

# テレビ番組に含まれるキーワードに基づく発話文生成手法

金子豊<sup>†1</sup> 星祐太<sup>†1</sup> 村崎康博<sup>†1</sup> 上原道宏<sup>†1</sup>

**概要:** 人と一緒にテレビを視聴する自律型のコミュニケーションロボットの開発を進めている。本報告では視聴中のテレビ番組に含まれているキーワードを基に、ロボットが自発的につぶやくためのテンプレートに基づく発話文生成手法を提案するとともに、その評価結果について述べる。提案手法では、過去のテレビ字幕文から「食べたい」「見たい」などの願望を表す動詞句、「美しい」「怖い」などの形容詞句を含む文を、発話文を生成するためのテンプレート文として抽出しておく。そして、テレビの視聴中には番組の字幕文に含まれるキーワードを抽出し、テンプレート文と組み合わせることで主観的な感情表現を含む発話文を生成する。テンプレート文の選択は、キーワードと、テンプレート文に含まれる動詞、形容詞との共起性を用いる。提案する手法を約6年分のテレビ番組の字幕文を用いて実験を行った。

**キーワード:** 発話文生成, テレビ視聴ロボット, テレビ字幕文, テンプレート, Word2vec, 感情表現

## Utterance Generation based on Keywords of TV Programs

YUTAKA KANEKO<sup>†1</sup> YUTA HOSHI<sup>†1</sup> YASUHIRO MURASAKI<sup>†1</sup>  
and MICHIIHIRO UEHARA<sup>†1</sup>

### 1. はじめに

家族や友人と一緒にテレビを視聴することは、単独で視聴する時には得られない効果がある。例えば、人が一緒にいることで、笑いの表情を促進する効果があるように[1]、他人の嗜好的な発言や行動は、気にもかけなかった事への気づきや、共感的なうれしさを与える効果があると考えている。

一方、平成27年国勢調査によると、日本の人口は減少傾向にあるのに対し、世帯数は増加している。1世帯当たりの平均人数は2.33人であり、数としては単独世帯、二人世帯の順となっている[2]。このため、複数人でテレビを視聴する機会が減っていると考えられる。我々のアンケート調査では、単独世帯の22%がテレビを複数人で見たいと回答している[3]。

高齢者に目を移すと、年齢の上昇とともにテレビ視聴時間が長くなると報告されており[4]、高齢者の約80%が「テレビが情報源」と回答している[5]。また、生活の中で充実感を感じる時として、「家族団らん」49%、「雑談」43%という調査結果がある[6]。一方で、高齢者の8%、単独世帯の高齢者に絞ると約25%が、「会話頻度は2~3日に1回以下」という調査結果[7]もあることから、「家族団らん」や「雑談」を望んでいても、実際には一人でテレビを見ている高齢者がある一定数いることが推察される。

現在、ロボット掃除機などのロボット家電、スマートスピーカー、高齢者向けロボット、ペット型ロボットなど、人とのコミュニケーション機能を持ったデバイスが家庭に

普及し始めている。今後、これらのデバイスとの対話機会が増えることが予想される。例えば、健康管理のためのセルフモニタリング機能をロボット掃除機に組み込んだ実験では、セルフモニタリング機能以外でも人からロボットへ話しかける行為が認められ、その頻度は年齢があがるにつれ多くなる結果が示されている[8]。

このようなロボットとのコミュニケーションでは、必ずしも複雑な対話を生成しなくても様々な効果が期待できる。例えば、ロボットから幼児的な言葉足らずの発話を用いることで、聞き手の新たな解釈や、能動的な解釈を引き出す可能性[9][10]や、2択疑問文のような答えやすい発話をするすることで、人が対話システムを受け入れやすくなる効果[11]などが報告されている。また、ロボットからの話題の提示[12]や、価値観を持った発話[13]は、ロボットとの対話を継続させる効果があることが報告されている。

以上のような背景から、我々は人と一緒にテレビを視聴するロボット(テレビ視聴ロボット)により、テレビを楽しく視聴する機会を増やすことを目標に、研究開発を進めている[14]。

テレビ視聴ロボットに搭載する機能の一つとして、視聴中の番組に関するちょっとしたつぶやきや、主観的な感情表現を含む言葉を周囲の人に向かって自発的に発話する機能を検討している。この発話機能による効果として、周囲の人の同士の会話のきっかけや、気づき、発話行為の誘引などを期待している。また、気まずさの改善や、周囲をなごませる効果なども期待している。

本稿では視聴中のテレビ番組に含まれているキーワードを基に、ロボットが自発的につぶやくための発話文生成手法を提案する。提案手法では、過去のテレビ字幕文から

<sup>†1</sup>NHK 放送技術研究所  
NHK Science & Technology Research Laboratories.

「食べたい」「見たい」などの動詞句や、「美しい」「怖い」などの形容詞を含む文を、テンプレート文として抽出しておく。そして、視聴中の番組キーワードとテンプレート文とを組み合わせることで、感情表現を含む発話文を生成する。組み合わせるテンプレート文は、基となるキーワードとテンプレート文内の感情表現語の共起性から選択する。

以下、2章では関連研究について、3章では提案する発話文生成手法について、4章ではテレビ字幕文を使った実験結果について述べ、5章でまとめる。

## 2. 関連研究

非タスク型対話システムとして様々なものが開発されている[15]。中でも「女子高生 AI りんな」は200万以上のユーザを獲得している[16]。このような雑談対話システムでは、単独の発話文生成手法にこだわらず、様々な手法を併用することで、ユーザを飽きさせることなく長期間利用し続けさせることができると考えられる。

発話文生成手法は辞書やモデルなどを構築するための学習時のデータと、文生成時のデータと密接に関係している。図1に発話文生成手法の分類と利用データとの関係を示す。

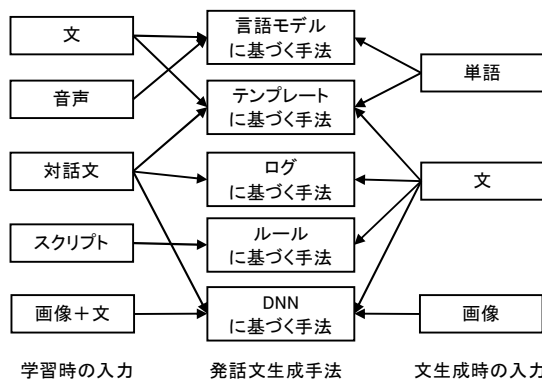


図1 文生成手法の分類と利用データ

言語モデルに基づく手法は、多数の学習用の文から  $n$  個の単語の接続を示す統計的な単語  $n$ -gram を言語モデルとして作成する手法である。 $n$  が高次になるほど、出力される文は自然になる反面、指数的に探索範囲が大きくなることや、生成文の確率が0になるデータスパースネスの問題もあるため、最大でも  $n=5$  といわれている[17]。 $n$  を可変長とする HPYLM (階層 Pitman-Yor 言語モデル) を応用した手法[18]や、HPYLM を利用し、連続音声から直接言語モデルを学習する手法[19]なども開発されている。また、ニューラルネットワークを利用した、Neural Probabilistic Language Model[20]や、RNN LM[21]などがある。

テンプレートに基づく手法では、学習用の文の中から、名詞や動詞などの一部を穴埋めテンプレート文として蓄積しておき、それに穴埋めすることで文を生成する。この手法では、テンプレートの生成方法、適切なテンプレートや置換語句の選択方法などが課題となる。例えば、SVM (サ

ポートベクターマシン) を利用して発話タイプを推定することで対話テンプレートを選択する手法[22]や、入力文の焦点語を抽出し、その焦点語と係り受け関係の語とを置換語句とする手法[23]などがある。

ログに基づく手法は、SNS などの対話文を蓄積しておき応答文に利用する手法である。この手法では、SNS などの大量の発話文の中から適切な対話文を抽出することが課題の一つとなる。例えば、話題語に対する正解発話と不正解発話に出現する単語の分布が異なるという仮定に基づいて話題語に関する発話文を抽出する手法[24]や、ユーザ発話の感情極性(Positive, Negative, Neutral)を判定し、あらかじめ極性で分類したツイート群の中から極性と同一文を選択する手法[25]、対話のメインテーマの連続性と、対話文の意味的な関連性から文を選択する手法[26]などがある。

ルールに基づく手法は、スクリプトに従って、入力文から出力文を生成する手法である。この手法は、テンプレートに基づく手法と併用されることが多く、タスク指向型対話システムで多く使われている。非タスク型対話システムの元祖とされる ELIZA[27]はこの方式に分類される。

DNN(Deep Neural Network)に基づく手法は、対話文や、画像から文を直接出力する End-to-End な方式である。特に、Sequence-to-Sequence(Seq2Seq)モデル[28]は、数多く検討されている。例えば、文献[29]では、Twitter の対話データによる学習後、特定のテレビ番組の字幕文による転移学習を行うことで、対話応答文に特定のキャラクタスタイルを付与している。文献[30]では、ユーザの過去数回の発話のコンテキストによって応答を選択する方式であり、過去のテレビ番組の会話を学習データとして用いている。さらに、文献[31]では、画像からキャプション文を生成する手法が提案されている。入力された画像は CNN (畳み込みニューラルネットワーク) により特徴量を求め、その特徴量と説明文を対して Seq2Seq モデルを学習させている。

テレビの視聴中に関連する発話を出力する方法として、Yahoo!リアルタイム検索のテレビ連携機能[32]や、ニコニコ実況[33]など、SNS の発言を利用する方法がある。また、SNS のコメントを用いてロボットに発話させることで、人の発話機会の増加を目指した検討がされている[34]。しかし、SNS の発言を用いた手法は、番組によって発言量が異なること、不特定多数の発言であるため一貫性のある発話にならないこと、誹謗中傷などの発言が多いことなど、直接ロボットの発話に用いるには課題が多い。

感情表現を含む文を生成する手法として、文献[35]の雑談対話システムでは、印象マイニング[36]を用いてニュース記事の興味語から印象値を求め、例えば、印象値として「悲しみ」と推定された場合には、「私も悲しみを感じます」のような感情表現を含む応答文を生成している。文献[37]では、感情生起表現辞書を構築し、入力文の感情判別を行い、それと同一の感情表現のテンプレート文を用いること

で入力文に対するレビュー文を生成している．感情判断システム[38]を用いて，物語文章から感想文を自動生成する手法[39]や，既存の文を模倣して新しい文を自動生成する手法[40]なども試みられている．また，発話文の自動生成の実現に必要な話し手の特徴と文型との関係を整理した辞書の構築が行われている[41]．

### 3. キーワードに基づく発話文生成手法

我々の検討しているテレビ視聴ロボット[14]では，テレビ番組を視聴中に，ロボットがテレビに向かってつぶやいたり，周囲の人へ話しかけたりする機能を検討している．例えば，ラーメンを紹介している番組を視聴中に，「このラーメンおいしそうだな」とつぶやいたり，「今度このラーメンを食べに行きたいね」と周囲の人へ話しかけるため，主観的な感情表現を含む発話文の生成を目標としている．

テレビ視聴時の発話生成に用いることができるテレビからの情報としては番組字幕文がある．しかし，2章で述べたように，発話文生成手法の多くはシステム対話を想定したものであり，システム構築には多くの対話データが必要となる．テレビ番組に対する発話文の生成を考えた場合，対話データに相当する対話対のデータが存在しない．SNSのコメントを利用することが考えられるが，字幕文から文を生成するためには，字幕文などのテレビの情報とコメントの対応をとったデータを構築する必要があり，現状は困難である．

そこで，提案手法では，視聴中の番組字幕文からキーワードを抽出し，そのキーワードから文を生成することとした．文の生成手法はテンプレートに基づく手法とし，テンプレート文も過去の字幕文から，感情表現を含む文を抽出する．

#### 3.1 提案手法の概要

図 2 に提案する発話文生成手法の動作概要を示す．本手法では，あらかじめ，発話の元となるキーワード辞書と，代表語辞書を準備しておく．代表語辞書は，代表語とその同義語で構成される．ここでは，願望を表す「動詞連用形 + "たい"」[42]（以降，動詞句）および形容詞のうち，事前に決めた語を代表語と呼ぶことにする．例えば，「食べたい」「行きたい」などが動詞句の代表語，「好き」「大きい」などが形容詞の代表語である．また，動詞の代表語に対して同義語を登録しておく．例えば，動詞句の代表語「食べたい」に対して，「食いたい」「食べちゃいたい」などが同義語である．

次に，キーワード辞書と代表語辞書を使い，過去の字幕文から，代表語ごとのテンプレート文を抽出しておく．テンプレート文は，過去の字幕文の中から，代表語辞書とキーワード辞書に登録された両単語を含む文を抽出し，キーワード語の部分に置換箇所とする．抽出したテンプレート文は，その代表語のテンプレート文として保管する．

テレビ視聴中は字幕文からキーワードを抽出し，そのキーワードと代表語との共起性からキーワードに近い代表語を選択し，その代表語のテンプレート文とキーワードを合成することで発話文を生成する．キーワードの代表語の共起性の判断として word2vec[43]による単語ベクトルのコサイン類似度を用いる．

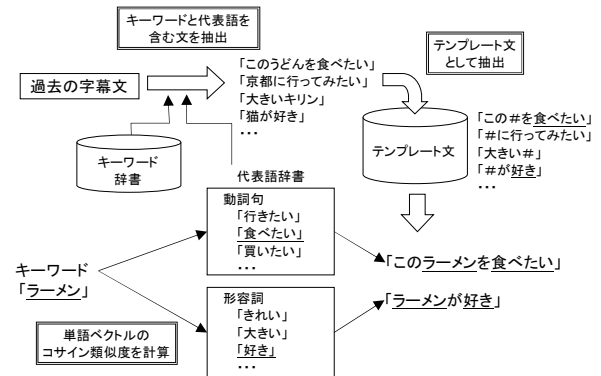


図 2 字幕に基づく発話文生成手法の動作概要

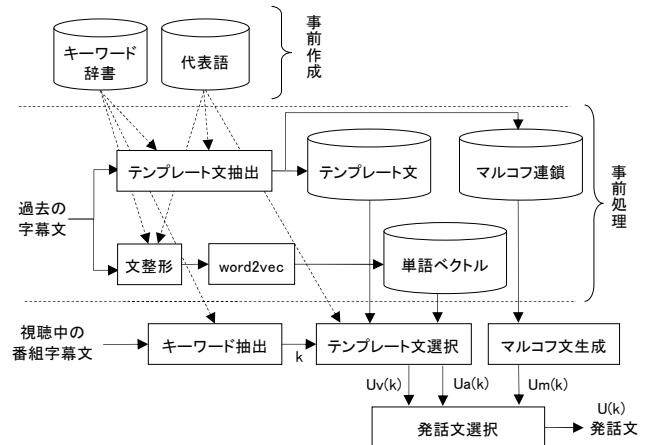


図 3 提案する発話文生成システム

#### 3.2 発話文生成システム

試作した発話文生成システムの構成を図 3 に示す．

試作システムでは，入力された番組字幕文からキーワード  $k$  を抽出し，そのキーワード  $k$  にコサイン類似度が近い動詞句と形容詞句の代表語を選択し，そのそれぞれの代表語に対して登録されているテンプレート文の中からランダムに 1 つを取り出し，動詞句から生成した文  $U_v(k)$  と，形容詞から生成した文  $U_a(k)$  を生成する．

また，試作システムでは，字幕文から生成したマルコフ連鎖の辞書(単語 2-gram)を使い，文  $U_m(k)$  を生成する．マルコフ連鎖による発話文生成は文献[44]のアルゴリズムを参考にした．

#### 3.3 単語ベクトルの生成

word2vec による単語ベクトルを生成するため，字幕文の単語分割を行う．始めに形態素解析(辞書 NEologd[45]を用いた MeCab[46])で分かち書き処理をする．代表語の動詞句が分割された箇所と，キーワード辞書に登録された語が分

割された箇所は、その部分を統合する。次に、動詞句が代表語の同義語な場合は、代表語の動詞句に置換する。例えば、「おしゃべりしたい」はその代表語の「話したい」に置換する。この整形処理を行った分かち書き文を word2vec の単語ベクトルの学習に用いる。

#### 4. 実験

##### 4.1 過去のテレビ字幕文を使用した事前処理

テレビの字幕データは、多くの放送局で年々増加しており、2016年のNHK総合テレビの総放送時間に閉める字幕時間の割合は84%となっている[47]。字幕文を基に発話文が生成できると、放送チャンネルを問わず多くの番組に対応することができる。我々が実験用に保管[48]しているNHK総合テレビの2016年の1日の字幕画面の平均は12,622画面、1画面の平均文字数は18文字であった。

キーワード辞書、代表語辞書、テンプレート文の生成には、2010年10月から2016年9月までの約6年分のテレビ番組の字幕文(NHK4チャンネル、民放5チャンネル)を用いた。人がロボットの発話を理解できる1回の発話長は、一画面に表示される字幕文と同程度と仮定し、20文字以内の字幕文44,693,884文を用いた。代表語には動詞句として66語、形容詞として23語を抽出した。

番組情報、NHKニュースの字幕文から抽出した名詞、Wikipediaの見出し語[49]とはてなキーワード[50]との共通の見出し語などから抽出した約15万語のキーワード辞書を構築した。word2vecによる単語ベクトル構築時のパラメータは、Skip-gram, size=200次元, window=8語, negative=25個とした。構築した単語ベクトル行列に含まれる単語数は227,728語、そのうち、キーワード辞書に含まれる単語は47,460語であった。

付録Aに今回使用した代表語とその同義語、代表語を含む字幕数、代表語のテンプレート数を示した。代表語を含む字幕文数に対して約6~7%がテンプレート文として抽出された。これは、利用した字幕文の0.17%であり、今後、字幕文からさらに多くの有効なテンプレート文を抽出することが課題である。

キーワードに対する共起性の強い代表語を調べた。全キーワードに対してコサイン類似度が大きい代表語の上位3個をカウントした。カウントした代表語の度数を図4に示す。動詞句、形容詞の上位10語の累積度数はそれぞれ34%、63%となっており、若干上位の代表語にキーワードが集中している傾向がある。

付録Bに各代表語に対して共起性の強い上位10個のキーワードを示した。動詞句の代表語では、「注目したい」「調べたい」「知りたい」など、どのようなキーワードでも文にしやすい代表語を除くと、「釣りたい」「読みたい」「育てたい」「飲みたい」「乗りたい」など、対象物が明確なものは、キーワードがうまく抽出されている。一方、「帰りたい」「あ

げたい」「当てたい」「取りたい」などの代表語では、代表語に合わないと思われるキーワードが多く抽出されている。形容詞は、比較的どのような語ともマッチする傾向があるため、キーワードは比較的よく抽出されているようであるが、「元気」「高い」「速い」などの代表語では、代表語に合わないと思われるキーワードが多く抽出されている。

##### 4.2 発話生成文の妥当性評価

2017年4月19日(水)の1日の字幕文15,610文を対象とし、そのうち5分以内に同じキーワードが抽出された場合を除いた8,767個のキーワードに対する発話文を生成した。1キーワードに対して、 $U_v(k)$ ,  $U_a(k)$ ,  $U_m(k)$ の3文を生成し、その中から妥当と思う文を1つ選択した。回答には3文とも妥当でないという選択肢を含めた。結果を表1に、回答例を図5に、生成文の例を図6示す。

表1 文選択の実験結果

	$U_v$ 選択	$U_a$ 選択	$U_m$ 選択	すべてなし	生成文なし
選択数	1,816	1,891	1,969	2,607	484
割合(%)	20.7	21.6	22.5	29.7	5.5
cos 類似度 avg. (動詞句)	0.303	0.294	0.294	0.287	
cos 類似度 avg. (形容詞)	0.267	0.267	0.267	0.266	

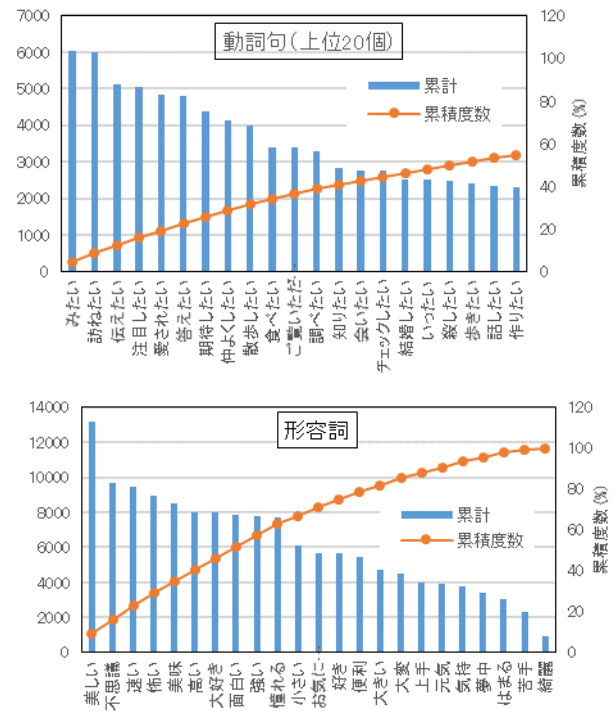


図4 全キーワードの共起性が強い代表語の度数  
 生成文  $U_v$ ,  $U_a$ ,  $U_m$  が選択された個数は、それぞれ全体の20.7%, 21.6%, 22.5%であり合計64.8%であった。これは、1キーワードから3候補文を生成し、その中から適当な文が選択できれば、約65%が妥当な発話となることを示している。

$U_v$  が選択された場合のコサイン類似度の平均値が0.303に対し、選択されなかった場合のコサイン類似度が0.287

であり、選択された文のコサイン類似度が大きい傾向が見られた。一方、Uaにはその差は見られなかった。この結果から、コサイン類似度の値から生成文の妥当性を評価することは困難であると考えられる。

この実験では5.5%のキーワードに対して文が生成できなかった。これは、事前処理に使った字幕文内にそれらのキーワードが含まれていなかったためである。これを改善するには、キーワード辞書や単語ベクトルを随時更新する必要がある。

動詞句の生成文が選ばれた(Uv 選択)例

うどん	Uv	○	今の時期うどん食べたいですね。
	Ua		絶景と美味うどんの極上湯宿
	Um		うどんだし。

形容詞の生成文が選ばれた(Ua 選択)例

千利休	Uv		飛び切りに愛されたい千利休。
	Ua	○	これははまるでしょうナイス千利休。
	Um		千利休みたいな臭いだ。

マルコフ連鎖の生成文が選ばれた(Um 選択)例

スマートフォン	Uv		スマートフォンも続けるし子供も産んで育てたい。
	Ua		山から下りるのはスマートフォンが便利。
	Um	○	スマートフォンも欲しい。

すべて良くない例

マニキュア	Uv		着たいなあのマニキュア
	Ua		やはり上手を取るとマニキュアですか
	Um		マニキュアみたいだよ濃いです

図 5 発話生成文評価の回答例

適当と判断した生成文	不適当と判断した生成文
<ul style="list-style-type: none"> <li>ここで1つ注目したいのが<b>原子力</b>空母です</li> <li>ところで<b>エクスポローラー</b>っていついかに</li> <li><b>フィギュアスケーター</b>も引き続き応援したいですね</li> <li><b>駐屯地</b>に入りたいです</li> <li>一生<b>居酒屋</b>で働きたいと思ってる</li> <li>絵を描きたいと<b>グラフィックデザイナー</b>を目指すも断念</li> <li><b>定食屋</b>で暮らしたい</li> </ul>	<ul style="list-style-type: none"> <li>味わいたいです<b>フレンチ</b>草</li> <li><b>メーカー</b>で一発作りたいと思います</li> <li>誰も見たことのない<b>カメラ</b>が撮りたい</li> <li><b>アナーキー</b>見たいね俺も回るとこ</li> <li>また<b>時系列</b>に書きたいんだって</li> <li><b>ヤマネコ</b>釣りたいです</li> <li>自分を持ちたいなら<b>二等辺三角形</b>になれ</li> <li><b>商社マン</b>で働きたいと考える外国人が参加する</li> <li>いやあ<b>バンドエイド</b>みたいな価値はないだろ</li> </ul>
(1) 動詞句による生成文 Uv	
<ul style="list-style-type: none"> <li>ここに驚きの<b>アフリカ</b>の不思議があるのです</li> <li><b>ショウガ</b>にしちゃ<b>小さい</b>な</li> <li><b>リュウキュウアユ</b>は美しいですね</li> <li>不思議な<b>イギリス</b>専門店</li> <li><b>青森空港</b>を楽しむのに便利な地図</li> <li>あーあ<b>パン</b>も冷凍できたら便利なのに</li> <li>実は今<b>東京都</b>にはまる外国人が急増中</li> </ul>	<ul style="list-style-type: none"> <li><b>法律事務所</b>の気持も</li> <li>坂上お気に入りの具は<b>ケージ</b>のミソ炒め</li> <li><b>ジョイント</b>は常備しておくとも便利</li> <li><b>コミュニケーション</b>絵は上手なんだ</li> <li><b>クレープ</b>君は何か好きな作品ございますか</li> <li><b>専門職</b>の子供が憧れる職業といたら</li> <li>あなたの大好きなお母さまの<b>フリーアナウンサー</b></li> <li><b>出資者</b>は2球目が大好き</li> </ul>
(2) 形容詞による生成文 Ua	
<ul style="list-style-type: none"> <li><b>コンソメ</b>味が好きな</li> <li><b>ホームページ</b>見ても面白いかも</li> <li><b>容疑者</b>に対し強い姿勢が面白いけどね</li> <li><b>陶磁器</b>って高いんです</li> <li><b>クレープ</b>大好きです</li> <li><b>シリアス</b>映画も見たい</li> <li><b>フィクサー</b>みたいなやつですか</li> <li><b>中国</b>も色々大変だしね</li> <li><b>フランス</b>に行きたいです</li> <li><b>牛乳</b>あんまり好きじゃなくてね</li> </ul>	<ul style="list-style-type: none"> <li><b>オープン</b>がないときにアンバーと話したいですね</li> <li><b>マニア</b>しか見ないみたいですね</li> <li><b>フィギュアスケート</b>大好きで桃の花みたい</li> <li><b>ビール</b>一番のお気に入りには白菜でした</li> <li><b>南関東</b>では早朝から元気な48歳になっちゃって大変な功績だと思います</li> <li><b>エッシャー</b>のだまし絵見ると出島みたい</li> <li><b>しょうが</b>がたっぷり載っけてみたい</li> </ul>
(3) マルコフ連鎖による生成文 Um	

図 6 発話生成文の例 (赤字はキーワード)

図 6 の生成例から分かるように、マルコフ連鎖では必ず生成文の文頭がキーワードになるが、提案方式ではそういうことがない利点がある。提案方式の不適当な例文を見ると、もともと選択されたテンプレートが不適当なもの他、「誰も見たことのない**カメラ**が撮りたい」などキーワード

と代表語の組み合わせは適当であってもテンプレートの選択が不適当なもの、「味わいたいです**フレンチ**草」のようなテンプレート文作成時に不要な語が残ってしまったもの、「**坂上**お気に入りの具は**ケージ**のミソ炒め」のように固有名詞が残ってしまっているものなどがある。テンプレート文の生成方法、選択方法に課題が多い。また、キーワードで置換した後の文の整形処理が必要である。

## 5. おわりに

テレビ視聴中の字幕文からキーワードを抽出し、感情表現を含む発話文生成手法について述べた。従来の対話システムでは、人からの入力文に応答してシステムが発話する文を生成するが、提案手法では、視聴中のテレビ番組の情報から発話文を生成できるため、人に対してシステム側から自発的に発話するシステムに利用できる特徴がある。また、字幕文をテンプレート文に用いることで、SNSのコメントを利用したときのように誹謗中傷するような文に成りにくい利点がある。

提案手法は、例えば「**ショウガ**にしちゃ**小さい**な」「**定食屋**で暮らしたい」などちょっと意外な組み合わせにより、主観的なちょっと面白い文が生成されることがある。文献[51]では用語の機械的な組み合わせから小説(ショートショート)のアイデアを見つけ出す方法が書かれているが、本手法も同様の効果があるのかもしれない。しかし、現状では「ヤマネコ釣りたいです」など意味不明な文が生成されることも多い。今後は生成文の妥当性を判定する手段、テンプレート文の生成・選択手法の改善、一貫性や番組内容に合った発話文の生成などが必要である。

## 参考文献

- [1] 大森慈子, 千秋紀子. 他者の存在が映像に対する面白さと笑い表情の表出に与える影響. 仁愛大学研究紀要. 人間学部篇 10, 2011, p.25-31.
- [2] “平成 27 年国勢調査”. 総務省統計局.
- [3] 村崎康博, 金子豊, 星祐太, 上原道宏. コミュニケーションロボットの社会的受容. 情処技報, 2018, vol. 2018-EIP-80, no. 2.
- [4] 齋藤健作. 高齢者のテレビ視聴(上)～視聴が増大する人・しない人～. 放送研究と調査, 2008, Sep., p. 2-17.
- [5] “平成 26 年高齢者の日常生活に関する意識調査”. 内閣府.
- [6] “平成 29 年国民生活に関する世論調査”. 内閣府.
- [7] “平成 23 年高齢者の経済生活に関する意識調査”. 内閣府.
- [8] 平野真理, 小倉加奈子, 坂本大介, 岩野裕利, 山下靖典, 土田剛生, 下山晴彦. セルフモニタリングデバイスとしてのお掃除ロボットの社会的受容. 信学技報, 2016, CNR2016-9, p.11-16.
- [9] 西脇裕作, 吉見健太, 岡田美智男. 言葉足らずの発話はなぜ人の心を引きつけるのか?. HAI シンポジウム, 2016, D-1, p.238-243.
- [10] 菅原董, 加納政芳. 人とロボットをつなぐために幼発話の活用可能か. 人工知能学会全国大会, 2016, 4C1-4.
- [11] 村上拓也, 李晃伸, 西川由理, 小島良宏. 音声対話システムにおけるシステムからの話しかけと他者性認知の関連性の調査. HAI シンポジウム, 2015, G-15, p.238-243.
- [12] 横山祥恵, 山本大介, 小林優佳, 土井美和子. 高齢者向け対話

- インタフェース。一雑談継続を目的とした。話題提示・傾聴の切替式対話法。情処技報, 2010, vol. 2010-SLP-80, no. 4.
- [13] 内田貴久, 港隆史, 石黒浩. 対話意欲を喚起する価値観肯定・否定割合に基づく自律対話ロボットの対話戦略. 人工知能学会全国大会 30, 2016, 115-1.
- [14] 金子豊, 星祐太, 上原道宏. 人と一緒にテレビを視聴するロボットの機能検討と試作. 第35回日本ロボット学会学術講演, 2017. RSJ2017AC112-04.
- [15] 狩野芳伸. コンピュータに話を通じるか 対話システムの現在. 情報管理, 2017, vol. 59, no. 10, p.658-665.
- [16] 呉先超, 伊藤和重, 飯田勝也, 坪井一菜, クライアン桃. りんな: 女子高生人工知能. 言語処理学会第22回年次大会, 2016, p.306-309.
- [17] 黒橋禎夫, 柴田知秀. 自然言語処理概論. サイエンス社, 2016, 2.4 マルコフモデルと言語モデル, p. 23-25.
- [18] 持橋大地, 隅田英一郎. Pitman-Yor 過程に基づく可変長 n-gram 言語モデル. 情処技報, 2007, vol. 2007-NL-178, no. 10, p. 63-70.
- [19] Graham NEUBIG, 三村正人, 森信介, 河原達也. ベイズ推論を用いた連続音声からの言語モデル学習. 情処技報, 2010, vol. 2010-SLP-82, no. 16.
- [20] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, Christian Jauvin. A Neural Probabilistic Language Model. Journal of Machine Learning Research 3, 2003, p. 1137-1155.
- [21] Tomas Mikolov, Martin Karafiát, Lukás Burget, Jan Cernocký, Sanjeev Khudanpur. Recurrent neural network based language model. INTERSPEECH 2010, p.1045-1048.
- [22] 久保田豊久, 若林啓. トピックモデルによる話題知識を考慮したテンプレート穴埋め型発話生成. 情処技報, 2017, vol. 2017-DBS-164 no. 10.
- [23] Hiroaki Sugiyama, Toyomi Meguro, Ryuichiro Higashinaka, Yasuhiro Minami. Open-domain Utterance Generation for Conversational Dialogue Systems using Web-scale Dependency Structures. SIGDIAL 2013, p. 334-338.
- [24] 稲葉通将, 神園彩香, 高橋健一. Twitter を用いた非タスク指向型対話システムのための発話候補文獲得. 人工知能学会論文誌, 2014, vol. 29, no. 1, p.21-31.
- [25] 田原俊一, 池田和史, 松本一則, 帆足敬一郎. ニュース関連ツイートの感情極性に基づく共感対話システムの提案と評価. DEIM Forum, 2018, D6-4
- [26] 柴田雅博, 富浦洋一, 西口友美. 雑談自由対話を実現するための WWW 上の文書からの妥当な候補文選択手法. 人工知能学会論文誌, 2009, vol. 24, no. 6, p.507-519.
- [27] Joseph Weizenbaum. ELIZA—A Computer Program For the Study of Natural Language Communication Between Man And Machine. Communications of the ACM, 1966, vol. 9, issue 1, p. 36-45.
- [28] Ilya Sutskever, Oriol Vinyals, Quoc V. Le. Sequence to Sequence Learning with Neural Networks. arXiv:1409.3215, 2014.
- [29] 赤間怜奈, 稲田和明, 小林颯介, 佐藤祥多, 乾健太郎. 転移学習を用いた対話応答のスタイル制御. 言語処理学会第23回年次大会, 2017, p.338-341.
- [30] 相楽翔太, 呉毘盧, 村上洋平, 川上皓平, 中辻真. ユーザに気づきを与える対話応答生成の試み. 人工知能学会全国大会, 2017, 2N1-4.
- [31] Oriol Vinyals, Alexander Toshev, Samy Bengio, Dumitru Erhan. Show and Tell: A Neural Image Caption Generator. CVPR2015, p.3156-3164.
- [32] “Yahoo! JAPAN, 「リアルタイム検索」に Twitter 上のテレビの情報を集約し, ダブルスクリーンで視聴を楽しむ新機能を公開”. [https://blogs.yahoo.co.jp/yj\\_pr\\_blog/24480583.html](https://blogs.yahoo.co.jp/yj_pr_blog/24480583.html)
- [33] “ニコニコ実況”. <http://jk.nicovideo.jp/>
- [34] 向井田一平, 西村祥吾, 川波弘道, 神原誠之, 萩田紀博. テレビ視聴型雑談ロボットのためのソーシャルメディア実況コメントの評価. 信学技報, 2017, CNR2017-9, p.1-6.
- [35] 武田悠, 熊本忠彦. 雑談対話を利用したニュース記事推薦システム. DEIM Forum 2018, P6-6.
- [36] 熊本忠彦, 河合由紀子, 田中克己. 新聞記事を印象とするテキスト印象マイニング手法の設計と評価. 電子情報通信学会論文誌, 2011, Vol. 94-D, No.3, p. 540-548.
- [37] 大田浩志, 山本和英. 感情生起表現を基にしたレビュー文生成. 言語処理学会第14回年次大会, 2008, p. 596-599.
- [38] 風間勇志, 植野義孝, 渡部広一, 河岡司. 常識的感情判断システムと主体語処理. 情報科学技術フォーラム, 2002, E-28, p. 137-138.
- [39] 町田啓悟, 奥村紀之. 感情判断に基づく物語文章からの感想文自動生成手法. 情報処理学会第74回全国大会, 2012, 3R-1, p.259-260.
- [40] 緒方健人, 佐藤理史, 駒谷和範. 模倣と置換に基づく超短編小説の自動生成. 人工知能学会全国大会, 2014, 1C3-OS-14b-2.
- [41] 夏目和子, 刀山将大, 佐藤理史. 発話文自動生成のための日本語表現文型辞書の作成, 言語資源活用ワークショップ, 2016, p. 126-139.
- [42] 益岡隆志, 田窪行則. 基礎日本語文法—改訂版—. くろしお出版, 1992.
- [43] “word2vec”. <https://code.google.com/archive/p/word2vec/>.
- [44] 秋山智俊. 恋するプログラム—Ruby でつくる人工無脳. 毎日コミュニケーションズ, 2005.
- [45] 佐藤敏紀, 橋本泰一, 奥村学. 単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討. 言語処理学会第23回年次大会, 2017, p. 875-878.
- [46] “MeCab”. <http://taku910.github.io/mecab/>.
- [47] “平成28年度の字幕放送等の実績”. 総務省.
- [48] 金子豊, 竹内真也, 黄珉錫, 和泉吉則. 放送時刻でファイルにアクセス可能な分散ファイルシステム. 電子情報通信学会論文誌, Vol. J97-B, No. 11, 2014.
- [49] “Wikipedia:データベースダウンロード”. <https://dumps.wikimedia.org/jawiki/>.
- [50] “はてなキーワード—一覧ファイル”. <http://developer.hatena.ne.jp/ja/documents/keyword/misc/catalog>.
- [51] 江坂遊. 小さな物語のつくり方. 樹立社, 2011.

## 付録

### 付録A. 代表語とテンプレート数

実験に使用した字幕文の中の代表語を含む字幕数と, 抽出されたテンプレート数を示す. 下記表は登録されたテンプレート数の多い順に並べている.

代表語	同義語	テンプレート数(Nt)	字幕数(Nc)	Nt/Nc
みたい		22,026	200,397	0.11
行きたい	いきたい, 行ってみたい, 伺いたい, 出かけた, 出かけてみたい	1,831	25,470	0.07
食べたい	食いたい, 食べちゃいたい, 食べてみたい, たべたい, 試食したい	1,451	17,605	0.08
買いたい	購入したい, 購入してみたい, 手に入れた, 手に入れてみたい, 注文したい, 買ってみたい, 頂きたい, 欲しい	1,424	24,144	0.06
なりたい		942	10,000	0.09
見たい	見せていただきたい, 見学したい, 見学をしたい, 拝見したい, 観たい, 観てみたい	894	15,614	0.06
いったい		782	27,638	0.03
聞きたい	聞いてみたい, お聞きしたい, 聴いてみたい, 聴きたい	360	11,884	0.03
話したい	おしゃべりしたい, しゃべってみたい, しゃべりたい, 言いたい, 言ってみよう, 話したい, 話してみたい, 話しをしたい	328	10,922	0.03
会いたい	お会いしたい	325	5,661	0.06
作りたい	創ってみたい, 創りたい, 造ってみたい, 造りたい, 作ってみたい	276	2,620	0.11
飲みたい	一杯やりたい, 飲んでみたい	258	2,097	0.12
知りたい		222	7,101	0.03
ありがたい		153	10,297	0.01

取りたい	取ってみたい	152	1,321	0.12
伝えたい		135	1,768	0.08
出たい	出てみたい, 出演してみたい	134	1,776	0.08
乗りたい	乗ってみたい	133	1,389	0.10
入りたい	入ってみたい	130	1,980	0.07
あげたい	買ってあげたい	123	2,457	0.05
帰りたい		120	2,374	0.05
期待したい		107	1,198	0.09
守りたい		107	1,144	0.09
使いたい	使ってみたい, 試したい, 試してみたい	98	1,682	0.06
勝ちたい		73	1,380	0.05
釣りたい	釣り上げたい, 釣ってみたい	52	410	0.13
撮りたい	写真を撮ってみたい, 写真を撮りたい, 写真撮りたい	47	810	0.06
歌いたい	歌ってみたい	42	777	0.05
応援したい		42	469	0.09
一緒にいたい	一緒にいたい, そばにいたい	41	908	0.05
働きたい	働いてみたい	39	543	0.07
着たい	身につけたい, 身につけてみたい, 着てみたい	36	608	0.06
注目したい		36	377	0.10
結婚したい	結婚してみたい	35	929	0.04
遊びたい	遊んでみたい	34	807	0.04
助げたい		31	758	0.04
打ちたい	打ってみたい	29	328	0.09
描きたい	描いてもらいたい	28	428	0.07
味わいたい	味わってみたい	28	366	0.08
飼いたい	飼ってみたい	27	235	0.11
育てたい	育ててみたい	25	394	0.06
持てたい	持ってみたい	22	321	0.07
歩きたい	歩いてみたい	22	368	0.06
触りたい	触ってみたい, 触れたい, 触れてみたい	21	520	0.04
走りたい	走ってみたい	19	326	0.06
信じたい		19	521	0.04
調べたい	研究したい, 研究してみたい	18	362	0.05
参加したい	参加してみたい	16	227	0.07
習いたい	習ってみたい	15	329	0.05
暮らしたい		15	426	0.04
踊りたい	踊ってみたい	14	235	0.06
書きたい	書いてみたい	13	441	0.03
読みたい	読んでみたい	12	509	0.02
殺したい		10	382	0.03
当てたい	当ててみたい	9	309	0.03
デートしたい	デートしてみたい	8	109	0.07
近づきたい	近づいてみたい	8	135	0.06
仲よくしたい	仲よくしてみたい, 仲良くしたい, 仲良くしてみたい	7	71	0.10
ご覧いただきたい		7	860	0.01
キスしたい	キスしてみたい	5	115	0.04
散歩したい	散歩してみたい, 散歩をしたい	5	60	0.08
訪ねたい	訪ねてみたい	4	74	0.05
答えたい		4	99	0.04
チェックしたい	チェックしてみたい	3	62	0.05
愛されたい	愛されてみたい	2	88	0.02
幸せになりたい		2	119	0.02
合計数		33,436	406,134	0.08

代表語(形容詞)	テンプレート数 (Nt)	字幕数 (Nc)	Nt/Nc	
好き	12,134	161,082	0.08	
大好き	5,311	40,095	0.13	
強い	4,421	47,279	0.09	
高い	3,785	43,520	0.09	
大変	3,126	66,182	0.05	
元気	2,141	57,790	0.04	
大きい	2,127	38,856	0.05	
上手	2,094	24,907	0.08	
美しい	1,587	15,002	0.11	
怖い	1,465	51,070	0.03	
速い	1,218	16,069	0.08	
面白い	1,112	40,401	0.03	
苦手	1,100	10,531	0.10	
小さい	1,082	15,678	0.07	
不思議	878	21,811	0.04	
気持ち	526	85,005	0.01	
夢中	475	4,309	0.11	
便利	407	7,932	0.05	
お気に入り	342	3,577	0.10	
懂れる	130	935	0.14	
美味	77	18,392	0.00	
はまる	52	997	0.05	
綺麗	19	291	0.07	
合計数		45,609	771,711	0.06

付録 B. 代表語に対する共起性が大きいキーワード

代表語に対するコサイン類似度の大きい上位 10 個のキーワードを以下に示す。代表語に対して筆者が主観的に文にしやすいと判断したキーワードと個数を赤字で示した。

代表語 (動詞句)	上位キーワード (数値はコサイン類似度)
みたい (10)	ドラクロア(0.529), サングリア(0.519), プレミアムモルツ(0.517), みにくいアヒルの子(0.516), 瞬間メタル(0.499), テーブルマジック(0.498), ノストラダムスの大予言(0.496), ボディランゲージ(0.495), ビール(0.493), 土用の丑(0.493)
訪ねたい (8)	小石川後樂園(0.494), 佐々木只三郎(0.475), 駿府城(0.462), 日本青年館(0.459), ホワイトホース(0.459), 大窪寺(0.458), 法輪寺(0.455), 若宮大路(0.448), 三笠山(0.447), 大國魂神社(0.446)
伝えたい (5)	ジロリアン(0.505), 夢のチカラ(0.491), 愛なんだ(0.486), カツゲン(0.482), 花は咲く(0.472), ジャスティン・ティンバレーク(0.470), 有間皇子(0.464), 住み開き(0.463), あなたにしかできないこと(0.462), 言葉にできない(0.462)
注目したい (10)	玉木幸則(0.451), フロントライン(0.446), 列車番号(0.433), ゴルゴンゾーラチーズ(0.430), トワイライトゾーン(0.424), ボトムアップ(0.421897), 温暖前線(0.417), 世界禁煙デー(0.417), 清水秀彦(0.417), 岸川聖也(0.416)
愛されたい (4)	さげまん(0.523), 大阪の女(0.520), 本歌取(0.512), シニカル(0.509), おっぱい星人(0.509), アーロンチェア(0.501), 現金主義(0.500), 乳糖不耐症(0.497), 納棺師(0.495), おめでた婚(0.494)
答えたい (8)	当て振り(0.474), 野沢直子(0.471), ずんのやす(0.470), 笑い飯 西田(0.465), E-Girls(0.460), 片平なぎさ(0.454), 美保純(0.454), マルモのおきて(0.451), 笑い飯(0.450), 藤井流星(0.449)
期待したい (9)	男子73キロ級(0.470), タキツバ(0.456), おさかな日村(0.452), ルンファア(0.448), トランボリン男子(0.446), 後藤輝基(0.443), 札幌オリピック(0.437), 岡村正広(0.433), 笑いが一番(0.433), 藤田征樹(0.433)
仲よくしたい (5)	恋の語と(0.523), 正直しんどう(0.508), 河内弁(0.498), 虎南有香(0.495), アスコマーチ(0.495), 幸せになろうよ(0.489), ドレディア(0.489), 金子信雄(0.484), 東京乾電池(0.483), アスペルガー(0.483)
散歩したい (7)	コニーアイランド(0.496), 小倉公園(0.483), 桜上水(0.447), パート・ランカスター(0.447), 高幡不動(0.437), 東京ダイナマイト(0.432), シーズ(0.430), 武蔵境(0.427), 常陸太田(0.426), 骨董屋(0.425)
食べたい (9)	スコッチエッグ(0.569), 故郷に帰りたい(0.528), てびち(0.519), 水茄子(0.513), たまごかけごはん(0.509), チゲ鍋(0.509), ホット鍋(0.506), ひなふら(0.501), ゴマペースト(0.495), はなの舞(0.494)
ご覧いただきたい (10)	狐の嫁入り(0.482), 木曜時代劇(0.467), 山城国一揆(0.448), 保谷市(0.442), メトロポリタン美術館(0.440), 東京ホルモ娘(0.439), 中国新聞(0.438), ドリー・パートン(0.438), 新ひだか町(0.436), 浜田省吾(0.435)
調べたい (10)	群馬大学(0.430), 国税庁(0.430), テトロドトケン(0.426), ベーバーカンパニー(0.423), バーチャル空間(0.422), 科学捜査研究所(0.418), 大宅 光子(0.415), スレンダーマン(0.414), 家計調査(0.409), 業務妨害(0.408)
知りたい (10)	ヴィーガン(0.465), 藤巻王(0.461), ハウツー(0.452), ヲイグル料理(0.450), グロブ座(0.442), ガスパッチョ(0.440), フィロソフィー(0.439), モバイルサイト(0.439), 組織犯罪対策部(0.437), 安定コソ薬剤(0.436)
会いたい (6)	シューブ(0.496), スロボキア人(0.481), 小宮浩信(0.472), カンサス(0.461), 松原千明(0.459), モンテレー(0.459), ノア(0.456), あの人に会いたい(0.441), 近衛十四郎(0.432), スウェイン(0.432)
チェックしたい (9)	イミダベプチド(0.454), 薬録案(0.433), ハードディスクレコーダー(0.432), ナスカン(0.428), アクシオンメラル(0.428), ダークチョコレート(0.422), 思考停止(0.420), エスリカ(0.413), アンチチームゼ(0.412), アンセリン(0.412)
結婚したい (4)	結婚(0.549), 結婚しない(0.500), シルバニアファミリー(0.463), かんちゃま(0.463), 結婚しようよ(0.449), 猫塾(0.448), ハイジャン(0.436), 種馬マン(0.435), てきちった婚(0.432), 運転免許試験場(0.427)
いったい (7)	ルベル(0.474), 離脱症状(0.455), ジェームズ・テイラー(0.452), 住民参加(0.450), プラシーボ(0.449), ヨノゾロ(0.446), ディミクスエンド(0.445), どんないこと(0.441), サモトラケのニゲ(0.439), ピンクパンサー(0.437)
殺したい (5)	処女懐胎(0.453), パルログ(0.447), スクィーズ(0.445), バックドラフト(0.442), ガノン(0.439), ムーランルージュ(0.433), 乙女ちゃん(0.432743), 佐藤和夫(0.431), 殺人鬼(0.418), テスラコイル(0.415)
歩きたい (6)	歩いていこう(0.467), クロックス(0.465), 青梅街道(0.439), ポジターノ(0.432), ソチミルコ(0.431), 駒沢通り(0.424), コアジサイ(0.417), 井ノ頭通り(0.415), オットマン(0.406), 深川不動(0.404)
話したい (6)	ドダイトス(0.496), 愛してた(0.475), 激おこぶんぶん丸(0.461), 知る権利(0.458), お客様は神様です(0.451), ゲームボーイ(0.446), ストリートギャング(0.444), 理想論(0.443), 渡辺ひとみ(0.441), スローター(0.439)
作りたい (9)	ビターチョコレート(0.448), 氷見うどん(0.439), セルブカパー(0.434), ローズウォーター(0.434), ゲイラカイト(0.431), ボルガ(0.431), パランスシート(0.427), ラタトゥイユ(0.426), ストロベリーパフェ(0.423), 鮭茶漬け(0.420)
勝ちたい (5)	ゴールパフォーマンス(0.457), 丹下左膳(0.442), ベストメンバー(0.438), テクニカルノックアウト(0.434), 完全燃焼(0.433), 中井貴裕(0.422), ちゃん掛け(0.421), ランサーズ(0.420), ライフセービング(0.418), はねトビ(0.418)
歌いたい (9)	Q U E E N(0.533), リミックス(0.509), ダークダックス(0.507), プリンセスプリンセス(0.503), B R E A K E R Z(0.501), 青江三奈(0.485), アニメソング(0.484), 手嶌葵(0.471), 渡り廊下走り隊7(0.469), ヘンゼルとグレーテル(0.469)
見たい (10)	オッサー(0.567), トゥモロランド(0.557), 社長シリーズ(0.467), テレビデオ(0.452), 池田成志(0.451), ラブドル(0.450), だいじょうぶだよ(0.446), ジェット・リー(0.445), エディ・マーフィー(0.440), キングシリロウモリ(0.436)
取りたい (1)	大塚高校(0.448), キム・ソンミン(0.439), 正木健人(0.431), 電気工事士(0.420), アマレス(0.417), バイエルンミュンヘン(0.409), モスクワオリンピック(0.405), 仙台育英高校(0.401), ゴトー(0.401), パファローズ(0.401)
キスしたい (3)	ひとりじゃない(0.493), ジョイスティック(0.451), キュアマミューズ(0.445), カギカコ(0.444), 時間よ止まれ(0.443), ゴニル(0.434), マイケル・ライ(0.433), 抱きしめたい(0.432), 伝えたいことがあるんだ(0.428), シュワルツェネッカー(0.427)
聞きたい (3)	ブルーノ・マース(0.487), スネオ(0.457), アマラオ(0.455), 後日談(0.453), やぶさかでない(0.452), かりゆし58(0.447), 尾田栄一郎(0.445), 林羅山(0.438), A K I K O(0.437), 岡崎公園(0.436)
ありがた	オキシドール(0.518), 温泉神社(0.484), コロンブスの卵(0.479), クリス習子

い (5)	(0.470),ハンカチノキ(0.456),水木プロダクション(0.451),すっぱん大学(0.450),ことてんでん(0.449),紋切り型(0.447),ひまわり(0.445)
働きたい (9)	マッキンゼー(0.461),託児所(0.455),個人事業主(0.442),コンサル(0.427),福利厚生(0.409),勤務機(0.407),土木作業員(0.407),介護ヘルパー(0.401),救世軍(0.399),造園業(0.397)
描きたい (8)	岸田劉生(0.512),クルールベ(0.469),トレーニングペーパー(0.466),少年マンガ(0.453),リアリズム(0.452),小川宇軒(0.447),ドローイング(0.443),無言館(0.442),鍾木清方(0.439),ルーベンス(0.437)
踊りたい (4)	フランチェン(0.487),かなめも(0.480),ボーイソプラノ(0.471),戦野波平(0.466),カマラン(0.454),あなたに夢中(0.443),バトミントン(0.425),成宮寛貴(0.423),大井競馬(0.417),ドゥルーズ(0.416)
守りたい (8)	聖家族(0.454),心の鐘(0.439),ダブラン(0.429),インセプション(0.429),社会秩序(0.422),神代凌牙(0.421),文明国(0.414),最高指揮官(0.411),イバライガー(0.409),ファイアボール(0.408)
なりたい (6)	あなたに会えてよかった(0.533),プリンシプル(0.532),逆ハーレム(0.491),ストックホルム症候群(0.488),セックスフレンド(0.483),誰も守ってくれない(0.477),マツキヨ(0.474),コアリズム(0.473),明日があるさ(0.472),ズブ友(0.471)
飼いたい (8)	サモエド(0.486),シーズー(0.473),セントバーナード(0.469),猫(0.456),ティーカップブードル(0.454),あなたに夢中(0.443),アラビアオックス(0.428),ねりきり(0.426),紀州犬(0.426),キンラン(0.419)
飲みたい (9)	サンクリア(0.519),プレミアムモルツ(0.517),ビール(0.493),ロイヤルミルクティー(0.483),ペダトナ(0.477),マキアート(0.475),コクリ(0.475),薬用酒(0.474),コービー(0.469),おとぞ(0.468)
書きたい (8)	喜多條忠(0.471),筆記体(0.465),私小説(0.458),暑中見舞い(0.439),岡本かの子(0.438),ミツバツジ(0.435),異体字(0.434),石川セリ(0.434),ヘブライ語(0.431),短編小説(0.428)
打ちたい (4)	イレギュラー(0.445),マルカーノ(0.427),サンドウェッジ(0.427),アイアン(0.422),ヌマワニ(0.417),バیتالエリア(0.413),岡島豪郎(0.412),リアップ(0.410),遺伝子治療(0.410),コ・ジョン(0.408)
信じたい (4)	スキュラ(0.440),エルガ(0.407),ゴジカ(0.403),理論武装(0.398),ピュウター(0.398),一夫一婦制(0.396),ソナタ(0.390),人相学(0.390),プライマー(0.389),たらいまわし(0.387)
助けたい (8)	アフェール猿人(0.462),ヒトヨタゲ(0.447),キニーネ(0.438),キミがいる(0.423),加藤誠(0.418),逆転移(0.416764),延命措置(0.416),山崎哲也(0.416),イングリッド・バーグマン(0.410),ボスドク(0.406)
走りたい (7)	軽車両(0.429),デューダ(0.411),ジャパンカップ(0.402),球技大会(0.399),吉野ヶ里歴史公園(0.398),有料道路(0.391),船の科学館(0.387),トムトム(0.385),ウサイン・ボルト(0.374),スポーツイベント(0.369)
乗りたい (9)	超小型車(0.491),テロリアン(0.469),トワイシクル(0.437),パレオエクスプレ(0.434),タワラ・オブ・テラー(0.432),鹿沼由理恵(0.432),スーパーひたち(0.431),しまかせ(0.421),都営浅草線(0.419),マスタング(0.418)
出たい (6)	大阪ドーム(0.451),イナバ物置(0.419),ラブライブ(0.418),カレカ(0.411),六大学野球(0.407),アッコにおまかせ(0.407),津川利愛美(0.403),ボラギノール(0.399),新国立劇場(0.394),冬のソナタ(0.394)
撮りたい (8)	ヒロヒト(0.490),ストップモーション(0.474),私の履歴書(0.460),お七夜(0.449),ダイヤモンド富士(0.442),長回し(0.438),デスクトップ(0.435),ブリクラ(0.435),魚眼レンズ(0.431),セルフタイマー(0.426)
参加したい (5)	三井不動産(0.433),餅つき(0.430),おじゃマップ(0.428),スペインオリピック(0.421),リエンソ(0.379),国家体制(0.377),ワダ芸(0.375),キューン(0.372),埼玉県警(0.369),不当解雇(0.368)
幸せになりたい (3)	あなたに会えてよかった(0.533),プリンシプル(0.532),逆ハーレム(0.491),ストックホルム症候群(0.488),セックスフレンド(0.483),誰も守ってくれない(0.477),マツキヨ(0.474),コアリズム(0.473),明日があるさ(0.472),ズブ友(0.471)
遊びたい (10)	ティンエージャー(0.446),大雑志(0.438),パーティーゲーム(0.424),砂遊び(0.413),ちゃんぽん(0.409),ファズ(0.407),ヨーギラス(0.405),デリバド(0.404),チェビ(0.403),ラブラドル(0.403)
買いたい (9)	リッキー・マートン(0.457),古細菌(0.443),菊田一夫(0.443),心拍計(0.424),きみさき(0.423),谷本知美(0.421),スターゲイザー(0.420),マリエル(0.416),浪漫飛行(0.413),ストレッチボール(0.412)
釣りたい (10)	メバル(0.513),クロソイ(0.507),マダイ(0.489),ヒラマサ(0.470),マアジ(0.462),エンゼルフィッシュ(0.453),アカマンボウ(0.452),クエ(0.450),イサキ(0.445),ヒラルク(0.445)
行きたい (4)	青葉台駅(0.446),デジタルサイネージ(0.440),浄妙寺(0.420),親権者(0.409),松重豊(0.407),新宿高校(0.406),ベネロビ(0.404),ワールド・ビジョン(0.396),せせんべい(0.390),一つの中国(0.390)
デートしたい (6)	パンティ(0.434),ミササギ(0.424),ボストンテリア(0.415),まいぶる(0.413),男闘呼組(0.401),毛麻伸(0.400),矢追純一(0.400),水嶋ヒロ(0.396),アナザーサイ(0.389),珠理奈(0.387)
読みたい (10)	田中ユギ(0.477),宮沢賢治(0.466),漫画(0.457),飯島奈美(0.453),つげ義春(0.449),久住昌之(0.447),エッセイ(0.429),入門書(0.429),夢占い(0.429),連載小説(0.429)
近づきたい (3)	トムラウシ(0.402),尊王攘夷(0.401),白瀬轟(0.386),ランゲージ(0.385),中華民族(0.380),国嫌(0.378),街道をゆく(0.375),肉食系(0.374),焼津漁港(0.373),伊東忠太(0.370)
帰りたい (0)	故郷に帰りたい(0.528),タイガーパーム(0.454),楽天カード(0.447),二番町(0.444),恋しくて(0.434),骨伝導(0.426),伊東屋(0.426),今はもうない(0.415),ケヤリムシ(0.410),ルサンチマン(0.407)
育てたい (10)	アスチルベ(0.434),無農薬(0.430),エビネ(0.430),ヒルガオ科(0.429298),ナガイモ(0.416),アブラナ科(0.411),カラコンコエ(0.409),シクラメン(0.399),宿根草(0.398),家庭菜園(0.398)
着たい (9)	パバシャツ(0.482),長襦袢(0.465),リパシーブル(0.460),十二単(0.456),ワンピース(0.444),カーゴパンツ(0.438),山崎真紀(0.437),黄八丈(0.435),ジャンパー(0.434),ベリル(0.428)
使いたい (7)	スマート미터(0.443),ポリ塩化ビニル(0.437),ケムリ(0.437),ラモーンズ(0.435),コレクターズ(0.429),フォークロア(0.426),尊敬語(0.410),ビッグライト(0.406),周辺機器(0.400),方眼紙(0.399)
暮らしたい (9)	ひとつ屋根の下(0.433),山添村(0.432),明野町(0.411),ナルニア(0.386),市町村合併(0.382),ルサント(0.377),ジョイマン高木(0.372),四合院(0.365),稲葉正成(0.363),ユウジ(0.362)
味わいたい (9)	弥生人(0.411),心象風景(0.409),台湾茶(0.408),モツ鍋(0.402),フグ刺し(0.396),鶏刺し(0.395),ネーブル(0.395),オランダ煮(0.384),西京味噌(0.383),ザボン(0.379)
入りたい (3)	島袋聖南(0.476),ランチラン(0.430),片倉館(0.406),Gたかし(0.392),塩原恒夫(0.389),佐藤夏希(0.388110),日帰り入浴施設(0.385),軽音楽(0.381),ニューエネレーション(0.378),クレーンジャック(0.378)
触りたい	ロミオ(0.400),アキバ(0.389),メンタルヘルス(0.386),シニオン(0.381),

(5)	マナティ(0.373),野性爆弾(0.371),団地妻(0.366),AR(0.362118),おそろぎ(0.358),幽霊部員(0.356)
当てたい (1)	ルウト(0.396),おやじの背中(0.394),三毛猫ホームズ(0.394),メグロコ(0.378),マジックアワー(0.377),ランニングアプローチ(0.376),ポニーキヤニオン(0.371),裏ビデオ(0.361),ふしぎ発見(0.358),設楽統(0.356)
応援したい (8)	夢を与える(0.453),歌の方(0.450),五目並べ(0.443),松村香織(0.430),草の根運動(0.422),頑張り日本(0.416),ブラックフライデー(0.398),ルミナス(0.380),別所キミエ(0.380),スカパラ(0.379)
習いたい (7)	古式泳法(0.429),民俗学(0.388),東ハト(0.388),教育番組(0.381),夜回り先生(0.377),ノーバンスヤブ(0.376),ふしぎ発見(0.358),設楽統(0.356),共通語(0.371),黒革の手帳(0.366)
持ちたい (3)	エコイスト(0.438),恋する日本語(0.379),ウォーズマン(0.367),一戸建(0.366),イギリス料理(0.356),ムーア人(0.355),コンセンサス(0.354),モルガン(0.350),中二階(0.350),アウトソーシング(0.348)
一緒にしたい (6)	ボール牧(0.496),j y A-M e(0.466),お兄ちゃんと一緒に(0.466),ベネロビ・クルス(0.440),ずっと好きだった(0.435),レスビーギ(0.433),いいなずけ(0.406),アメリカン(0.406),さっさと大丈夫(0.406),仮面夜(0.403)
あげたい (1)	リーマス(0.481),チェーントピア(0.437),ゆとりちゃん(0.422),富地真緒(0.411),F U J I W A R A 原西(0.409),平野友朗(0.406),ベネロビ(0.404),小矢部(0.404),専修念仏(0.401),聖一園師(0.398)

代表語 (形容詞)	上位キーワード (数値はコサイン類似度)
美しい (10)	ルネサンス様式(0.563),ルドベキア(0.560),ジャンボール城(0.559),マデイラ島(0.552),ジャカララン(0.544),ヤマツツジ(0.544),日本の名峰(0.543),白米の千枝田(0.535),ものあわれ(0.533),イオニア海(0.531)
不思議 (8)	摩訶不思議(0.547),不可思議(0.540),金木犀(0.517),ホーンテッドマンショ(0.517),高田謙長(0.514),不気味(0.513),アエニス(0.503),サンビラー(0.496),エキゾチック(0.495),コンセンサス(0.493)
速い (2)	ライヒ(0.536),金田和也(0.521),サイドウェイ(0.510),バイエル(0.503),りんたろう(0.500),与一(0.492),ハッセル(0.491),池田亮介(0.484),キャベツ太郎(0.473),反応速度(0.472)
怖い (7)	オオメジロザメ(0.524),エイトマン(0.522),ダグトリオ(0.522),アールン(0.518),困鬼(0.516),イレブンナイン(0.515),飛蚊症(0.511),シニカル(0.508),故郷に帰りたい(0.506),おっぱい星人(0.506)
美味 (6)	土佐日記(0.562),海の幸(0.547),関あじ(0.532),どじょう鍋(0.522),成川美術館(0.521),KG(0.517),大町温泉郷(0.509),おっぴらメ(0.496),かりんどう饅頭(0.496),大鰐温泉(0.493)
高い (4)	ハルラ(0.531),タンザナイト(0.528),ユーフォニアム(0.481),コウヤマキ(0.479),生殖細胞(0.478),由布院温泉(0.478),杜の都(0.477),キカイダー(0.474),ルメール(0.474),取崩期血圧(0.466)
大好き (10)	小橋建太(0.516),八日市場(0.499),松下宣夫(0.491),フランク永井(0.478),ブルースハープ(0.476),大地の子(0.473),松本大洋(0.470),アニメオタク(0.469),グーズベリー(0.466),リリース(0.462)
面白い (9)	すっぱん大学(0.566),スポーツ漫画(0.557),瞬間メタル(0.533),オルメカ(0.528),ハンカチノキ(0.510),高野寛(0.506),サティスティック(0.502),ゲイライカイト(0.501),有間島子(0.499),テレビオ(0.499)
強い (8)	ジフテリア(0.496),アスチルベ(0.492),キム・ジョン(0.470),空気流(0.463),アガパンサス(0.459),判官びいき(0.458),ナライ(0.453),下っ風(0.449),自明の理(0.448),モリス(0.448)
懂れる (9)	橋本祥(0.501),マダニ(0.483),エディ・マーフィ(0.480),バンクロク(0.470),恋の詩と(0.464),萩原明太郎(0.462),コンテンポラリー(0.458),江戸川コナン(0.455),レーシングドライバー(0.451),ペルリオーズ(0.447)
小さい (8)	ヒメリンゴ(0.579),かんころもち(0.533),レーザーディスク(0.523),カマツカ(0.495),岸壁の母(0.492),ミニチュアダックスフント(0.491),ハシブトガラス(0.491),パンパズ(0.491),ブラック・サブス(0.489),ハオコゼ(0.479)
お気に入り (10)	売れ筋(0.471),トム・ペティ(0.461),ジンジャーブレッド(0.439),伊藤みちよ(0.430),グリーンガラス(0.424),ケンズ(0.421),ベルサイユのばら(0.420),ゼンダマン(0.414),上野台地(0.413),クラシックギター(0.413)
好き (10)	小橋建太(0.516),テレビっ子(0.516),八日市場(0.499),松下宣夫(0.491),森大(0.490141),財津一郎(0.489),フランク永井(0.478),手裏剣(0.477782),ブルースハープ(0.476),文系系(0.474)
便利 (7)	保存食品(0.522),フリック入力(0.497),コレクターズ(0.488),保険者(0.475),プリントゴッコ(0.468),ピストルズ(0.466),三原圭(0.459),ワープロソフト(0.458),スマート미터(0.458),カチューム(0.452)
大きい (10)	バンド型(0.525),ヌカエビ(0.524),オニオニオニマキエ(0.502),オナガバズナギドリ(0.460),パラフェダ(0.455),オニゴモ(0.452),プライオサウルス(0.451),メガウルス(0.447),キチジ(0.446),キンギョハナダイ(0.443)
大変 (7)	児童福祉司(0.504),ハウンド・ドッグ(0.485),人民裁判(0.449),華まめ(0.439),ことでん(0.439),ミルクiewicz(0.437),タイプミス(0.437360),四苦八苦(0.437),不当判決(0.433),暁羅雄・光雄(0.432)
上手 (9)	琴吹(0.579),内無双(0.499),が一まるちよ(0.492),ママちゃん(0.480),北天佑(0.471),極大海(0.460),うちやちや(0.452),後ろもたれ(0.451),旭富士(0.448),ちゃん掛け(0.441)
元気 (5)	笑って許して(0.442),イラプー汁(0.429),色事師(0.414),アルケマ(0.412),佐武宇崎(0.411),ミッションスクール(0.408),コペン(0.408),カーミット(0.406),フレンドリー(0.404),ロケック(0.404)
気持 (6)	タネボー(0.513),あー夏休み(0.485),小室徳幸(0.460),スネオヘアー(0.454),トウゴクミツバツジ(0.452),ゲロー・クーパー(0.448),空がきれい(0.444),湯あたり(0.440),お客様は神様です(0.437),理論武装(0.430)
夢中 (5)	野いちご(0.452),大捜査線(0.440),天文学的数字(0.439),グラウンドゴルフ(0.436),ヒサカキ(0.436),あなたがた(0.434),アナホリクロウ(0.433),ブルーバック(0.433),修二と夢(0.424),アズマヒキガレ(0.421)
はまる (9)	ボーイズラブ(0.430),コロポックル(0.405),サイバー空間(0.393),固定価格買取制度(0.385),ニューミュージック(0.380),結婚できない男(0.379),催眠術(0.375),地下鉄博物館(0.372),サブリーミナル(0.372),百田夏菜子(0.371)
苦手 (10)	あがり症(0.477),論理学(0.451),シーライオン(0.447),強迫神経症(0.432),マルカーノ(0.431),運動音痴(0.425),高所恐怖症(0.409),そっち系(0.404),ローボールヒッター(0.400),マッシュ・ペリー(0.399)
綺麗 (6)	キレイキレイ(0.459),アズマキナグ(0.395),おに(0.394),植木等(0.387),オールマイデー(0.365),フェミニン(0.360),南部美人(0.354),ロマンチスト(0.353),しづか(0.350),防空法(0.346)