

係り受け構造及びシソーラスによる対話文生成と簡易演出技法を用いた Web コンテンツの受動的視聴

灘本 明代[†] 林 正樹^{††} 道家 守^{††} 浜口 斉周^{††} 田中 克己^{†,††}

[†] 独立行政法人情報通信研究機構

〒 619-0289 京都府相楽郡精華町光台 3-5

^{††} 日本放送協会放送技術研究所

〒 157-8510 東京都世田谷区砧 1-10-11

^{†††} 京都大学大学院 情報学研究科社会情報学専攻

〒 606-8501 京都市左京区吉田本町

E-mail: †nadamoto@nict.go.jp, ††{hayashi.m-hk,douke.m-eq,hamaguchi.n-go}@nhk.or.jp,

†††ktanaka@i.kyoto-u.ac.jp

あらまし これまで我々は Web コンテンツからキーワード群を抽出し、これらに基づいた対話文自動生成を行い、キャラクターアニメーションと音声合成を用いて Web コンテンツをトークショー型コンテンツへ変換することを行ってきた。本論文では、Web コンテンツの文の係り受け構造を用いて、文の分割、質問・応答やうなづき等の対話手法を決定するとともに、これら対話文を自動生成する。さらに、Web ページから難しいと思われる単語を抽出し、これらをシソーラスを用いて、上位/下位語、同義語、類義語等による、より解りやすい単語に言い換える手法も提案する。また、以上の手法により生成された対話文をトークショー型コンテンツに変換する際、筆者らが提案している APE (Automatic Production Engine) を用いる。これにより、対話者のもつ感情や、対話の種類(つつこみ、質問等)に即した演出を容易に付与したり、差し替えたりすることができるフレキシブルな演出技法を実現する。

キーワード Web コンテンツ, トークショー型コンテンツ, 対話文生成, 言い換え, 演出付加技術

Transforming Web Content into TV-program-like Content Based on Automatic Creation of Dialogue and Easy Direction

Akiyo NADAMOTO[†], Masaki HAYASHI^{††}, Mamoru DOUKE^{††}, Narichika HAMAGUCHI^{††}, and

Katsumi TANAKA^{†,††}

[†] National Institute of Information and Communications Technology

3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-0289 Japan

^{††} Japan Broadcasting Corporation, NHK Science & Technical Research Laboratories

Kinuta 1-10-11, Setagaya-ku, Tokyo, Japan

^{†††} Department of Social Informatics, Graduate School of Informatics, Kyoto University

Yoshida Honmachi, Sakyo, Kyoto 606-8501, Japan

E-mail: †nadamoto@nict.go.jp, ††{hayashi.m-hk,douke.m-eq,hamaguchi.n-go}@nhk.or.jp,

†††ktanaka@i.kyoto-u.ac.jp

Abstract We have described transforming declarative-based web content into humorous dialog-based TV-talkshow-like content that is presented through cartoon animation and synthesized speech. Our system has created dialog automatically based on only keywords of the web content. In this paper, we describe the method of dividing long sentence, and question and answering based on dependency relation. Then, the system transforms into dialog automatically by using these method. Moreover, we extract difficulty-degree of words, and paraphrase difficult word to the easy understand word by using the difficulty-degree. When the system transforms into dialog-based TV-talkshow-like content, we use the simple direction method called APE(Automatic Production Engine). By using APE, we can create direction easily and we can reuse the directions.

Key words Web Content, TV-talkshow-like content, dialogue based content, paraphrase, direction method

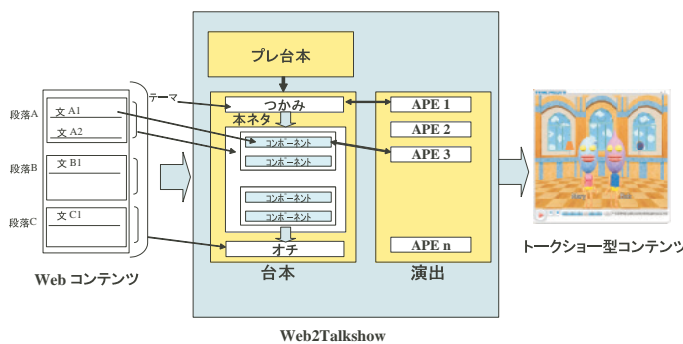


図1 Web2Talkshow の構成
Fig.1 Architecture of Web2Talkshow

1. はじめに

現在の Web コンテンツは利用者が「クリックする」「スクロールする」「読む」といった能動的操作を行うことにより情報を取得している．これに対し、テレビは利用者が「見る」「聞く」という受動的動作により情報を取得することを可能としている．そこで我々は、このような能動的操作を要求する Web コンテンツをキャラクターアニメーションと音声合成を用いてテレビ番組のような放送型コンテンツに変換することにより、「容易に」「楽しく」「片手間に」Web コンテンツを受動的に取得することが可能になると考え、Web コンテンツの受動的視聴を提案してきた [1]．しかしながら、通常の Web コンテンツは平叙文で書かれているため、放送型コンテンツにおいて Web コンテンツの平叙文をそのまま音声合成により読み上げたのでは、わかりやすく親しみのあるコンテンツになるとは言いがたい．そこで我々はさらに、Web コンテンツをよりわかりやすく親しみのあるコンテンツに変換するために、平叙文で記述されている Web コンテンツから対話文を生成しトークショー型コンテンツに変換することを提案してきた [2] [3]．本研究の目的は、対話文変換による Web コンテンツの理解支援と興味の喚起であり、つまりは、Web コンテンツを一般の人々誰にでもわかりやすく親しみのあるトークショー型コンテンツへの変換を目的としている．そこで漫才は一般に誰にでも親しみのあるコンテンツの一つであると考え、漫才メタファを用いて、ユーモアのある対話文生成を行うシステムである Web2Talkshow [2] [3] を提案してきた．Web2Talkshow の構成を図 1 に示す．

これまで Web2Talkshow において提案してきた対話文生成は、我々の提案する話題構造を用いて、Web コンテンツ全体のキーワードとなる単語群を抽出し、この単語群に基づく簡単な質問応答を生成してきた．しかしながら、単にキーワードに基づいた質問応答では、表層的な対話しか生成することができず、本当に利用者の理解支援及び興味の喚起を行うためのコンテンツに変換したとは言いがたい．そこで本論文では、さらに Web コンテンツの内容に踏み込んだ対話文生成手法を提案し、利用者にとってよりわかりやすく、より興味が沸くトークショー型コンテンツへの変換を行う．実際には、Web コンテンツを構成する文の係り受け構造と Web コンテンツのキーワードを用いて、

文の分割、質問・応答やうなづき等の対話手法を決定するとともに、これら対話文を自動生成する．さらに、理解支援の一つとして、Web コンテンツから難しいと思われる単語を抽出し、これらをシソーラスを用いて、上位/下位語、同義語等に置き換え、より解りやすい単語に言い換える手法も提案する．本論文で変換する対象の Web コンテンツは 1 ページとし、ニュースページを対象とする．

一方、我々の提案するトークショー型コンテンツはキャラクターアニメーションを用いて、テレビ番組のように視覚的にも訴え、利用者にとってより親しみのあるコンテンツに変換されている．このようにトークショー型コンテンツでは、キャラクターの動作やカメラワーク等コンテンツの演出も重要な要素になっている．Web コンテンツの内容に応じて動的に詳細な演出を生成するのが好ましいが、実際の演出は種々の要素が絡み合い複雑化しており、内容に応じた動的な演出付加は困難である．そこで Web2Talkshow では図 1 に示すように、プレ台本という台本のフレームワークを用いて、対話の種類に応じた演出を決定する．本論文では、演出の再利用と演出付加作業の軽減を目的とし、我々が提案している APE (Automatic Production Engine) を用いることを提案する．これにより、対話者のもつ感情や、対話の種類(つかみ、質問等)に即した演出を容易に付与したり、差し替えたりすることができるフレキシブルな演出技法を実現することが可能となる．

以下、2 章ではこれまでの Web2Talkshow とその実験を、3 章では対話文生成を、4 章では演出を、5 章でまとめと今後の課題について述べる．

2. これまでの Web2Talkshow と実験

2.1 これまでの Web2Talkshow

これまで我々は、我々の提案する Topic Structure を用いて Web コンテンツから対話文生成を生成を行ってきた [2] [3]．我々のこれまで提案してきた手法の概要を以下に述べる．

漫才の形式化

Web コンテンツからトークショー型コンテンツへの自動変換を容易にするために、漫才を形式化した．形式化された漫才は、前処理、台本、演出からなる．台本は「つかみ」「本ネタ」「オチ」からなる．「つかみ」は挨拶と Web コンテンツのテーマを述べる部分であり、「本ネタ」は対話文生成により Web コンテンツの内容を伝える部分であり、「オチ」は最後の笑いと終わりの挨拶を行う部分である．

キーワード群の抽出

本研究では対話生成時に Web コンテンツの話題の構造を示すキーワード群として我々の提案する Topic Structure を用いた．Topic Structure は以下のとおりである．

ページ P における Topic Structure TP はトピック t_i , $i \in \{1, \dots, n\}$ からなり、 t_i は主題語 s_i と内容語の集合 C_i の 2 つの組からなる．また、 C_i は複数の内容語 c_{im} , $m \in \{1, \dots, k\}$ で構成される．すなわち、 TP は以下のとおりである．

$$TP = \{t_1, \dots, t_i, \dots, t_n\}$$

$$t_i = (s_i, C_i)$$

$$C_i = \{c_{i1}, \dots, c_{im}\}$$

< 主題語の抽出 >

主題語は単語の出現頻度を用いて抽出する．対象となる単語は名詞のみとする．すなわち，主題語の候補となる単語 t は

$$tf(t) \times weight(t) > \alpha$$

となる．ここで， $tf(t)$ は P における t の出現頻度を示し， $weight(t)$ は品詞による単語の重みを示し， α は閾値を示す．

< 内容語の抽出 >

内容語は主題語との共起度の高い単語とする．我々はあらかじめニュースにおける単語の共起辞書を作成し，この共起辞書を用いて共起度を求める． P の主題語を $\{s_1, \dots, s_i, \dots, s_n\}$ とすると，各々の主題語 s_i において内容語の集合である $C_i = \{c_{i1}, \dots, c_{ij}\}$ を求める． c_{ij} は s_i との単語の共起度がある閾値 (β) 以上の単語である．内容語も名詞のみを対象とする．

このように，ページ P における主題語と内容語を決定する．よって， P の TP は $\{t_1 = (s_1, C_1), \dots, t_n = (s_n, C_n)\}$ となる．

対話文生成

対話文生成では以下のパターンの複数の対話タイプを提案した．

- 質問・応答

これまでは主題語を含む文において Topic Structure のみを使用した質問・応答を行ってきた．実際には，「イチロー（主題語）といえば？」「マリナーズ（内容語）」といったように，主題語と内容語に関する簡単な質問応答を行う対話を生成した．また，この主題語と内容語のセットをわざと間違えて使用することにより，笑いのある対話を生成することも行ってきた．

- 誇張

Web コンテンツ内の文に含まれる数字の部分の桁を一桁上げることにより，誇張を用いた対話を生成し，驚きを表してきた．例えば，「松井が 15 本目のホームランを打った」という記事に対し，A「松井がね，150 本目のホームランを打ったんだよ．」B「え？150 本目？」A「あ間違えた，15 本目だった」といったような対話を生成した．

- うなずき

Topic Structure や数字を含まない文に関しては，うなずきのみで対話を生成してきた．

2.2 ユーザ評価実験

これまで提案した Web2Talkshow をもちいて，Web コンテンツ（ニュースページ）から対話文を生成し漫オメタファを用いたトークショー型コンテンツに変換することの有用性を図るために 172 名の被験者を対象にしたユーザ評価予備実験を行った．Figure 2 に年齢別被験者数を示す．被験者のうち約 10% がコンピュータの研究者またはエンジニアであり，80% は一般のユーザであり，10% がまったく Web を使用したことのない人であった．すべての被験者は関西地区に住んでおり，日ごろから漫才に慣れ親しんでいる人々である．我々は，2 種類のニュースページを漫オメタファを用いたトークショー型コンテンツに変換し被験者に提示して以下のアンケートをとった．

(1) Web2Talkshow は面白かったですか？

(2) ニュースを漫才にすることはとっつきやすいですか？

(3) ニュースを漫才にすることはニュースの内容をわかりやすくしましたか？

各々の質問に対するアンケート結果を図 Figure 2 に示す．また，好意的なコメントの代表的なものを以下に示す．

- 10 歳未満の子供のコメント

Web 上のニュースサイトを見ないが，Web2Talkshow は漫才形式でニュースを伝えてくれるので，楽しくニュースを見ることができた．

- 60 歳以上のお年寄りのコメント

コンピュータを使うことができず Web を閲覧したことがないが，Web2Talkshow はテレビのように Web コンテンツを提示してくれるので，これだったら Web の情報を取得することができるようになりうれしい．

- その他の被験者のコメント 2

Web コンテンツのキーワードを用いた対話文生成は，オリジナルの Web コンテンツの内容を理解することが可能であるとともに，楽しみながら情報を取得することができた．

厳しいコメントの代表的なものを以下に示す．

● 対話の「間」が変である．漫才にとって「間」は大切であるので，対話の「間」をもっと上手に表現してほしい．

● 音声合成のアクセントが自然ではなくおかしいのが気になった．

● 対話が長くなると，オリジナルの Web コンテンツが何を言っていたのかわかりにくくなる．

これらのコメントから，対話の「間」と対話の長さの考慮をしなければならぬことがわかった．

本ユーザ評価実験では，我々のシステムを提示した後，実際の漫才師に漫才を演じてもらい比較を行った．90% の被験者が実際の漫才師の演じる漫才の方が面白いと回答したが，10% の被験者は Web2Talkshow の方が面白いと回答した．我々のシステムと漫才師との比較では以下のようなコメントをいただいた．

- 実際の漫才はやはりおもしろい．

● 実際の漫才ではその場で指定した Web ページから漫才を演じることが困難であるのに対し，Web2Talkshow はその場で興味のあるページを漫才に変換できるので面白い．

● コンピュータと人とを比べてみて初めて人間のすばらしさがわかった．

3. 対話文生成

本論文ではニュースサイトにおける Web コンテンツ（ニュースページ）1 ページを漫オメタファを用いてトークショー型コンテンツへ変換する．前節で述べた，これまでの Web2Talkshow とユーザ評価実験の結果に基づき，本論文ではさらに Web コンテンツの内容に踏み込んだ対話文生成を行う．実際には，変換時に変換元の Web コンテンツの話題の構造を示すキーワード群と文の構造に基づき文の分割や質問応答による対話文生成を行う．ここでは，Web コンテンツを要約した対話文を生成するのではなく，Web コンテンツすべての内容を対話文生成によ

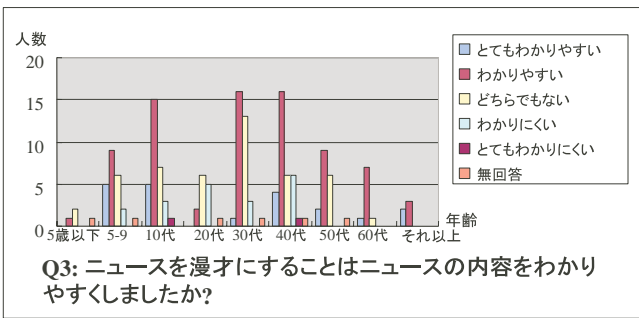
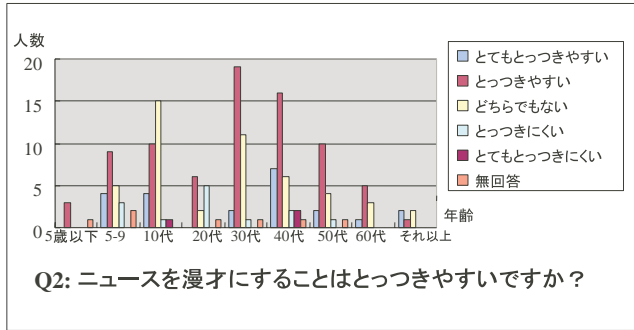
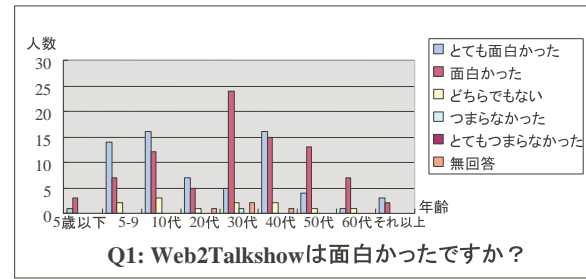
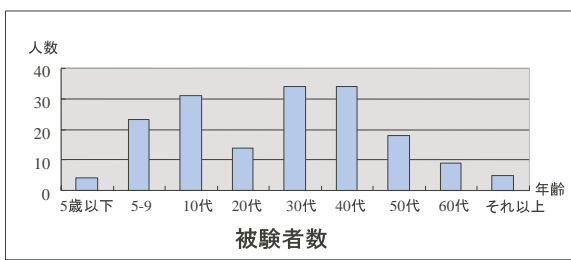


図2 ユーザ評価実験結果
Fig.2 Results of User Evaluation

てトークショー型コンテンツに変換する。

3.1 文の分割

Web2TalkshowではWebコンテンツの1文から一つの対話文を生成するが、一般にニュースページ等は1文が長い場合が多い。この長い文から対話を生成したのでは利用者にとってわかりにくく退屈な発話になってしまう。そこで、長い文は複数の文に分割し、その分割された文に基づいた対話文を生成することを行う。文には単文、重文、複文がある。本論文では、重文や複文を複数の単文に分割することを提案する。重文、複文は複数の動詞からなる文である。そこで、本論文では読点で区切られた各々の領域に複数の動詞を持つ文を分割する。

例えば

「イチローはジョージ・シスラーの記録を破り、年間262安打の新記録を作った。」

という文は、「破り」「作った」という2つの動詞からなり、これらの動詞は読点により分割されているため、上記の例は以下のように2つの単文に分割される。

- (1) 「イチローはジョージ・シスラーの記録を破り」
- (2) 「年間262安打の新記録を作った。」

ここで、(1)の文の「破り」は連用形であるため、これを基本形に変換するとともに、分割した(2)の文が過去形のため、さらに過去形に変換し「破った」とする。また、(2)の文は分割されたことにより主語が欠落している。分割前の文の係り受け解析を行うことにより、「作った」の主語は「イチロー」であることがわかるため、(2)の文に主語を付加する。このようにして、上記の例文は最終的に以下の2文となる。

- (1) 「イチローはジョージ・シスラーの記録を破った。」
- (2) 「イチローは年間262安打の新記録を作った。」

3.2 単語の言い換え

Webコンテンツは平叙文で書かれているため、一般的に対話では使用しない単語や理解するのが困難と思われる単語が含まれている場合が多くある。そこで本論文では、対話文生成を行う前に、文中に含まれる難解と思われる単語をよりわかりやすく親しみのある単語に言い換えることを行う。言い換えに関しては種々な研究が行われている[13][14]。本論文ではわかりやすい対話文生成のための言い換えを行うことを目的とする。そこで、単語のなじみ度と難解度を提案し、この難解度の高い単語をシソーラスを用いて、より難解度の低い単語に言い換え、対話文生成を行うことを提案する。

なじみ度

なじみ度とはこれまでによく見聞きする単語であり、一目見ただけで理解できる単語であると考えられる。つまりは、ニュースの場合、これまでによくニュースに出てくる単語であるといえる。しかしながら、過去におけるニュースの単語の出現頻度を求めることは実際困難である。そこで、共起辞書において共起する単語の共起度の総和の高い単語はこれまでに多くニュースに出現したと考える。つまりは、単語 w におけるなじみ度 (Friendly-degree) $FR(w)$ は

$$FR(w) = \sum_{k=1}^n cooc(w, X_k)$$

となる。ここで、 $X_k, k \in \{1, \dots, n\}$ は共起辞書に登録されている、単語 w の共起語すべてを示し、 $cooc(w, X_k)$ は単語 w と単語 X_k との共起度を示す。すなわち、 $FR(w)$ は w のすべての共起語各々の共起度の総和である。

例えば、「公務員」という単語の共起語は、我々の共起辞書では共起度の高い順から「準拠」(0.51)、「推進」(0.33)、「勤務」

(0.25)、「ボーナス」(0.25)、「是正」(0.25)、「戦後」(0.25)、「今夏」(0.125)、「大幅」(0.17)、「態勢」(0.16)となっている(括弧内は共起度を示す)。単語 w のなじみ度はこれらすべての共起語の共起度の総和であると仮定しているため、「公務員」という単語のなじみ度は 2.295 となる。提案するなじみ度は固有名詞の場合は満足の行くものであるが、我々の使用する共起辞書はこれまでのニュースサイトの単語による共起辞書であるため、例えば「棄却」といったように、普段我々は使用しないがニュースにおいてよく使用されるかたい単語はなじみ度が高くなってしまふという現象がおこり、なじみ度については今後更なる検討が必要である。

難解度

難解度が高い単語とは、なじみのない単語であり、その単語を見ただけで意味を理解することが困難な単語であると考え。また、一般に平叙文では用いるが、対話文では用いない単語も対話において使用すると理解しにくい単語になる可能性が高い。つまりは、理解しにくい単語は、漢字が難しかったり、意味的に難しかったり、対話においてあまり使われない単語であるといえる。そこで本論文では、以下のように難解度を測る指標を提案する。本研究では利用者の対象を特に子供やお年寄りに限定するのではなく、小学生からお年寄りまで幅広い利用者を対象としている。したがって、ここで述べる難解度が高い単語とは、小学生が理解できない程度の単語レベルを指す。

● 漢字が難しい単語

ニュースにおける読み仮名なしで使用する漢字は、例えば朝日新聞の場合常用漢字 1945 字に 66 字足して 2011 字となっている [4]。一方、小学生で習う漢字は学習漢字と呼ばれ現在 1006 字である。つまりは、小学生で習う漢字は新聞における約半数を網羅していることになり、学習漢字が理解できれば、新聞の半数の漢字は読めることになっている。そこで、我々は、年齢別に学習する漢字のレベル分けをすることを考え、日本漢字能力検定 [5] における級別に重み付けを行うことを提案する。

● 意味的に難しい単語

小学生の語彙数は [6] によると 5000 から 20000 語である。これらはすべて学校で習う単語ではなく、日常生活から知識として得ている単語もあるため、小学生がどの単語を理解してどの単語を理解できないかを示す基準がない。そこで我々は日本語を母語としない人の日本語能力を測り認定する試験である日本語能力試験 [7] において、レベルごとに使用される単語を用いて意味的に難しい単語を決定する。

固有名詞の場合、漢字が難しくとも一般に知れている土地名や人名、組織名である場合は難解度の高い単語であるとは言えないため、本論文では、固有名詞と固有名詞以外の品詞の単語に分類した難解度を提案する。

単語 w における難解度 (Difficulty-degree) $DF(w)$ は <固有名詞以外> 固有名詞以外の単語は漢字が難しいと判断された単語でも、ひらがなに直すと意味的にやさしい漢字であったり、またその逆もありうる。そこで、我々は漢字が難しい指標と意味的に難しい指標の両方を用いて難解度を決定する。

$$DF(w) = Kw + Sw + \frac{1}{FR(w)}$$

となる。ここで、 Kw は漢字検定級による重みを、 Sw は日本語能力検定の級による重みを示し、 $FR(w)$ は単語 w のなじみ度を示す。

<固有名詞> 難解な固有名詞は一般に知られていない地名や人名であると考え。ニュースにおいて頻繁に報道されている地名や人名は一般の人々にとってなじみがあり、難解な固有名詞ではないと考え、なじみ度の逆数を難解度と定義する。

$$DF(w) = \frac{1}{FR(w)}$$

単語の言い換えの決定

上記で求めた難解度がある閾値以上の単語を難解度の高い単語とし、シソーラスを用いて上位/下位語、同義語に言い換えることを行う。実際には、難解度の高い単語をシソーラスを用いて上位/下位語、同義語を求め、これらの中から最も難解度の低い単語に言い換えて対話文を生成する。

例えば、以下のような例文の場合

「箱根駅伝は、往路は東海大学が優勝した。」

ここでは、「往路」「優勝」が難解度が高い単語となった。この場合、シソーラスによると「往路」の上位語は「道」下位語は「往航」同義語は「行き」「行く道」となった。ここでこれらの上位/下位語、同義語の難解度の最も低い単語は「行き」となり、言い換えの単語と決定される。また、同様に「優勝」の上位語は「1位」下位語は「連覇」同義語は「制覇」であり、「1位」が難解度の最も低い単語となった。そこで、上記の例題は以下のように言い換えられる。

「箱根駅伝は、行きは東海大学が1位になった。」

上記のように単語の言い換えを行った後の文で対話文生成を行う。

実験

上記の提案に基づき予備実験を行った。予備実験では形態素解析に mecab [15] を使用し、共起辞書は我々が過去 2 年間で蓄積してきたニュースサイトにおける共起辞書を使用した。また、漢字が難しい単語の判定には漢字検定常用漢字辞典 [8] を、意味的に難しい単語の判定にはリーディングチュー太 [9] を、シソーラスには言語学研究所のシソーラス [10] を使用した。Yahoo ニュースページ [11] の 12 ページをこれらのツールを使用し予備実験を行った。予備実験に使用したニュースページのカテゴリは国内、海外、経済、スポーツであり、カテゴリごとに 3 ページづつ実験を行った。助詞や助動詞を除いた総単語数は 1547 単語である。我々の提案する難解度の手法でうまくいった例が 65% であり、うまく行かなかった例は 35% であった。固有名詞以外の単語でうまくいかなかった例は、経済のニュースで頻繁に使用されている「オプション」に代表されるような外来語に関しては漢字の難解度、意味的難解度とも低く、且つ頻繁に経済ニュースで使用されているため共起度が高く、したがって難解度も低くなるといった現象がおき、言い換える対象にならなかった。また、固有名詞は地理名に関しては、ほぼうまく難解度が得られたが、人名においては「鈴木」のように一般に多い

人名は難解度が低く提示されるといった現象が起こった。またシソーラスはあまり固有名詞が登録されていないため、人名や土地名の言い換えについては、あらたにシソーラスに登録しなければならぬことがわかった。

3.3 対話手法の決定

Web コンテンツのキーワードを Topic Structure を用いて抽出し、その抽出されたキーワード群と文の係り受け構造を用いて、質問・応答やうなづき等の対話手法を決定する。我々は、Web コンテンツ内において出現頻度の高い単語は、その Web コンテンツ内を構成する文においても重要であると考え、我々の定義する Topic Structure の主題語に注目し文の係り受けによる分類を行い対話手法を決定する。

主題語を含む文

< 主題語が主語の場合 >

- 内容語が主題語に係る場合

この場合、この文において主題語と内容語の関係は強いと考え、主題語と内容語の関係を問う質問・応答を行う。例えば「マリナーズのイチローがセフィコ球場で世界記録を作った」という文に対し主題語が「イチロー」であり、内容語が「マリナーズ」の場合、以下のような質問・応答を生成する。

A: イチローって誰だか知ってる？

B: マリナーズのイチローでしょう。

A: そのとおり、セフィコ球場で世界記録を更新したんだよ。

- 内容語が主題語に係っていない場合

この場合、この文においては主題語と内容語の関係はそれほど強くないと考え、主語である主題語に対して、目的語や述語を回答とする質問・応答を行う。例えば「イチローはセフィコ球場でマリナーズ対レンジャーズ戦に出場した」という文に対し主題語が「イチロー」であり、内容語が「マリナーズ」である場合、主語である「イチロー」に「マリナーズ」は係っていない。そこで、述語である「出場した」や目的語である「マリナーズ対レンジャーズ戦」を応答とする以下のような質問・応答を生成する。

A: イチローがどうしたのか知ってる？

B: 知ってるよ、出場したんしょう。

A: 何に出場したのか知ってる？

B: セフィコ球場でマリナーズ対レンジャーズ戦に出場したんだよ。

< 主題語が目的語の場合 >

この場合、目的語がこの文において重要であると考え、応答を主題語とする質問・応答生成を行う。例えば「マリナーズはイチローを解雇しないと決定した」という文に対し主題語が「イチロー」であり、内容語が「マリナーズ」であった場合「イチロー」が目的語であるため、「イチロー」を応答文とする質問・応答を生成する。

A: マリナーズは誰を解雇しないと決定したの？

B: イチローだよ。

A: ふーん

主題語を含まない文

< 場所や時間を含んでいる場合 >

場所や時間を含んでいる文はニュース記事において重要であると考え、Where や When 質問生成により、場所や時間を強調することをを行う。例えば「レンジャーズは 10 月 30 日、3-1 で勝った」という文に対し主題語と内容語は上記と同様であった場合、この文には主題語である「イチロー」が含まれていない。しかしながら、日時を示す「10 月 30 日」はこの文においては重要であると考え、以下のような質問応答を行う。

A: いつレンジャーズは勝ったの？

B: 10 月 30 日だよ

< その他の場合 >

文中に主題語が含まれていなく、且つ場所や日時を示す単語も含まれていない場合は、特に強調すべき単語のない文であるとみなし、Yes/No の回答の質問や付加疑問の質問・応答やうなづきを生成する。例えば「レンジャーズは 3-1 で勝った」という文の場合以下のような付加疑問文を生成する。

A: レンジャーズは勝ったんだよ。

B: そうだよ

実験

上記の提案に基づき予備実験を行った。予備実験では係り受け解析に cabocha [16] を使用した。先の予備実験で使用した Yahoo ニュースページ [11] の 12 ページを実際に係り受け解析を行い対話文を上記のルールに基づき予備実験を行った。ニュースの場合は、主題語が固有名詞の場合が多く、このときはほぼうまく対話文を生成することができた。しかしながら、主題語は出現頻度が多い単語であるため、何度も同じタイプの対話が生成されてしまうという現象が起こった。また、固有名詞以外の単語が主題語になった場合は、主題語のほとんどが目的語であるため、こちらも同様に同じタイプの質問が繰り返される現象が起こった。この結果から、対話と対話とのつながりやつながぎを考慮しなければならないことがわかった。

3.4 プレ台本

上記のように Web2Talkshow は対話文を生成するが、実際には「ふーん」や「そうだよ」といった対話中のうなづき等は対話にとって重要であるが、自動で生成するのは困難である。また、我々は Web2Talkshow では漫才メタファを用いてトークショー型コンテンツに変換する際、「つかみ」「本ネタ」「オチ」と行ったように漫才を形式化し、話の流れが明確なコンテンツに変換することを提案してきた [3]。「つかみ」ははじめの挨拶であり、「本ネタ」で対話文を自動生成し「オチ」は最後の挨拶という構成である。この「つかみ」と「オチ」の挨拶の部分も自動生成することは困難である。そこで、我々はプレ台本を作成し、これらの問題を解決することとする。プレ台本ではコンテンツの再利用を考え XML による記述を行う。プレ台本は構造タグとコンテンツタグと演出タグからなる。プレ台本の構造タグとコンテンツタグを表 1 に示し、プレ台本の例を図 3 示す。演出タグは次章にて説明する。

4. 演出

我々の提案するトークショー型コンテンツでは台本に基づい

表 1 プレ台本の XML タグ

Table 1 Pre-scenario Content Tags

構造タグ	
Initialize	プレ台本はこのタグにより囲まれる。
Intro	つかみを示すタグ。Intro タグは始めの挨拶と、変換元の Web コンテンツのテーマを述べる対話で構成される。
Dialogue	本ネタを示すタグ。Dialogue タグは対話を自動生成する部分であり、各対話の骨組みからなる。
Conclusion	オチを示すタグ。Conclusion タグは最後の笑いであるオチと終わりの挨拶を示す対話で構成される。
Content Tags	
line	キャラクターの台詞を示すタグ。chara 属性でどのキャラクターが話す台詞なのかを指定する。
question	質問・応答のフレームワークを指定するタグ。type 属性によりどのタイプの質問・応答かを指定する。
exaggeration	誇張を示すタグ。type 属性により誇張のタイプを指定する。

```

<?xml version="1.0" encoding="Shift_JIS" ?>
<initialize>
  <set up type="1">
</initialize>
<Intro>
  <dintro type="1">
    <line chara="1"> こんにちはボブです。</line>
    <line chara="2"> こんにちはマリです。</line>
    <line chara="1"> 今日は何の話題の漫才か知ってる?</line>
    <line chara="2"> 知ってるよ。今日は$themeの話題の漫才でしょう。</line>
    .
  </dintro>
</Intro>
<Dialogue>
  <question type="1-1" key="what">
    <LookAtCamera>
      <line chara="1"> ねえ、$sentence1知ってる?</line>
      <nod>
        <line chara="2"> もちろん!! $mis-answerでしょう!! </line>
        <line chara="1"> え!! アホやな。$answerじゃないか。 </line>
      </nod>
    </question>
    <question type="1-2" key="who">
      <line chara="1"> $sentence1?</line>
      .
    </question>
  </Dialogue>
<Conclusion>
  .
</Conclusion>

```

図 3 プレ台本の例

Fig. 3 Example of Pre-scenario

たキャラクターアニメーションを生成する。本論文では、キャラクターモデルの決定やキャラクターの動き、カメラワーク等の映像化するための設定を演出と呼ぶ。Web2Talkshow では、演出の設定の容易さと再利用を考え、APE (Automatic Production Engine) [17] と呼ばれるテレビ番組自動生成エンジンを用いる。

4.1 APE

APE は、テレビ番組の生成に必要な最低限の情報を XML 形式で入力することにより、これを TVML [18], [19] に変換する、一種の文字列変換処理をモジュール化したものである。図 4 の例では、簡単なニュース番組を生成する APE を示す。たとえば、< title >、< speech > などのタグを使ってニュースの情報を APE に与えると、TVML スクリプトを自動生成し、これが TVML プレイヤーによってニュース番組として映像化される。ここで、TVML はテレビ番組を記述する特別な言語であ

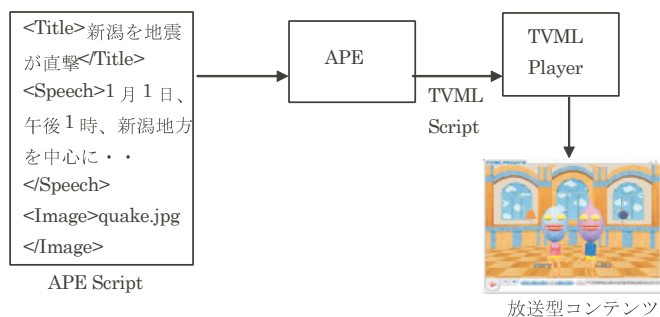


図 4 APE

Fig. 4 APE

り、TVML プレイヤーが TVML スクリプトを読み取り、キャラクターアニメーション、合成音声などを使いリアルタイムでテレビ番組を生成するものである。TVML は、比較的物理層に近い記述レベルだが、APE はより抽象度の高い記述レベルで番組を記述することが可能となる。TVML では、キャラクターのセリフ文字列のような「意味的内容」と同時に、たとえばキャラクターの動きの指定] や、カメラワークのパラメータ指定などの「映像化情報」である「演出」が混在している。そのため、テレビ番組を自動生成する際、「台本」と「演出」を切り分けることが難しい。APE というスキームを用いることにより、台本を自動生成するフェーズと、生成された台本に演出を自動付与するフェーズを切り離すことが出来るという長所がある。なお、APE は文字列を処理するブラックボックスであり、その実体は、C 言語や Pearl で書かれた文字列処理プログラムの実行モジュール (Windows であれば exe ファイル)、あるいは XML の変換処理を記述するスタイルシート (XSLT) などである。

4.2 Web2Talkshow への APE の適用

本論文の目的であるコメディ番組の自動生成に適するように、APE の入力 XML のタグを定義し、これをプレ台本の演出タグとする。表 2 にタグの一覧を示す。プレ台本に記述されたこれらのタグは APE に入力され、テレビ番組として出力される。ここで用いる APE そのものの作成は、表 2 の右側の欄に示されるような一連の CG アニメーション作成作業を必要とする。いわば、CG アニメーターへのタスクの指示を、表 2 に示したタグによって行っていると言える。台本の自動生成と、CG アニメーターの作業を、タグによって切り離すことで、完全に分業化することができる。このような APE のスキームを使う長所は、APE (すなわち演出スタイル) を同一にして、台本を次々に入れ替えるといった演出の再利用を可能とする。また逆に、台本を同一にして、APE を入れ替えることで好みの演出スタイルで映像化するという台本の再利用も可能とする。短所としては、限られた数のタグで台本と演出を切り離してしまったことで、台本内の意味的な情報に反応したキャラクタの動き、といった微妙な演出がこぼれ落ちてしまう。結果として、比較的単調な、定型化された映像になりがちである。

5. ま と め

本論文では、これまで提案してきた対話文自動生成手法から

表 2 プレ台本の演出タグ

Table 2 Pre-scenario Direction Tags

setup	CG セット, CG キャラクタ選択, カメラのセットアップ, 照明のセットアップ等を設定する.
dintro	コメディ開始までのイントロ部分の演出の設定を行うタグ. BGM と共にキャラクタが出てきて, 配置について会話を始めるまでの部分.
ending	会話を終えて, BGM と共に舞台を去っていく, エンディング部分の演出の設定.
LookAtCamera	カメラ視線
LookAtFello	対話の相手を見る
Hataku	ハリセンで相手をはたく
Tsukkomi	ツッコミの動作
nod	うなずく
no	首を振る
wait	間を取る
Lauph	楽しくわらう
Smile	明るくほほえむ
Hang	悲しげにうつむく
Fdown	暗い表情をして下を向く

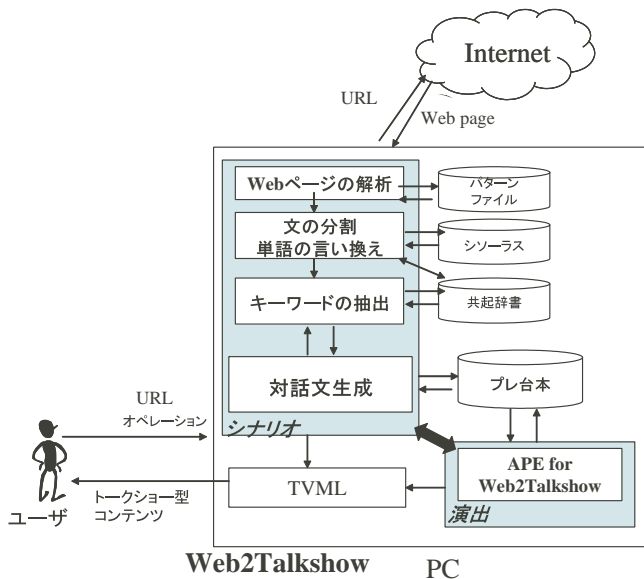


図 5 Web2Talkshow システム構成図

Fig. 5 Web2TV System Overview

さらに Web コンテンツの内容に踏み込んだ対話文生成手法を提案し, 利用者にとってよりわかりやすく, より興味が沸くトークショー型コンテンツへの変換を提案した. 実際には, 単語の難解度を提案し, この難解度を用いて, Web ページから難しいと思われる単語を抽出し, これらをシソーラスを用いて, 上位 / 下位語, 同義語等からより理解しやすい単語に言い換える手法も提案した. また, Web コンテンツを構成する文の係り受け構造と Web コンテンツのキーワードを用いて, 文の分割, 質問・応答やうなづき等の対話手法を決定の提案も行った.

今後の課題を以下に示す.

● 言い換え技法の更なる検討

本論文では言い換えの最初の段階として, 漢字の難しさと意味

的難しさからなる難解度を提案し, 難しいと思われる単語の言い換えを行った. しかしながら, 実際には難しい単語の言い換えだけでなく, 一般的な単語や注目度の高い単語もより親しみのある単語に変換し対話文を生成することが望ましいと考えられる. そこで, 今後より親しみのある単語へ変換する手法の検討を行う.

● より深層的な対話文の生成

本論文では, 係り受け解析を行うことにより, これまでより深層的な対話文を生成することができた. しかしながら, 実際の人間同士の対話はさらに深層的であるため, 今後より深層的な対話文の生成の検討を行う.

● 演出の自動付加の検討

本論文では演出付加作業の軽減と演出の再利用を考え APE による演出を提案した. しかしながら, 実際には Web コンテンツの感情や意味に沿った演出をある程度自動化できると望ましいと考え, 今後演出の自動付加技術の検討を行いたい.

文 献

- [1] 瀬本 明代, 服部 多栄子, 近藤 宏行, 沢中 郁夫, 田中 克己, “ Web コンテンツの受動的視聴のための自動変換とスクリプト作成マークアップ言語 ”, 情報処理学会論文誌: データベース (TOD8) Vol.42No.SIG1, pp.103-116, 2001 年 .
- [2] 蓬萊博哉, 瀬本明代, 田中克己, “ 理解しやすさとユーモアを考慮した Web コンテンツの対話番組変換 ”, 日本データベース学会 Letters, Vol.2, No.2, pp. 29-32, 2003 年 9 月.
- [3] 瀬本明代, 田中克己, “ 対話文自動生成による Web コンテンツの受動的視聴 ”, 情報処理学会研究報告, Vol.2004, No.72 2004-DBS-134(I), pp.183-190 2004 年 7 月 .
- [4] <http://tb.sanseido.co.jp/kokugo/kokugo/column/living-lang/bn/020401-1.html>
- [5] 漢字検定ホームページ <http://www.kanken.or.jp/index.html>
- [6] 「 図解日本語 », 林大監修, 角川書店
- [7] 日本語能力試験ホームページ <http://www.jees.or.jp/jlpt/>
- [8] 漢検 常用漢字辞典, 日本漢字教育振興会 (編集)
- [9] リーディングチュー太ホームページ <http://language.tiu.ac.jp/>
- [10] シソーラス辞書検索ホームページ <http://digilib.silkroad.net/thesaurus/index.php>
- [11] Yahoo ニュースサイトホームページ <http://headlines.yahoo.co.jp/hl>
- [12] Akiyo Nadamoto, Ma Qiang, and Katsumi Tanaka, “ Concurrent Browsing of Bilingual Web Sites By Content-Synchronization and Difference-Detection ”, Proceedings of the 4th International Conference on Web Information Systems Engineering (WISE2003), pp.189-199, Roma, Italy, Dec 2003.
- [13] 佐藤 理史, “ 論文表現を言い換える ”, 情報処理学会論文誌, Vol.40, No.7, pp.2937-2946, 1999
- [14] 峰脇さやか, 新見 道治, 河口 英二, “ SD 式意味モデルにおける概念体系を利用した短文の言い換えと評価方法の提案 ”, 情報処理学会研究報告, Vol.2002, No.104, 2002-NL-152, pp.71-78, 2002.
- [15] Mecab site homepage <http://chasen.org/taku/software/mecab/>
- [16] Cabocha site homepage <http://chasen.org/taku/software/cabocha/>
- [17] M.Hayashi, M.Douke, and N.Hamaguchi, “Automatic TV Program Production with APEs”, The 2nd International Conference on Creating, Connecting and Collaborating through Computer(C5), IEEE Press, pp20-25, 2004, Jan.
- [18] NHK 放送技術研究所: TVML ホームページ: <http://www.str1.nhk.or.jp/TVML/indexj.html>
- [19] 林 正樹: 番組記述言語 TVML を使った情報の番組化, 情報処理学会 DBS 研究会技術報告, Vol.2000, No.10, 2000-DBS-120-13, pp.91-98, 2000 年 1 月 .