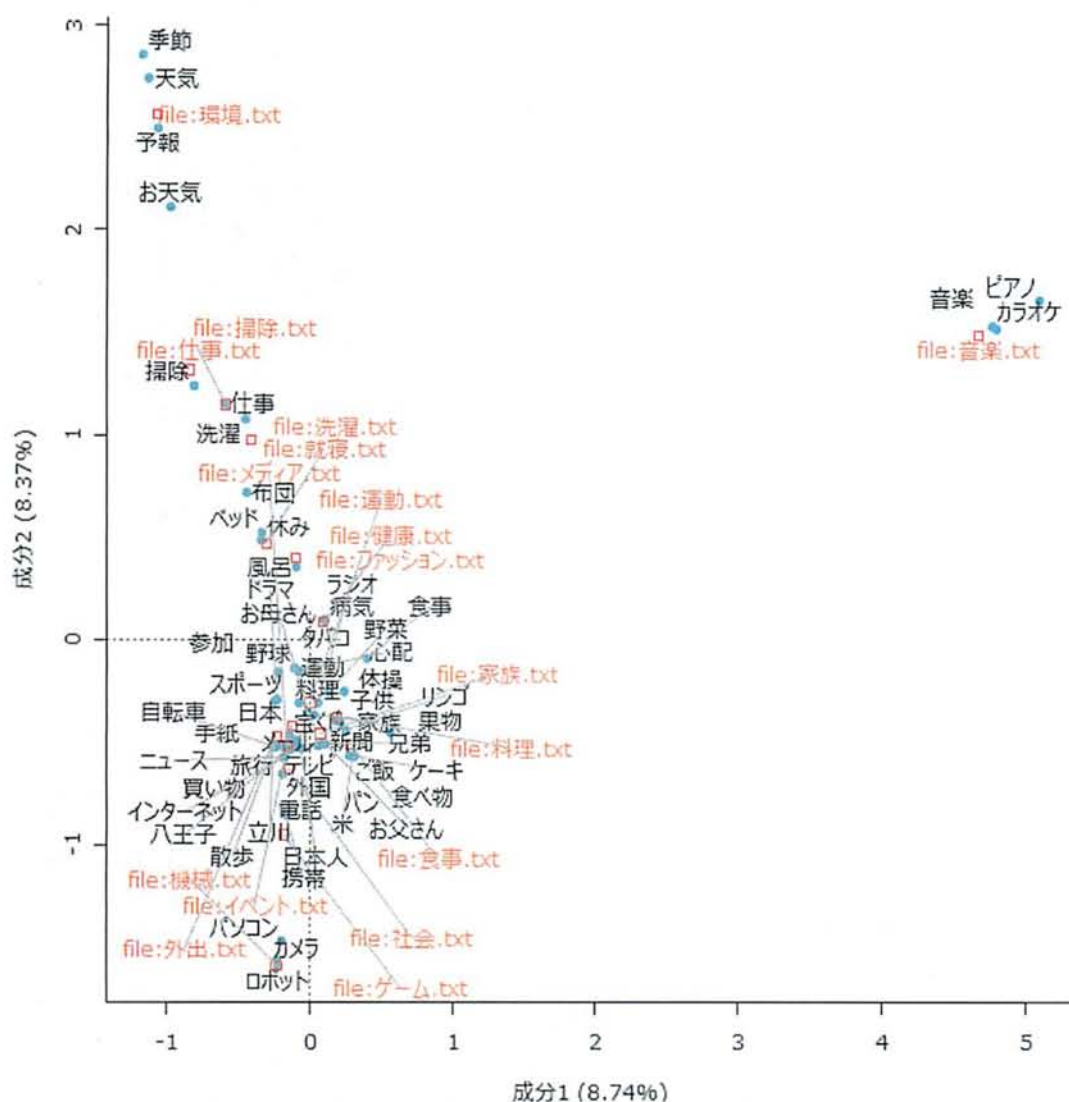


図 14 全カテゴリの対応分析結果

図 14 より、「環境・音楽」のカテゴリが特徴カテゴリ群として獲得できる。そこで、図 15 に「環境・音楽」のカテゴリに属している雑談文を除いた結果を示す。



### 5.3 ユーザプロフィール獲得に向けた特徴キーワードマップ

ユーザプロフィール獲得に向けた特徴キーワードマップを作成するために、本研究では、KHcoderの対応分析を用いて解析を行った。対応分析では内容が近いもの同士が近くに集まり、内容が遠いものが離れて配置される。また、原点から離れたものがより特徴的な内容が含まれている。そのため、この解析を行うことで、原点から離れたものが「特徴キーワード」として考えられる。

本研究では、雑談文内に出てくる名詞を「特徴キーワード」として定義している。そこで、「名詞・サ変名詞・固有名詞・組織名・人名・地名」の6種類を用いて行い、最小出現回数を5回に指定し解析を行った。

実際に獲得した結果として、図16に高齢者1、図17に高齢者6の結果を示す。

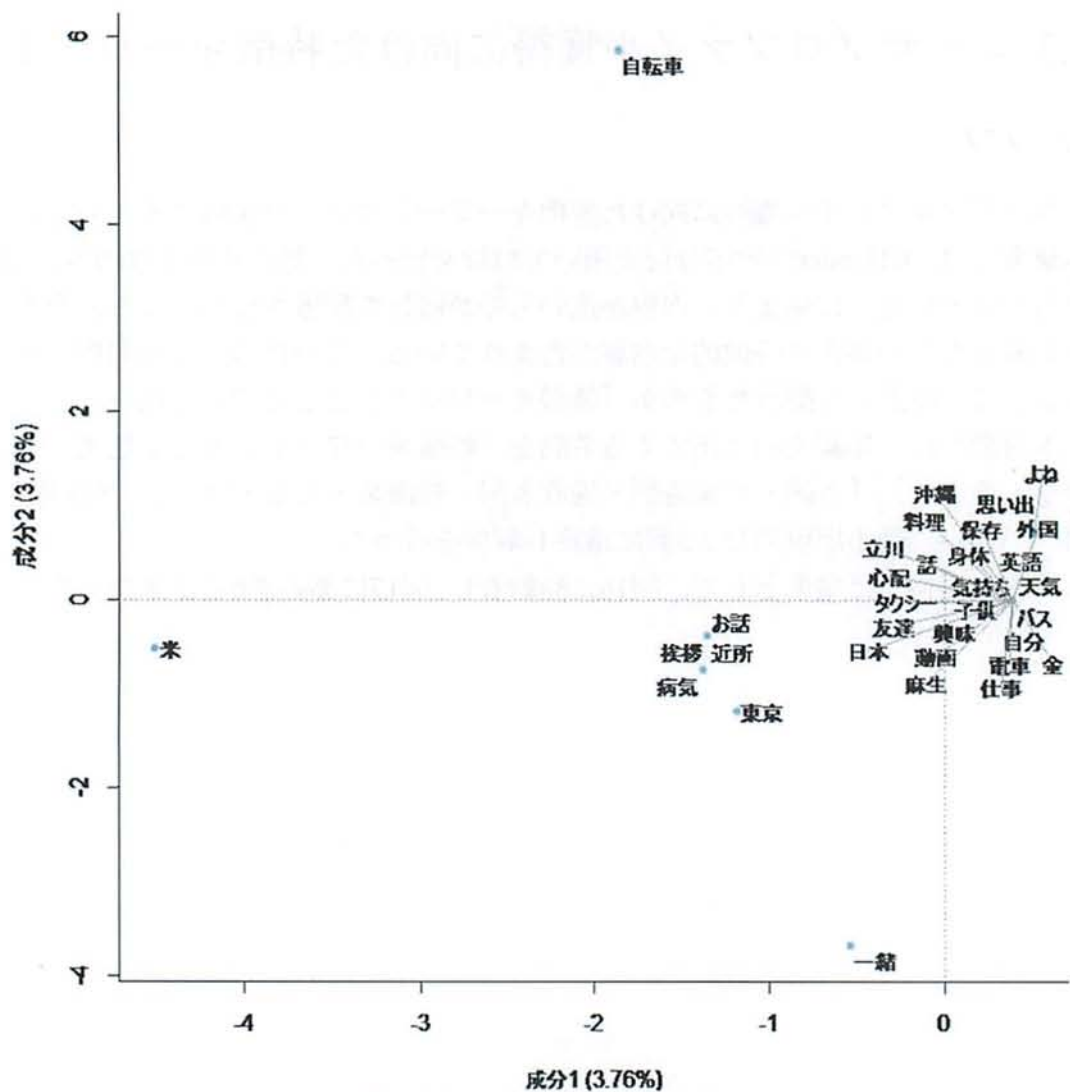


図16 高齢者1の対応分析結果

図16より，高齢者1の特徴キーワードとして「お話・近所・挨拶」や「病気」が獲得できる．これにより高齢者1は「近所付き合いが好き」で「健康に関する話もする」人であると考えられる．



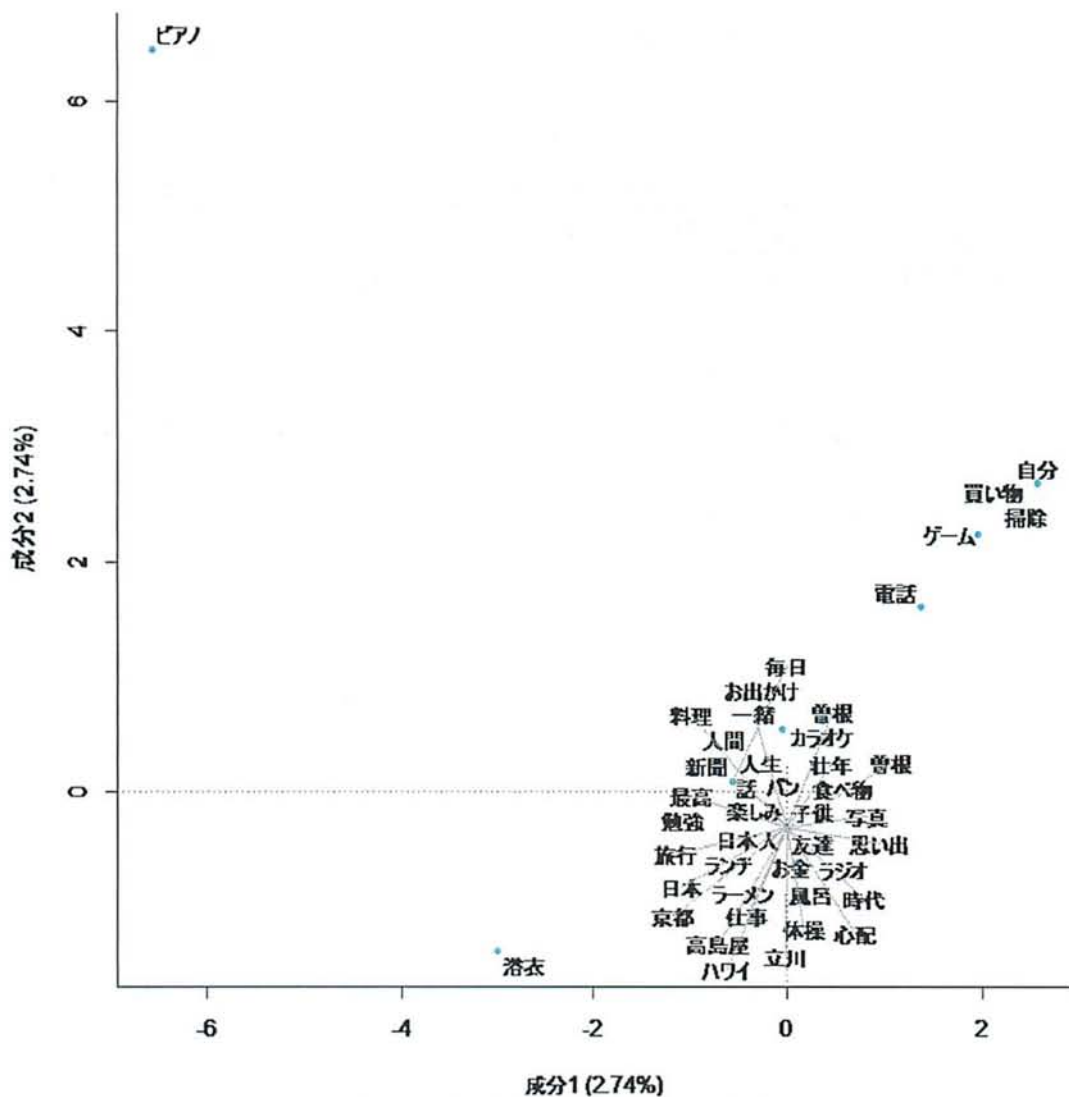


図 17 高齢者 6 の対応分析結果

図 17 より，高齢者 6 の特徴キーワードとして「買い物」，「掃除」，「ゲーム」や「ピアノ」が獲得できる．これにより高齢者 6 は「買い物や掃除など日常生活に関する会話をよくする」で「ゲームやピアノなどの趣味がある」人であると考えられる．このようにユーザごと特徴キーワードを獲得することでユーザプロフィールが行える事を獲得した．

また高齢者に心理テストを行い、「外向性・情緒不安定性・開放性・誠実性・調和性」の5点から高齢者の特徴の獲得を行った。高齢者1・6の結果を表8に示す。

表8 心理テスト結果

	外向性	情緒不安定性	開放性	誠実性	調和性
被験者1	54	45	47	45	46
被験者6	60	52	46	55	50

この結果より、どちらの高齢者も「外向性」が高いことが分かる。よって特徴キーワードマップから獲得できた「近所付き合いが好き」や「買い物や掃除など日常生活に関する会話をよくする」といった他者との関わりをよく行う人物であることが分かり、特徴キーワードマップから獲得できた特徴が有効であることが分かる。

以上のことより、対応分析によって特徴キーワードマップを作成することが、ユーザプロファイル獲得に有効であることが獲得できた。

## 6. おわりに

本研究では音声対話ロボットを用いて高齢者と特定のテーマを定めた会話ではなく雑談を行った。その結果、現在のカテゴリ付けを検証することで、音声対話ロボットのカテゴリ付けがどの程度出来ているかの確認を行うことが出来た。また、雑談履歴からユーザ毎の趣味・趣向などの個人属性であるユーザプロフィールを獲得できることを検証し、特徴キーワードマップを作成することでユーザプロフィールに必要な特徴キーワードを確認することが可能であることを獲得した。

## 参考文献

- [1] 内閣府 “平成 26 年度 高齢化の状況 高齢者の現状と将来像”,  
[http://www8.cao.go.jp/kourei/whitepaper/w-2014/gaiyou/sl\\_1.html](http://www8.cao.go.jp/kourei/whitepaper/w-2014/gaiyou/sl_1.html)
- [2] 松日楽信人, 小川秀樹, 吉見卓, 人と共存する生活支援ロボット, 東芝レビュー, Vol.60 No.7, 2005, pp.112-115
- [3] TOSHIBA “生活支援ロボットの開発について”,  
[http://www.toshiba.co.jp/about/press/2005\\_05/pr\\_j2001.htm](http://www.toshiba.co.jp/about/press/2005_05/pr_j2001.htm)
- [4] 山本大介, 土井美和子, ユーザーと家電機器をつなぐインターフェースロボット, 東芝レビュー, Vol.64 No.1, 2009, pp.28-31
- [5] 横山祥恵, 山本大介, 小林優佳, 土井美和子, 高齢者向け対話インターフェース-雑談継続を目的とした話題提示・傾聴の切替式対話法-, 情報処理学会研究報告, Vol.2010-SLP-80 No.4, 2010
- [6] Hirotaka Aoki, Yasunari Fujimoto, Satoshi Suzuki, Eri Sato-Shimokawara, and Toru Yamaguchi : Physiological Responses on Greeting with Robot under Difference of Culture, Automatic Control of Physiological State and Function, Vol. 1, March 2012
- [7] Joseph Weizenbaum. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM 9, 1 (January 1966), 36-45. DOI=10.1145/365153.365168
- [8] Alicebot <http://www.alicebot.org/>
- [9] Goh Ong Sing, Kok Wai Wong, Chun Che Fung, and Arnold Depickere. 2006. Towards a more natural and intelligent interface with embodied conversation agent. In Proceedings of the 2006 international conference on Game research and development (CyberGames '06). Murdoch University, Murdoch University, Australia, Australia, 177- 183.
- [10] Cheongjae Lee, Sangkeun Jung, Seokhwan Kim, Gary Geunbae Lee, Example-based dialog modeling for practical multi-domain dialog system, Speech Communication, Volume 51, Issue 5, May 2009, Pages 466-484, ISSN 0167-6393,
- [11] Cheongjae Lee, Sangkeun Jung, Kyungduk Kim, Gary Geunbae Lee, Hybrid approach to robust dialog management using agenda and dialog examples, Computer Speech & Language, Volume 24, Issue 4, October 2010, Pages 609-631, ISSN 0885-2308,
- [12] Yang, Z., Levow, G.-A., Meng, H. Predicting user satisfaction in spoken dialog system evaluation with collaborative filtering (2012) IEEE Journal on Selected Topics in Signal Processing, 6 (8), art. no. 6362156, pp. 971-981.



- [13] Noh, H., Ryu, S., Lee, D., Lee, K., Lee, C., Lee, G.G. An Examplebased approach to ranking multiple dialog states for flexible dialog management (2012) IEEE Journal on Selected Topics in Signal Processing, 6 (8), art. no. 6361259, pp. 943-958.
- [14] 近藤 直人, 内田 理, Twitterを用いた個人の特徴抽出とその情報提供ポータルサイト構築への応用, 第20年次大会 言語処理学会(2014年3月)
- [15] 榊剛史, 松尾豊, ソーシャルブックマークとしてのTwitterリスト機能の応用, 2010年度人工知能学会全国大会(第24回) JSAI2010, 3B3-2
- [16] 石井 健一, 新聞記事データからキーワードを自動抽出する試み: 朝日新聞のデータを用いて, ドクメンテーション研究, 33(11), 543~548, 1983.11
- [17] 田中 貴志, 小山 聡, 角谷 和俊, 田中 克己, ユーザ間の対話からユーザプロフィール抽出と動的情報探索, 電子情報通信学会
- [18] 稲葉通将・鳥海不二夫・石井健一郎, 対話からの話者のプロフィール情報自動獲得, 電子情報通信学会, 2011年度人工知能学会全国大会(第25回) JSAI2011, 3C2-OS19-6
- [19] 平成26年度修士論文, 野村 駿“ロボットコミュニケーションのための発話内容のカテゴリ付けと新規カテゴリの獲得”

## 発表文献

- [1] 篠田 遥子, 野村 駿, 李 海妍, 和田 一義, 下川原 英理, 山口 亨,  
"高齢者と音声対話ロボットの雑談履歴の解析"  
日本ロボット学会第 32 回記念学術講演会 (RSJ2014), CD-ROM 講演番号  
3C1-02 (2014/9/4-6).
- [2] Yoko Shinoda, Shun Nomura, Haeyeon Lee, Tomoya Takatani, Kazuyoshi Wada, Eri  
Shimokawara, and Toru Yamaguchi  
"A Dialogue Analysis of Elderly Person with a Chat Robot,"  
2015 RISP International Workshop on Nonlinear Circuits, Communications and  
Signal Processing NCSP'15, Kuala Lumpur, Malaysia, 28PM1-2-2(February 27 -  
March 2, 2015).

## 謝辞

本研究の遂行及び論文の執筆にあたり，終始暖かく，丁寧にご指導をして頂いた首都大学東京大学院 山口亨教授に深く心から感謝申し上げます。

また同大学院 高間康史教授，小町守准教授には本研究をまとめるにあたりご指導いただきましたことを心から感謝いたします。

さらに，本研究を進めるにあたり丁寧かつ的確にご助言をしていただいた首都大学東京大学院 下川原英理助教，多大なるご協力・ご助言をいただいたトヨタ自動車株式会社 李海妍様，高谷智哉様，首都大学東京大学院 和田一義准教授，藤本泰成研究員深く感謝をいたします。

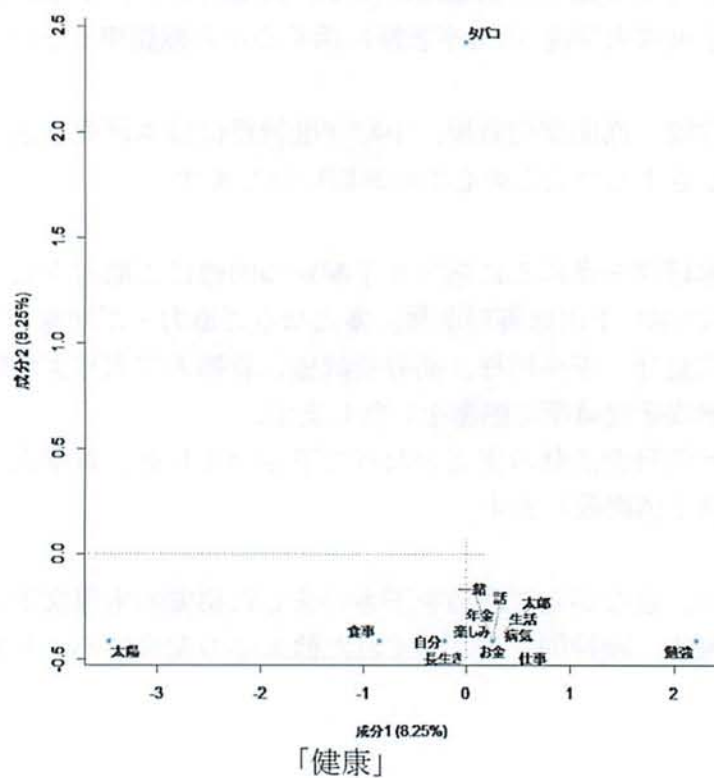
そして日々の研究活動の支えとなって下さいました，首都大学東京山口研究室の皆様に深く感謝致します。

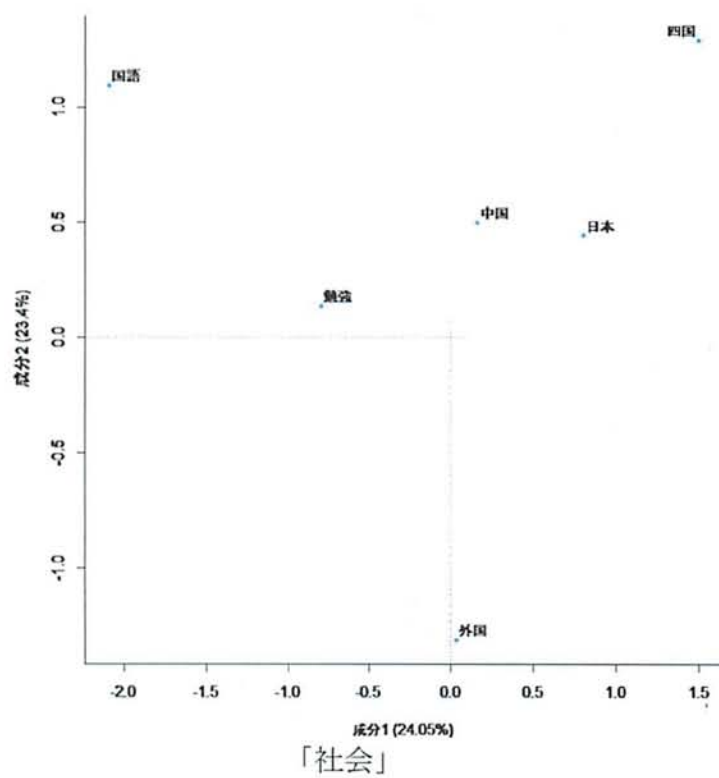
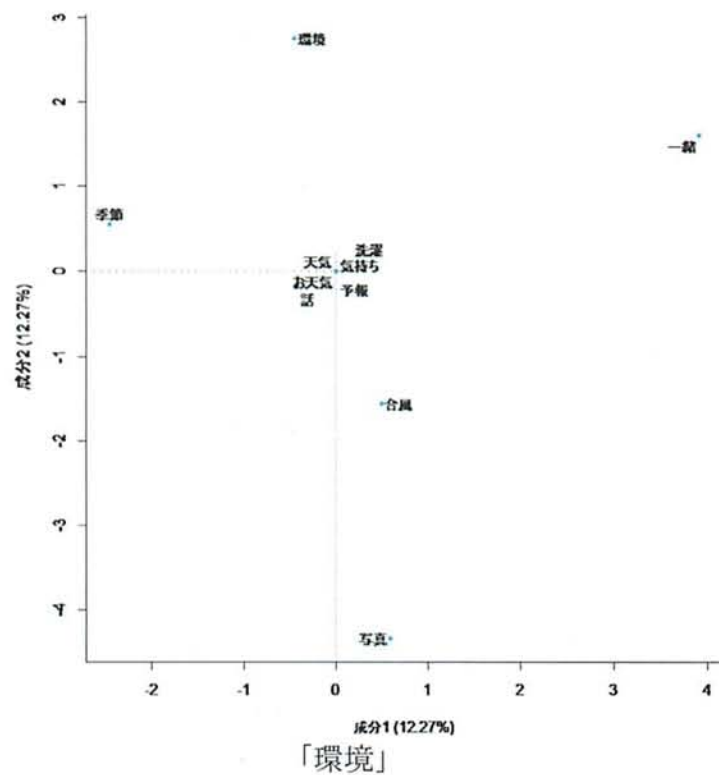
また最後に，陰ながらご助力を下さいました秘書の木原文子さん，生活を支えてくれた家族，精神的に支えてくれた他大学の友達に心より感謝の意を表します。

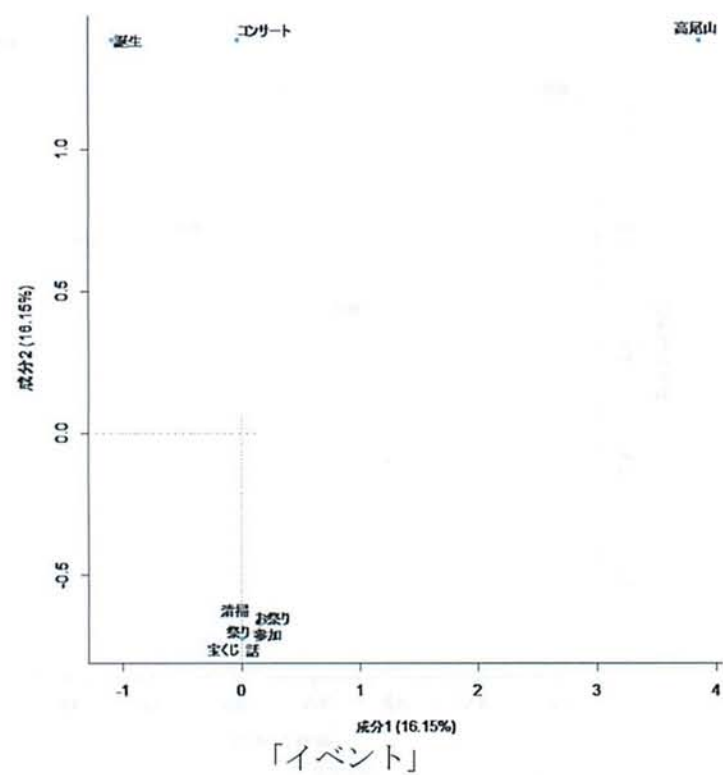
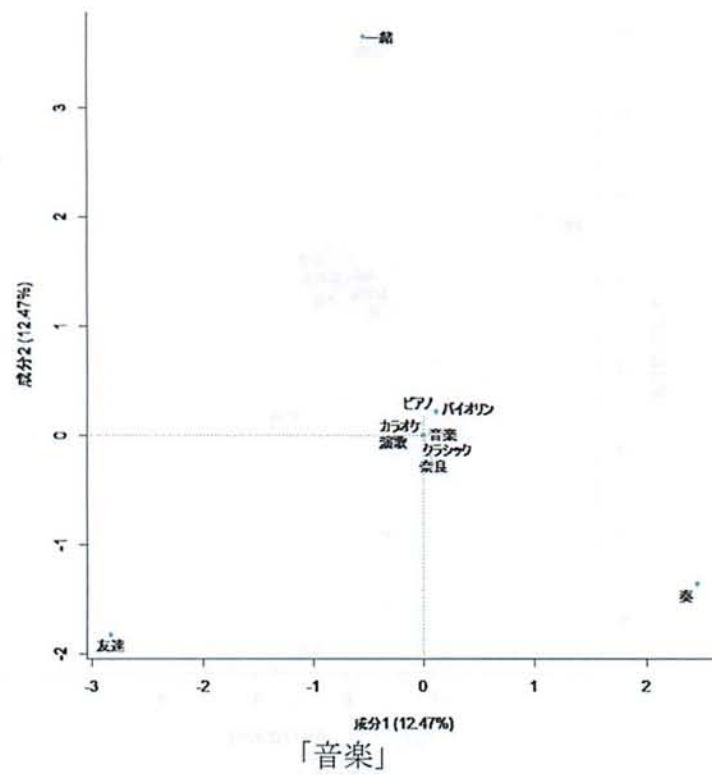


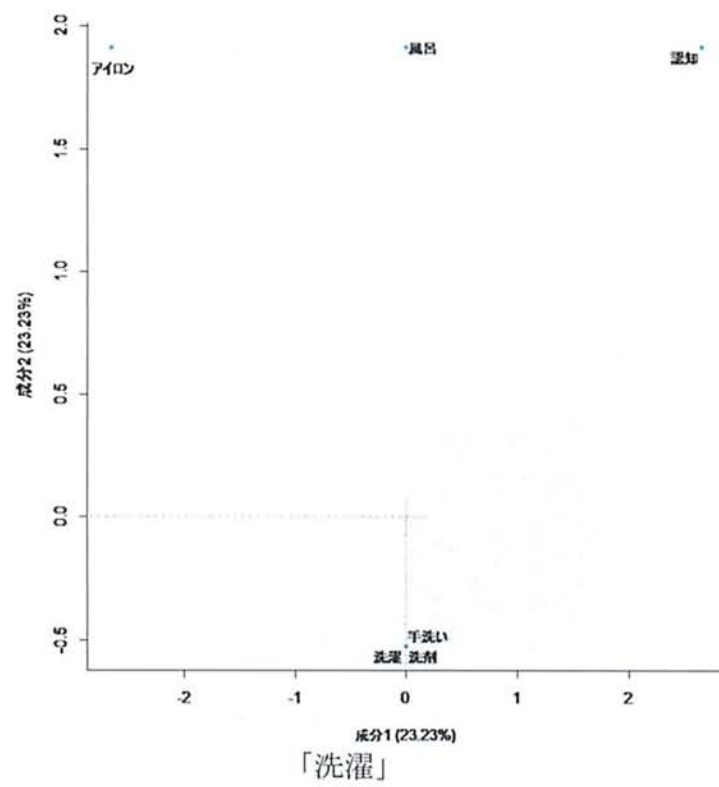
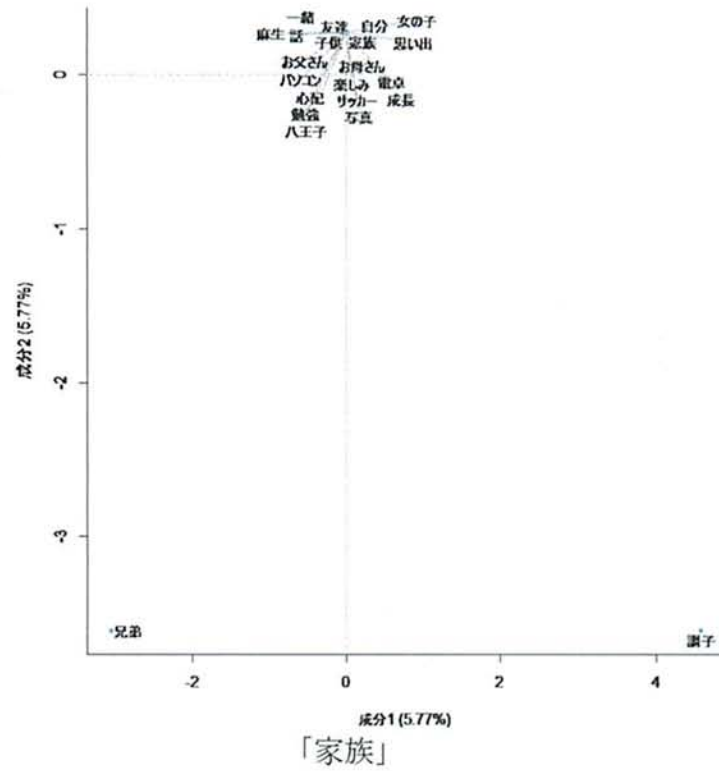
## 付録

### カテゴリごとの特徴キーワードマップ



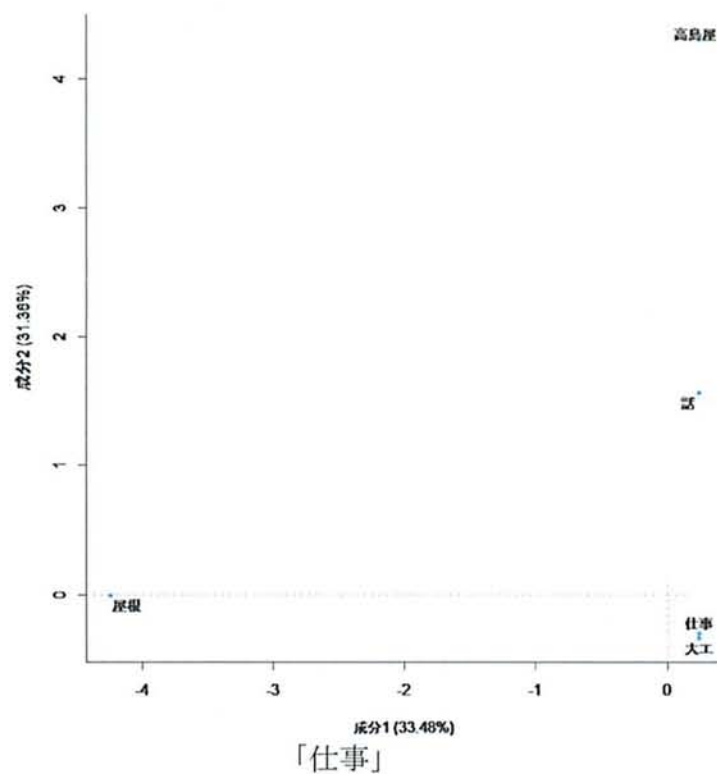
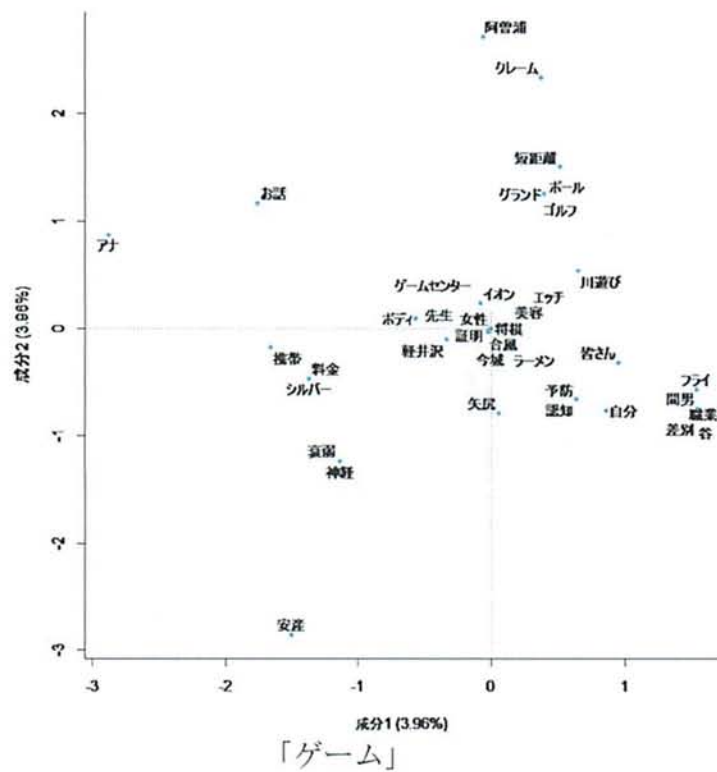


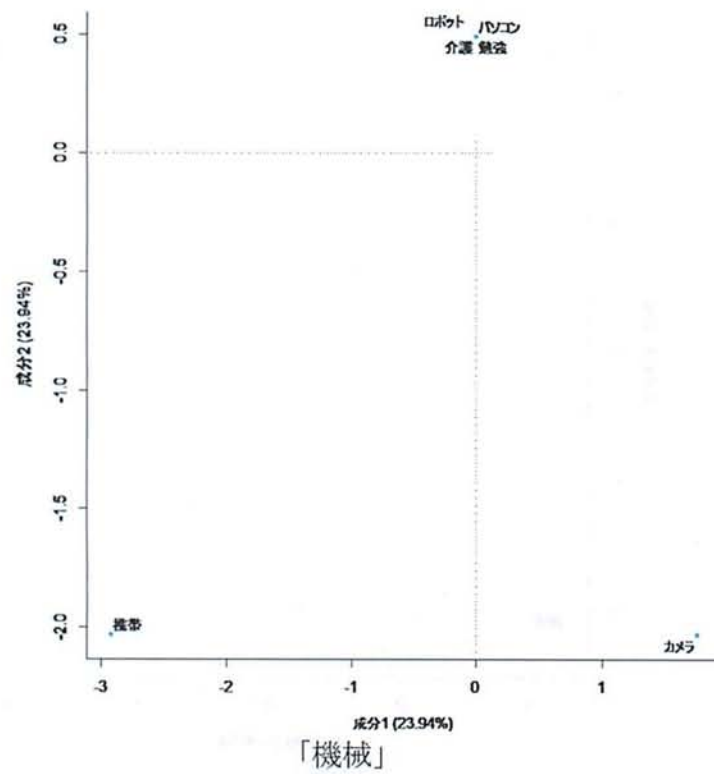
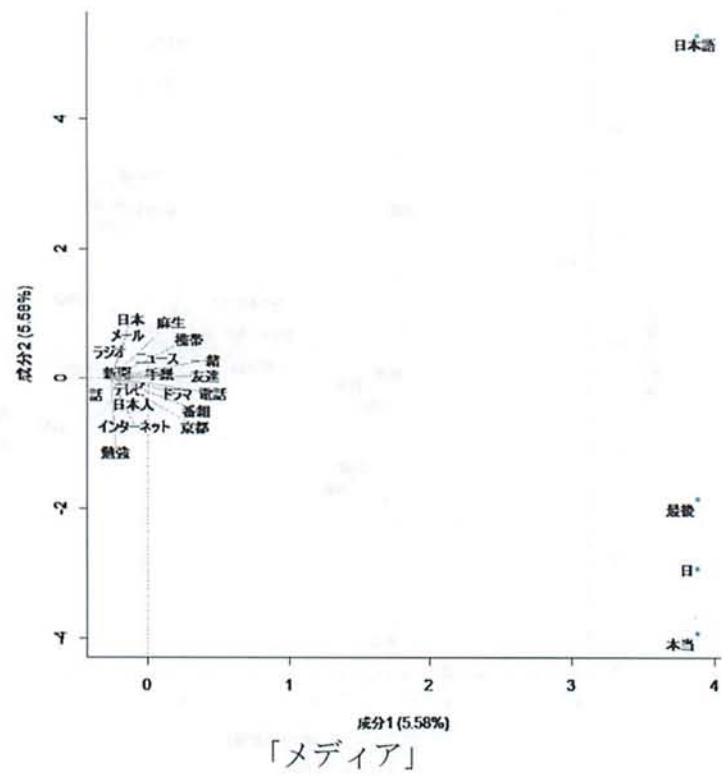


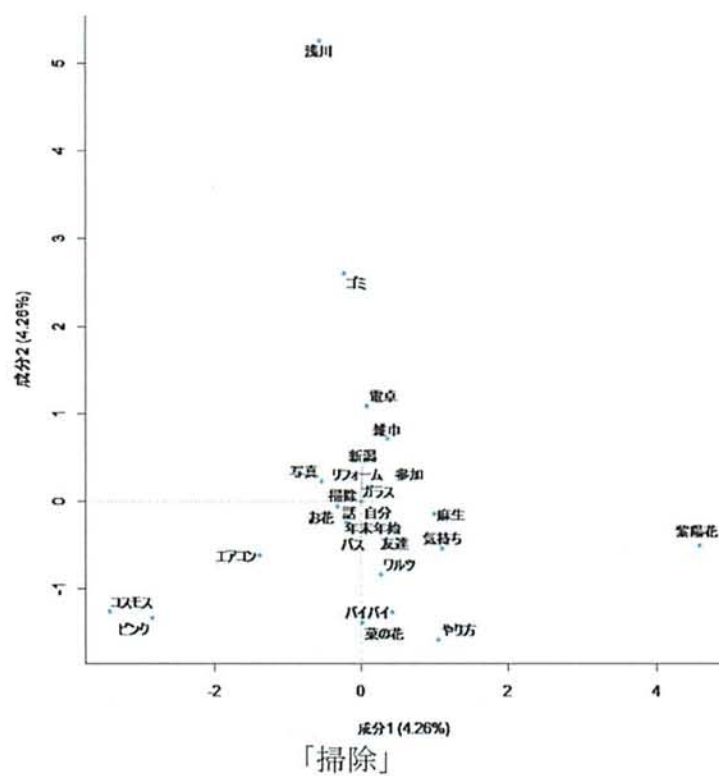
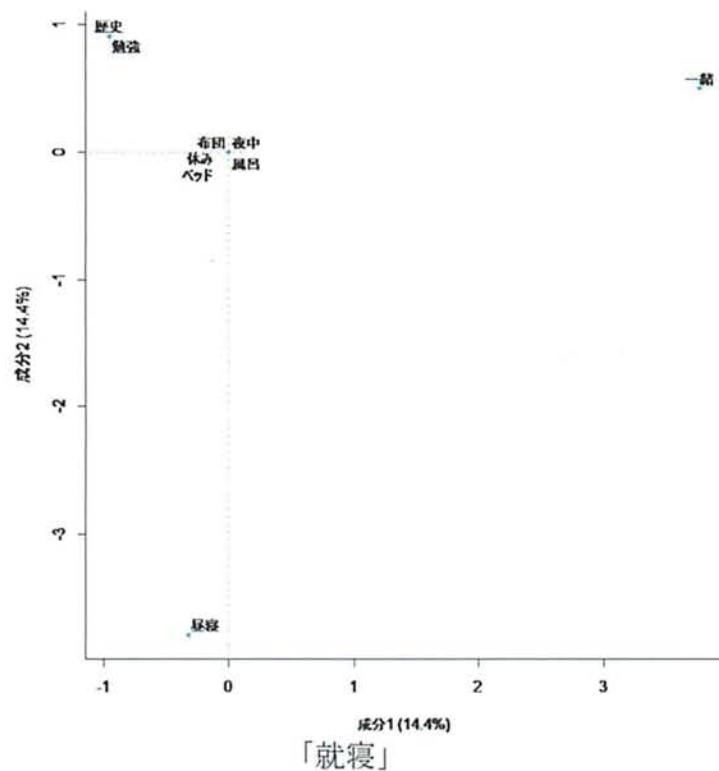




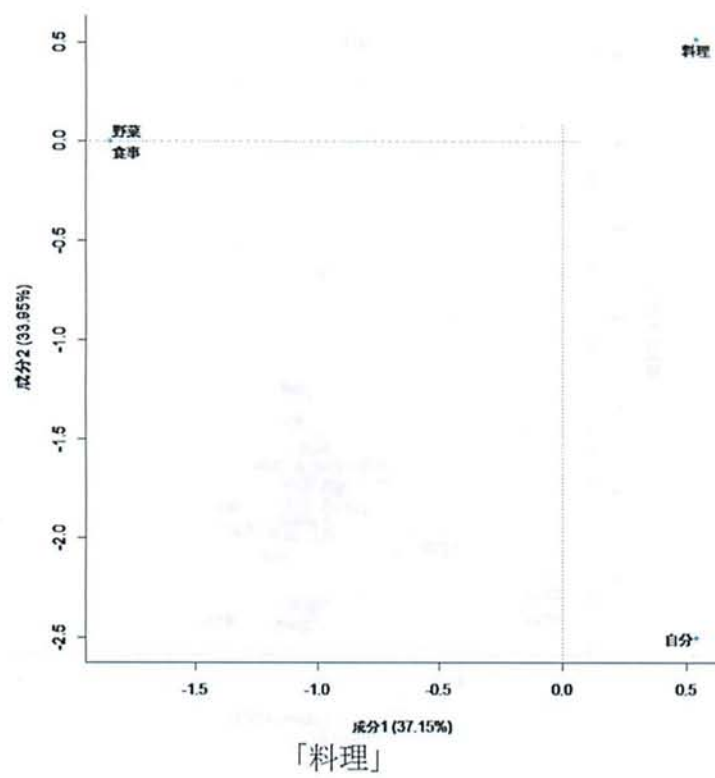
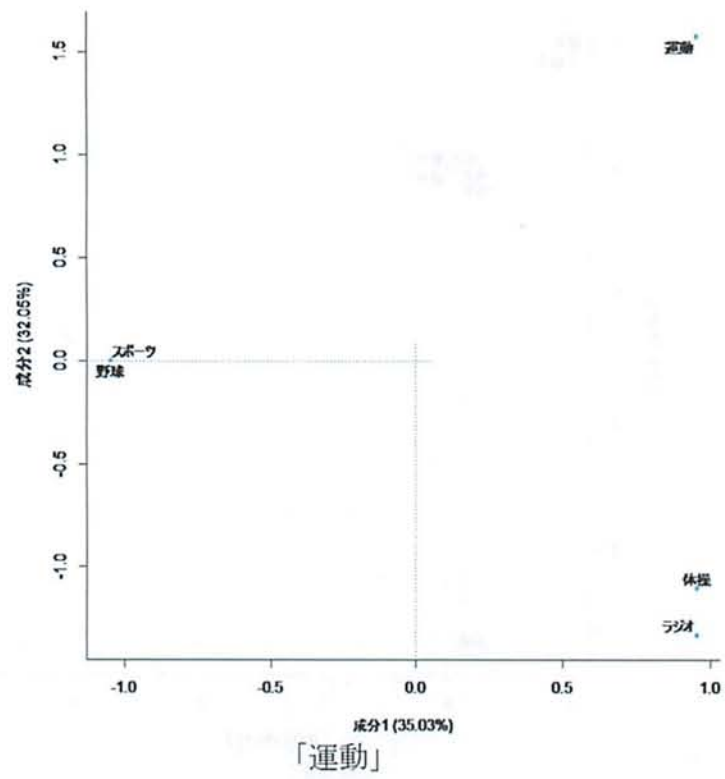




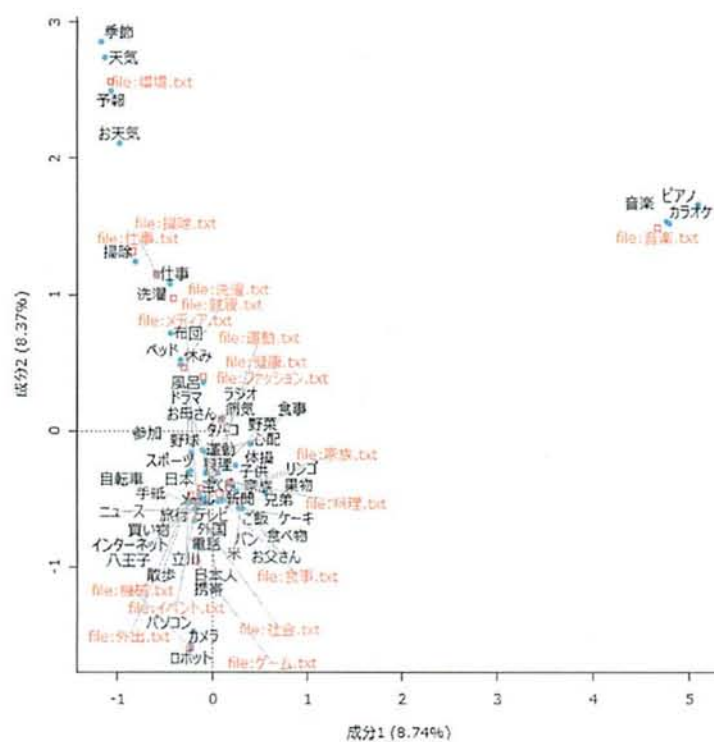




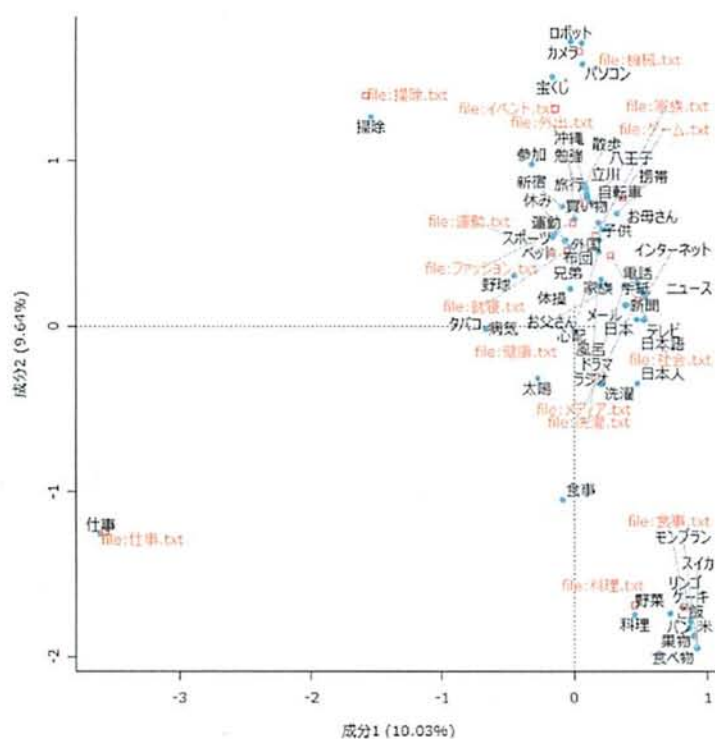




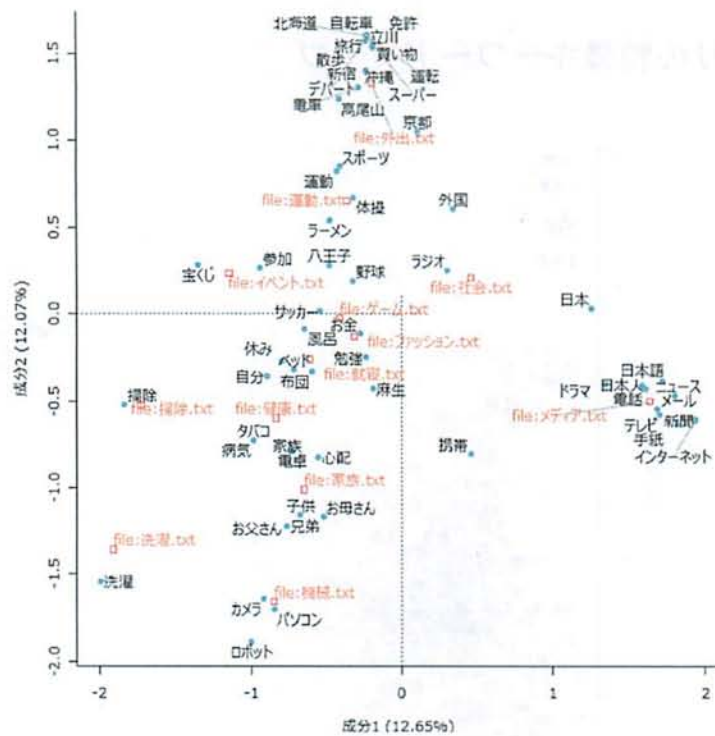
## 18 カテゴリの特徴キーワードマップ



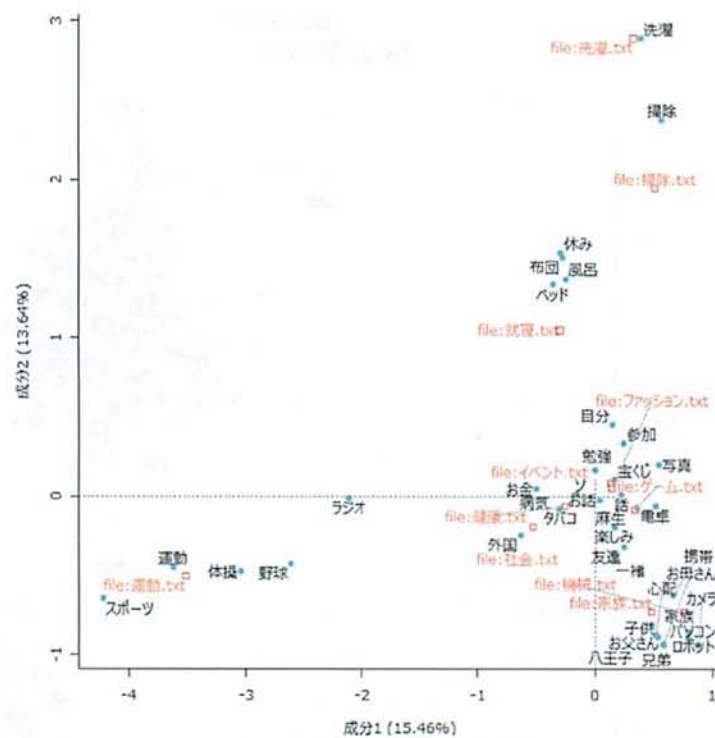
18 カテゴリ



18 カテゴリー「環境&amp;音楽」



16 カテゴリー「仕事&食事&料理」

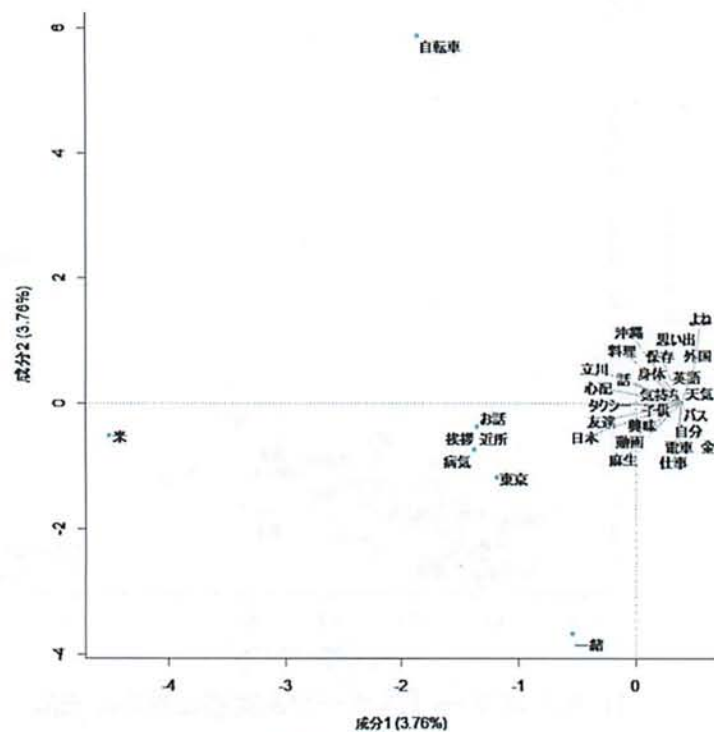


13 カテゴリー「メディア&外出」

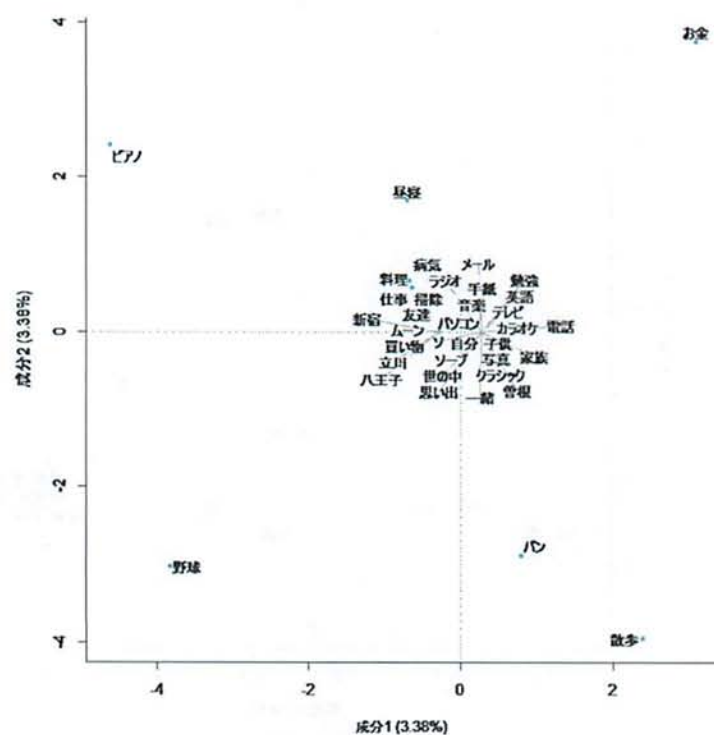




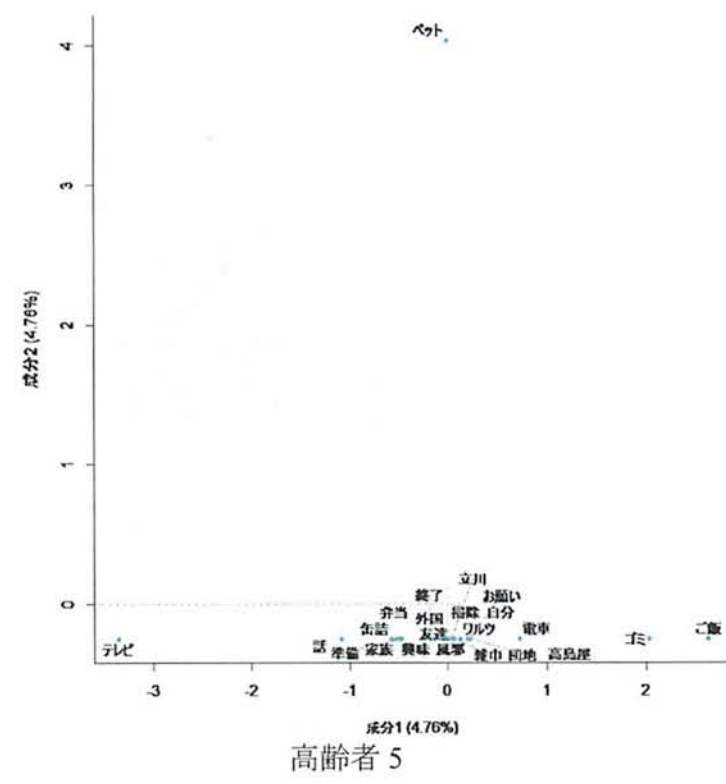
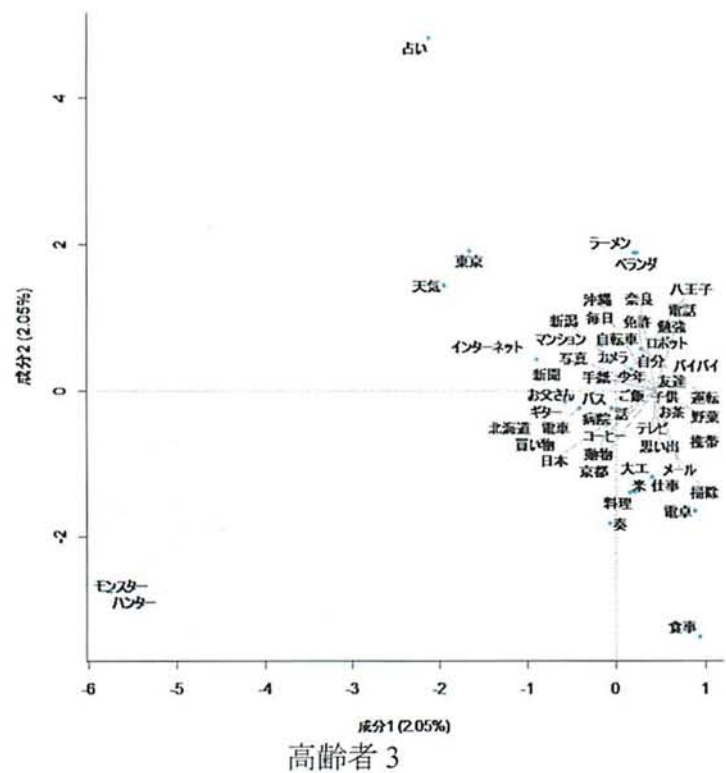
## 高齢者ごとの特徴キーワードマップ

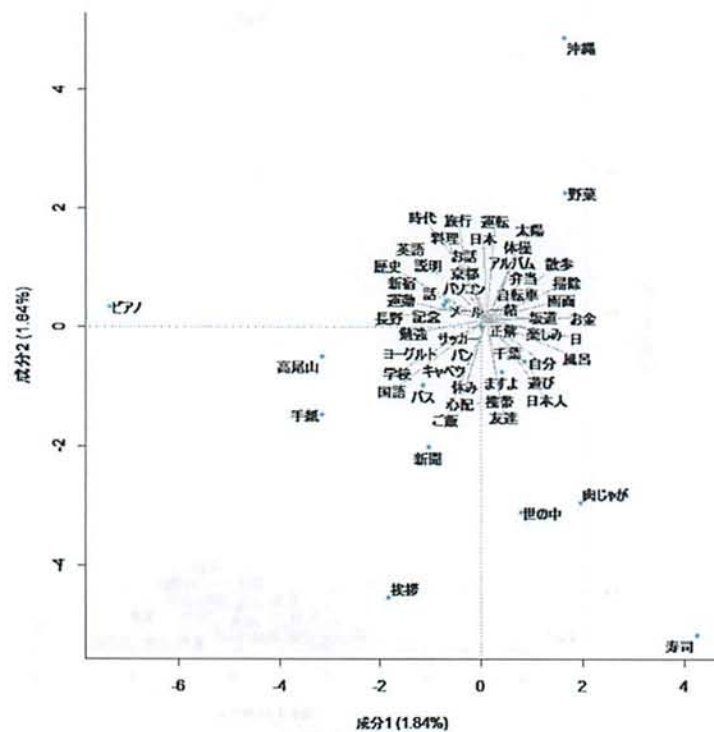
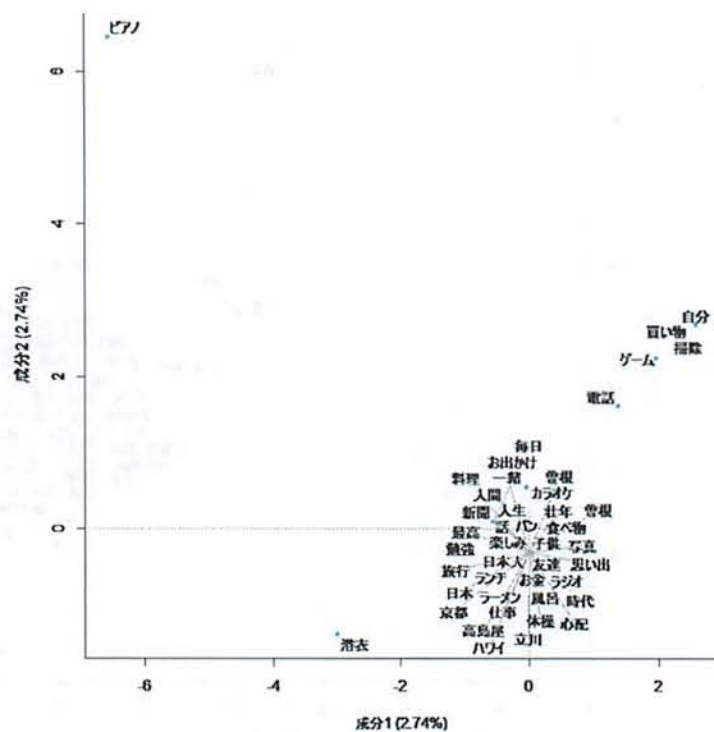


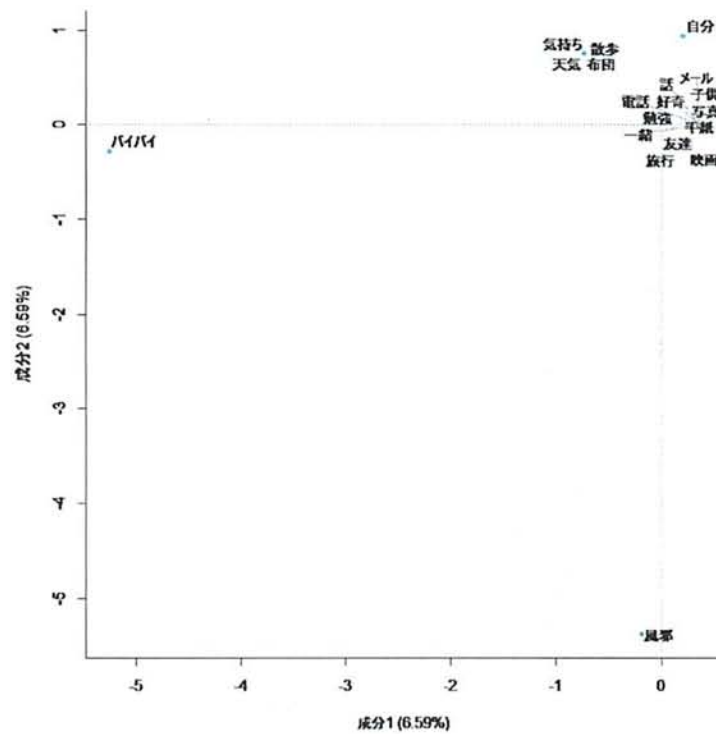
高齢者 1



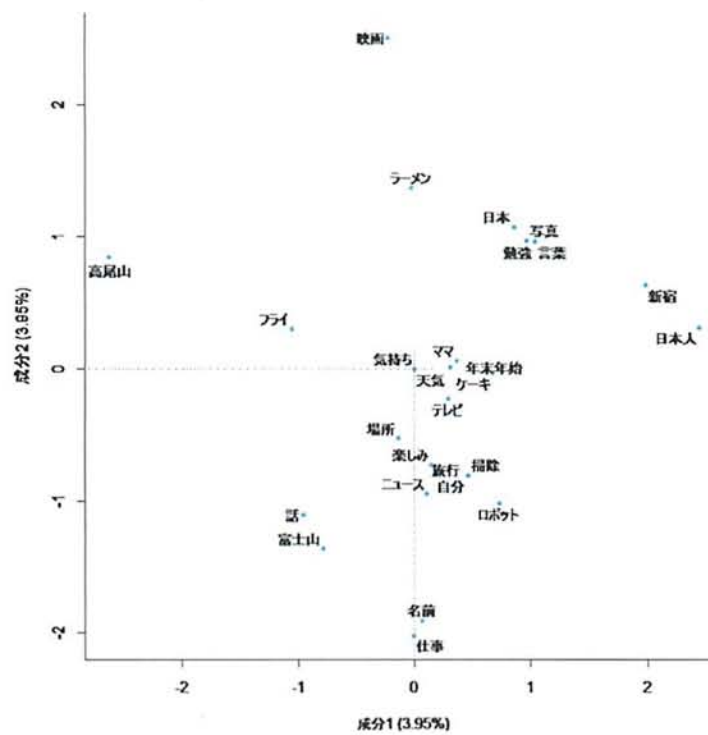
高齢者 2





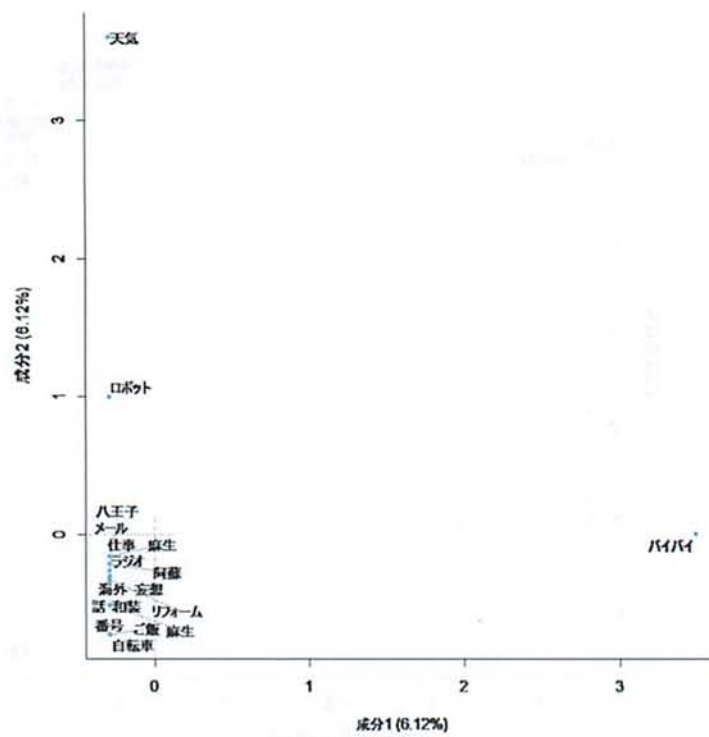


高齢者 8



高齢者 9





高齢者 10