

会話によるニュース記事伝達のための 口語化における述語の書き換え

高津 弘明¹ 福岡 維新¹ 藤江 真也^{1,2} 林 良彦¹ 小林 哲則¹

早稲田大学¹ 千葉工業大学²

{takatsu,fukuoka}@pcl.cs.waseda.ac.jp, shinya.fujie@p.chibakoudai.jp,
yshk.hayashi@aoni.waseda.jp, koba@waseda.jp

1 はじめに

書き言葉で記述された情報を話し言葉に変換して伝える際に必要となる述語の書き換えについて検討した。

我々は、ニュース記事に代表されるまとまった量の情報を利用者の理解や興味に応じて会話でインタラクティブに伝える音声対話システムの研究・開発を行っている [1]。聴覚メディアで未加工のニュース記事を伝えるとき、聞き手はたとえ興味がない内容であっても最後まで黙って聴くしかなく、フラストレーションの大きいものとなる。要約した内容を会話的に質疑を交えながらインクリメンタルに伝えることでこの不都合は解消できると考えるが、そこでは書き言葉を話し言葉に変換する必要が生じる。

書き言葉と話し言葉の最も基本的な違いとして、話し言葉の同時性がある [3]。話し言葉によるコミュニケーションでは、話し手が情報を送信すると同時に聞き手が受信することになる。話す内容や話すペースは話し手が制御することになるため、聞き手の負担は、話し手の裁量に依存する。そのため、話し言葉では書き言葉に比べ、聞き手にとって負担の小さい表現が使われる傾向にある。

本稿では、ニュース記事の内容を会話で伝達する際の話し言葉への書き換えについて、上述の話し言葉の特性を考慮しながら、特に、述語表現に焦点を絞って検討した。しかしながら、述語だけを取り上げてみても、鍛治ら [8] が指摘しているように話し言葉と書き言葉で共通して使われる表現も存在し、どの表現をどのように書き換えればよいかは明らかでない。そこで、人がニュースの内容を会話で伝える際、どのような言い換えを行うか調査するために口語化コーパスを作成した。そして、このコーパスを使用して、機械学習の手法に基づく書き換え実験を行った。

以下、2章で関連研究を概観した後、3章で設計したコーパスを紹介し、4章で各種検討した書き換え方法について説明する。

2 関連研究

書き言葉を話し言葉に書き換える研究は、主に音声合成の分野で行われてきた。音声合成を用いたアプリケーション開発では次の二つが問題となる。一つは、韻律やイントネーションが不自然であるという音響的な問題であり、もう一つは、音声化されるテキストの言語

的な問題である。書き言葉から話し言葉への書き換えに関する研究は後者の問題に着目したものである。音声合成分野で話し言葉が扱われる場合、ほとんどの場合、聞き手を特定しない不特定多数の人に (一方的に) 情報を伝えることを想定しているため、敬体 (ですます調) への書き換えが行われる。林らは、ニュース音声の自動配信アプリケーションを想定し、新聞記事を対象に文体変換規則による変換と体言止めの補完による話し言葉への変換を行った (e.g. 論争を展開 ⇒ 論争を展開しました) [4-6]。大泉らは、格解析結果を用いた名詞化された用言を含む表現の変換 (機種依存文字 ⇒ 機種に依存する文字) と、国語辞典を利用した漢語調の平易表現への変換 (e.g. 決定する ⇒ はっきり決める)、普通体及び文語調の口語調への変換 (e.g. 仕方ない ⇒ 仕方ありません) を行った [7]。鍛治らは、用言を対象として、国語事典から用言の言い換えペアを学習し、WWW から自動収集した書き言葉コーパス (新聞記事など) と話し言葉コーパス (チャットや掲示板など) を用いて、書き言葉語彙から話し言葉語彙への言い換え (e.g. 異様だ ⇒ 違う) を行った [8]。

外国人にもわかりやすい日本語「やさしい日本語」を作る研究が行われている [13]。「やさしい日本語」は、「難しい言葉を避け、簡単な語を使う」「二重否定を避ける」、「文末表現はなるべく統一する」などの規則に従い作成される。やさしい日本語を作るという点において話し言葉への書き換えと近いところがあるが、「やさしい日本語」は外国人にもやさしい読み物の作成を志向した研究であり、利用者とのカジュアルな会話 (多少砕けた言い回しも許容される) の実現を目指す本研究とは異なる。

対話システム研究における発話生成は、事前に人手で用意した発話文を使用したり、Twitter 上のツイートをそのまま利用したりすることが多い [9-11]。ヒューリスティックな変形ルールでツイートの語尾の口調を修正するなどの書き換えは行われている [12] が、文書のような書き言葉で書かれたまとまりある情報を会話で話すための表現に書き換える研究はない。

3 口語化コーパス

まず、次の二つの観点でニュース記事から伝えるべき内容の選択を行った。一つは、ニュースの見出し的な内容またはニュース記事の核となる情報を 1~4 文選択したものである (以降、要約 A)。もう一つは、ユー

表 1: 要約・口語化の例

原文	テニスのメキシコ・オープンが28日、メキシコのアカプルコで行われ、男子シングルス決勝で第1シードの錦織圭は第2シードのダビド・フェレールに3-6、5-7のストレートで敗れ、ツアー通算9勝目はならなかった。
要約	テニスのメキシコ・オープンが、アカプルコで行われ、男子シングルス決勝で錦織圭はダビド・フェレールに敗れた。
口語	テニスのさ、メキシコ・オープンがアカプルコで行われたよ。錦織圭がね、男子シングルス決勝でダビド・フェレールに負けたんだって。

ザーの関心を引くような内容または記事への理解を促進させる内容を要約 A に加えて 1~3 文選択したものである (以降、要約 B)。文選択を行った後、作業者は選択した文から不要な情報を文節単位で取り除く。文節の区切りは KNP で与えた。次に、作業者は選んだ情報を会話表現に書き換える。口語化作業では、要約 A, B で選ばれた情報を増やしたり損ねたりしないように書き換える。また、情報の伝達相手としては友達のようなキャラクターを想定し、話し手は中性的なキャラクターとする (「～らしいわよ」(女性的)「～だぜ」(男性的)のような、特定の性別・キャラクターを想起させるような表現は使わない)。口語化の操作として、述語や助詞の書き換え、指示詞や接続詞の挿入、助詞や句点の削除、文節の順序交換、連体修飾節の述語化、述語の含意表現への書き換えなどを許容した。以降、要約 A の口語化を口語化 A、要約 B の口語化を口語化 B と呼ぶ。100 トピックについて 3 人の作業者が要約と口語化を行った。ニュース記事は毎日新聞、産経新聞、読売新聞、NHK、nifty から収集したものを利用した。トピックは同じ話題について記述した 2 つ以上の記事から構成され、各記事は 5 文以上のものを選んだ。以降の実験では作業員 1 の口語化 B のデータセットを使用する。

実際に作成された要約と口語化の例を表 1 に示す。表の「原文」が書き言葉で書かれた元々の文である。会話で話す際、この内容をそのまま読み上げる人はいないだろう。相手が友達なら、表の「口語」のような話し言葉に直してから話すと考えられる。この例では、終助詞「さ」や「ね」が文節の切れ目に挿入されていたり (間投詞的用法)、堤題助詞「は」が主格「が」に書き換えられていたり、文節の位置が入れ替えられていたり (「男子シングルス決勝で」と「錦織圭は」) するが、本稿では、述語の書き換え (「行われ、」⇒「行われたよ。」、「敗れた。」⇒「負けたんだって。」) に焦点を当てる。

4 述語の書き換え

述語の書き換えは、次の 5 つの処理で行う。

- (1) 書き換えを行う述語を特定する。述語の中には書き言葉でも話し言葉でも利用される表現が存在するため、コーパスベースで書き換え対象となる述語を特定する。
- (2) 書き換え前後で、時制やモダリティは統一する必要

表 2: 述語の書き換え判定の実験結果

	精度	再現率	F 値
B	88.3	82.3	85.1
I	91.6	85.4	88.4
O	97.1	98.7	97.9
平均	92.3	88.8	90.5

があるため、述語の時制とモダリティを解析する。

(3) 文中の連用形述語や順接・逆接述語を終止形に直すかどうかを判定する。話し言葉では一回の発話で伝えられる量が限られるため、長い文では連用形の述語などを終止形に直し、発話を区切る必要がある。

(4) 述語を含意表現に書き換え、堅苦しい表現を緩和する。主に動詞を対象とする。例えば、「死去する」や「敗北する」などの書き言葉に見られる堅苦しい表現を「亡くなる」や「負ける」のように会話でもよく使われる柔らかい表現へと書き換える。

(5) 末尾表現を伝聞口調や断定口調に書き換える。書き言葉では句点を伴い終止形で文を終えることが多いが、話し言葉では発話末尾の口調を変えることで聞き手に対して情報伝達の意図を表現することが多い。

4.1 述語の書き換え判定

書き言葉と話し言葉では両方で共通して使われる表現もあるため、コーパスベースで書き換え対象となる述語の特定を行う。述語は文末だけでなく、文中に連用形として出現したり、連体修飾として出現したりする。これらの述語について、書き換えるべき述語を特定するのがここでの目的である。

識別モデルとして CRF を使用し、10 分割交差検定で評価した。ラベルとして IOB タグを定め、素性には JUMAN の形態素情報 { 基本形, 品詞大分類, 品詞細分類, 活用型, 活用形 } を用いた。実験結果を表 2 に示す。F 値はおよそ 90% であった。

4.2 述語の時制・モダリティ解析

書き換え前後で述語の時制やモダリティを合わせる必要がある。現在、時制は表層表現や係り受け関係を利用して推定している。義務や推量などのモダリティ、受動態や使役態などのヴォイスは、形態素列の情報を用いて判定している。

4.3 終止形化判定

会話では一回の発話で伝えられる量が限られる。そのため、長い文は連用形の述語などを終止形に直し、発話を区切る必要がある。そこで、ここでは、文中の連用形述語や順接・逆接述語を終止形に直して、文を区切るかどうかの判定を行う。この判定結果は、「アカプルコで行われ、」を「アカプルコで行われたよ。」のように書き換えるかどうかにつながる。線形カーネル SVM を使用し、5 分割交差検定で実験を行った。素性として、対象文節の各形態素の形態素情報 { 表層形, 基本形, 品

表 3: 終止形にするかどうかの実験結果

	精度	再現率	F 値
終止形にする	51.7	81.7	63.3
終止形にしない	98.1	88.6	93.1
平均	74.9	85.1	79.7

詞大分類, 品詞細分類, 活用形, 活用型 } と文の文字数・文節数, 対象文節前後の文字数・文節数を用いた。

実験結果を表 3 に示す。終止形にすると判定する精度が終止形にしないと判定する精度よりも低かったが、この結果の差は、データの偏りが原因だと考えられる。作業 1 の口語化 B のデータセットに含まれる連用形などの文中述語は 200 個あるが、この内、終止形に直されたものの数はわずか 14 個だった。

4.4 含意表現への書き換え

会話では使わないような堅苦しい表現を柔らかい表現に書き換える。例えば、「死去する → 亡くなる」や「敗北する → 負ける」のような言い換えが挙げられる。特に、述語の内、動詞を対象とする。

知識源として ALAGIN の『動詞含意関係データベース』を使用した。ここで、動詞 1 が動詞 2 を含意する(動詞 1 → 動詞 2)とは、動詞 1 の表す事態が成立するならば、同時かそれ以前に、動詞 2 の表す事態も成立しているということの意味する。このデータベースでは、正例群は「含意が成り立つ類義/上位下位関係」(e.g. 挑戦する → チャレンジする)「文字列上包含関係にあり、含意が成り立つ類義/上位下位関係」(e.g. あざ笑う → 笑う)「前提関係」(e.g. 酔っぱらう → 飲む)「作用反作用関係」(e.g. 借りる → 貸す)から構成される。この内、「含意が成り立つ類義/上位下位関係」と「文字列上包含関係にあり、含意が成り立つ類義/上位下位関係」のデータを利用した(計 49401 ペア)。

実験では、作業 1 の口語化 B のデータセットを使用し、頻度や分散表現の類似度、RandomForest の含意書き換え確率によりランキングしたときの一位正解率で評価した(表 4)。なお、平均書き換え候補数は 4.18 であった。Freq(wiki) は Wikipedia 上での頻度、Freq(tweet) は Tweet コーパス(Streaming API で約一年間収集したツイート)上での頻度、Sim(sg(wiki)) と Sim(sg(tweet)) は Wikipedia または Tweet コーパス上で学習した分散表現(skip-gram, 1000 次元)の類似度、RF{sg(wiki)} と RF{sg(tweet)} は、動詞 1,2 の分散表現(sg(wiki) または sg(tweet))を結合したベクトルで RandomForest を学習し、10 分割交差検定で評価した結果である。また、RF{(sg(wiki), Freq(tweet))} は、RF{sg(wiki)}の素性に Freq(tweet)を加えたときの結果である。結果は、RF{(sg(wiki), Freq(tweet))}が最も高い正解率を示した。他に注目すべき結果として、Tweet コーパス上での頻度のみでランキングしたときの正解率(Freq(tweet))も比較的高いことがある。これは、Twitter への書き込みに、会話に近い表現が多く使われているためだと考えられる。

表 4: 含意動詞への書き換え正解率

手法	一位正解率
Freq(wiki)	47.4
Freq(tweet)	73.7
Sim(sg(wiki))	40.8
Sim(sg(tweet))	57.9
RF{sg(wiki)}	75.0
RF{sg(tweet)}	68.4
RF{(sg(wiki), Freq(tweet))}	76.3

表 5: 「情報のなわ張り理論」(神尾, '90)における文末形式の分類

		話し手のなわ張り	
		内	外
聞き手のなわ張り	外	直接形	間接形
	内	直接ね形	間接ね形

このデータベースで全ての動詞の言い換えをカバーできるわけではないが、分散表現を用いた動詞含意知識の獲得 [17] や推移律による自動拡張 [18] などの手法により、データベースを拡張することが可能である。

4.5 末尾表現の書き換え

神尾は話し手と聞き手がある文の情報をどの程度知っているかによって日本語の文末形式が変わるとする「情報のなわ張り理論」を提唱した [19, 20]。1990 年のモデルでは、聞き手にとってその情報がなわ張りの内か外か(既知か未知か)、話し手にとってその情報がなわ張りの内か外かで異なる文末形式が顕在化するとした¹ [19](表 5)。直接形は断定・言い切り(e.g. 「昨日は動物園に行ってきた」)、間接形は伝聞(e.g. 「バリの冬は寒いらしい」)、直接ね形は直接形に終助詞「ね」を伴ったもの(e.g. 「いい天気ですねえ」)、間接ね形は間接形に終助詞「ね」を伴ったもの(e.g. 「君は退屈そうだね」)である。

システムがユーザーにニュースを伝える場合について考える。ニュースの内容はユーザーにとって概ね未知の情報であるため、なわ張りの外側の情報として受け取られると考えられる。一方、システムは、その内容を知っているため、なわ張りの内側として伝えることができるが、その内容の真偽を確認したわけではないので、なわ張りの外側としても伝えることができる。したがって、ニュースの内容をユーザーに伝える際には、断定口調または伝聞口調がよいと考えられる。

以上のことを踏まえ、ここでは、述語の末尾表現を断定口調または伝聞口調に書き換える。まず、書き換え規則として JUMAN の活用形と末尾表現のパターンをコーパスから取得する。例えば、「売れ始めたんだっ

¹神尾は、なわ張りを [0,1] の連続空間で表現し、発話形が 6 つの状況で変わるとするモデルを 2002 年に提案している [20]

表 6: 述語末尾表現の正解率

	(1)			(1)+(2)			ランダム	平均候補数
	一位	MRR	五位	一位	MRR	五位		
作業者 1	20.7	34.6	61.7	23.3	37.3	64.1	7.67	18.4
作業者 2	35.0	55.3	90.0	48.7	65.8	91.4	9.45	10.2
作業者 3	56.4	70.7	97.8	63.3	79.5	98.3	14.9	6.63

て」から「タ形+んだって」、「先制されちゃったんだって」から「未然形+れちゃったんだって」のようなパターンが獲得できる。そして、「更新した」(タ形)のような述語の末尾表現の候補として、「んだって」「んだ」「んだってさ」「よ」「そうだよ」などが得られる。この内どの表現が良いか決定するのがここでの課題である。ただし、書き換え前後の時制やモダリティ、肯定/否定を揃える。

次の二つの観点でスコアを計算し、その積でランキングする。

(1) 述語の主辞と末尾表現の結束性 (e.g.「超えた」+「んだって」): 3層のフィードフォワードニューラルネットワークで学習する。中間層のユニット数を 50 に設定し、中間層の活性化関数には tanh、出力層の活性化関数には softmax を使用した。入力層には述語文節の形態素の形態素情報 { 表層形, 基本形, 品詞大分類, 品詞細分類, 活用型, 活用形 } を素性として与え、出力層には述語末尾表現の 1-of-k 表現を与える。

(2) 末尾表現の共起性 (e.g.「んだって」+「らしいよ」): 3層の Elman 型リカレントニューラルネットワークで学習する。中間層のユニット数を 50 に設定し、中間層の活性化関数には tanh、出力層の活性化関数には softmax を使用した。入力層には現在の述語末尾表現の 1-of-k 表現を与え、出力層には次に出現する述語末尾表現の 1-of-k 表現を与える。

口語化 B のデータセットを使用し、(1)のみで述語末尾表現を決定したときの正解率、および、(1)と(2)の積をスコアとしてランキングしたときの正解率で評価した。実験結果を表 6 に示す。この結果から、主辞との結束性だけでなく、前に出現した末尾表現を考慮することで正解率が向上することが分かる。末尾表現の候補が多い作業者 1 の正解率は低いが、候補が少ない作業者 3 の正解率は他と比べて高い。

誤り例を見てみると「勝ち点 3 だって」を「勝ち点 3 らしいよ」、「行われるそうだよ」を「行われるんだって」としてしまうような誤りが多かったが、少なくとも時制や肯定/否定が揃っていれば、それほど致命的な誤りにはならないことが分かる。しかしながら、同じ口調でしゃべり続けてしまうと会話に魅力がなくなってしまうという問題もあるため、口語化の良し悪しは会話タスク上で評価した方がよいと考えられる。

5 おわりに

文書のような視覚メディアは書き言葉で書かれている。そのため、ただその内容をそのまま読み上げても聴覚メディアが持つ快適さを享受することはできない。そこで、ニュース記事のような文書を会話で伝える状

況を想定し、人手でニュース記事の内容を会話表現に書き換え、口語化コーパスを作成した。本稿では会話文を話し言葉たらしめている述語に着目し、このコーパスを利用して述語の書き換えに関する実験を行った。口語化では、述語の書き換えだけでなく、助詞の書き換え、接続詞や指示詞の挿入、文節順序の入れ替え、連体修飾節の述語化なども行われるため、今後はこれらの書き換えについても検討していく。

参考文献

- [1] 高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則: “快適な情報享受を可能とする音声対話システム”, 言語処理学会第 22 回年次大会発表論文集, 2016.
- [2] 高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則: “会話によるニュース記事伝達のための情報選択”, 言語処理学会第 22 回年次大会発表論文集, 2016.
- [3] 島弘巳: “話しことばの特徴 - 冗長性をめぐって -”, 国文学解釈と鑑賞, 52(7), pp.22-34, 1987.
- [4] 林由紀子, 松原茂樹: “聞きやすい読み上げ音声出力のためのテキスト変換の検討”, 情報処理学会第 69 回全国大会講演論文集, Vol.2, pp.581582, 2007.
- [5] 林由紀子, 松原茂樹: “ニュース記事の自然な音声出力のためのテキスト変換”, 言語処理学会第 14 回年次大会発表論文集, pp.790-793, 2008.
- [6] 林由紀子, 松原茂樹: “自然な読み上げ音声出力のための書き言葉から話し言葉へのテキスト変換”, 情報処理学会研究報告, 音声言語情報処理, Vol.47, pp.49-54, 2007.
- [7] 大泉敏貴, 鍛冶伸裕, 河原大輔, 岡本雅史, 黒橋禎夫, 西田豊明: “書きことばから話しことばへの変換”, 言語処理学会第 9 回年次大会発表論文集, pp.9396, 2003.
- [8] 鍛冶伸裕, 岡本雅史, 黒橋禎夫: “WWW を用いた書き言葉特有語彙から話し言葉語彙への用言の言い換え”, 自然言語処理学会論文誌, Vol.11, No.5, pp.19-37, 2004.
- [9] 別所史浩, 原田達也, 國吉康夫: “リアルタイムクラウドソーシングと Twitter 大規模コーパスを利用した対話システム”, 情報処理学会研究報告, Vol.2012-NL-206, No.13, 2012.
- [10] 目黒豊美, 杉山弘晃, 東中竜一郎, 南泰浩: “ルールベース発話生成と統計的発話生成の融合に基づく対話システムの構築”, 人工知能学会全国大会論文集, Vol.28, pp.1-4, 2014.
- [11] 高津弘明, 小林哲則: “対話エージェントのための性格モデル”, 言語処理学会第 21 回年次大会発表論文集, pp.191-194, 2015.
- [12] 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, Vol.29, No.1, pp.21-31, 2014.
- [13] “「やさしい日本語」”, 理前大学人文学部社会言語学研究室, <http://human.cc.hirosaki-u.ac.jp/kokugo/EJ3mokuji.htm>
- [14] “やさしい日本語の手引”, 島根県・(公財)しまね国際センター, <http://www.sic-info.org/sic/post-1318/>
- [15] T.Mikolov, K.Chen, G.Corrado and J.Dean: “Efficient Estimation of Word Representations in Vector Space”, in Proceedings of the International Conference on Learning Representations, pp.1-12, 2013.
- [16] T.Mikolov, I.Sutskever, K.Chen, G.Corrado and J.Dean: “Distributed Representations of Words and Phrases and their Compositionality”, Advances in Neural Information Processing Systems, Vol.26, pp.3111-3119, 2013.
- [17] 高津弘明, 小林哲則: “分散表現を用いた動詞・フレーズの含意関係認識”, 言語処理学会第 21 回年次大会発表論文集, pp.928-931, 2015.
- [18] 橋本力, 鳥澤健太郎, 黒橋禎夫, ステイン・デ・サーガ, 風間淳一, 藤田篤: “動詞含意関係データベースの自動拡張”, 言語処理学会第 16 回年次大会発表論文集, pp.47, 2010.
- [19] 神尾昭雄: “情報のなわ張り理論”, 大修館書店, 1990.
- [20] 神尾昭雄: “続・情報のなわ張り理論”, 大修館書店, 2002.
- [21] 岩畑貴弘: “「情報のなわ張り理論」再考”, 神奈川大学人文研究, 2010.
- [22] 島津明, 中野幹生, 坂浩二, 川森雅仁: “話し言葉対話の計算モデル”, 電子情報通信学会, コロナ社, 2014.