

# 吹き出し型字幕の読みやすさ向上のための改行・改ページの挿入

木村 拓朗 南條 浩輝 吉見 毅彦

龍谷大学 理工学部 情報メディア学科

e-mail: kimura@nlp.i.ryukoku.ac.jp

## 1 はじめに

近年,さまざまな場面において障がい者に対する情報保障の充実が求められている。聴覚障がい者に対する情報保障では,聴覚的情報を視覚的情報に変えて伝達する手段が採られており,例として手話やノートテイクなどが挙げられる。現在これらの方法を用いた聴覚障がい者の情報保障は人手作業で行われており,時間や労力といったコストがかかってしまう。そこでコストを減らす技術として,音声を文字に変換する音声認識技術が挙げられる。この音声認識技術を用いた情報保障には,字幕付与システムなどがある。

字幕付与システムの中には発話者の発話を画面の下部や横に表示する字幕表示方式がある。しかしこの表示方式では話者が複数人存在する場合,発話者の特定が困難で,適切に情報保障を行えない可能性がある。これに対して,我々は発話者の顔の付近に字幕を吹き出し型で表示する方式(以下,吹き出し型字幕表示方式)を提案しており,発話者の特定が容易であることから,吹き出し型字幕表示方式の有効性を確認している [1][2][3]。

しかし吹き出し型字幕表示方式において,単にテキストを吹き出しに表示するだけでは読みにくさが生じ,聴覚障がい者の読解に支障をきたす恐れがある [4]。読みやすい字幕を生成するためには,吹き出しに表示するテキストの適切な位置で改行を行ったり,適切な位置で現在の吹き出しから新しい吹き出しに切り変える改ページを行ったりすることが求められる [5]。本研究では,吹き出しのテキストの適切な位置に改行や改ページを挿入することで字幕を読みやすくする方法を提案する。

## 2 吹き出し型字幕表示

吹き出し型字幕表示方式は,発話者の顔付近に字幕を吹き出し型で表示する方式である。吹き出し型字幕

表示方式の特徴として,発話者の特定が容易であることが挙げられる。既存の吹き出し型字幕表示システム [1] は 1 行に表示できる文字数が 6 文字であり,表示できる行数は 3 行である。これは,人間が 1 秒間に読む文字の量,吹き出しの大きさ,文字の大きさを考慮した結果,表示できる量を 18 文字としたものである。これまでテキストの表示方式は単純に 6 文字ごとに改行を行うという方式を採っていた。

一つの吹き出しに入る文字数の制限が強い吹き出し型字幕表示方式では,単純に 6 文字ごとに改行を行うだけでは読みにくさや分かりづらさが生じ,聴覚障がい者の情報保障に支障をきたしてしまう可能性がある。本研究ではこの 1 行 6 文字 × 3 行の制限に従いかつ,読みやすい表示方式の検討を行う。

## 3 改行・改ページ挿入

既存の吹き出し型字幕が読みにくい理由として,形態素の途中で改行が行われたり,新しい吹き出しに変わったりすることが原因と考えられる。そのため,改行や改ページ(吹き出しの切り替え)を行う位置が重要といえる。本研究では,形態素や文節の途中で改行や改ページが行われることをできるだけ避ける方法を検討する。

提案手法では,入力テキストの形態素境界,文節境界,節境界 [5] および係り受け構造に基づいて改行や改ページを行う位置を決定する。まず,節境界と係り受け構造に基づいて入力テキストを文節列という単位に分割する。その後,各文節列について,改行や改ページを行うかの判定を行う。本章では,文節列の生成方法,改行・改ページ条件,改行・改ページ挿入処理について述べる。

### 3.1 入力テキストの文節列への分割

提案手法では、文節列境界を、改ページを行う位置の候補とする。文節列境界は節境界と係り受け構造に基づいて決定する。

節境界を文節列境界とする理由は、節境界は、統語的にも意味的にもある程度完結した単位であり [5]、同じページ（吹き出し）に表示するのが望ましいためである。例えば、図1では、文節「いたしまして」の直後が節境界と判定されており、「いたしまして」と文節「店舗情報検索システムこれを」の間を文節列境界とする。

係り受け構造による文節列境界の決定では、隣接する文節が係り受け関係を持たない場合に別の文節列の構成要素とする。図1の例では、隣接する2つの文節「具体的に」と文節「本研究で」は係り受け関係を持たないので、この2つの文節の境界を文節列境界とする。次に、4つの文節「本研究で」、「取り上げた」、「タスクと」、「いたしまして」はそれぞれ隣接文節間で係り受け関係を持つので、これを1つの文節列とする。さらに、2つの文節「店舗情報検索システムこれを」と「構築いたしました」も隣接し、係り受け関係を持つので、1つの文節列とする。

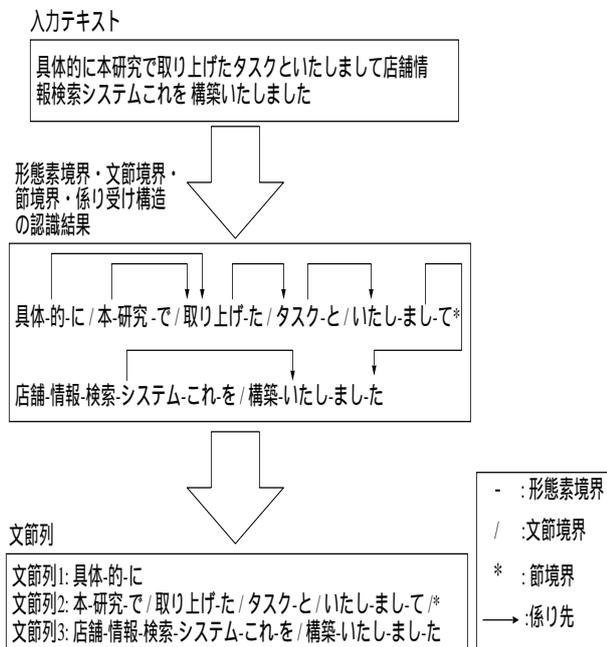


図1: テキストの文節列への分割

### 3.2 改行・改ページ挿入処理

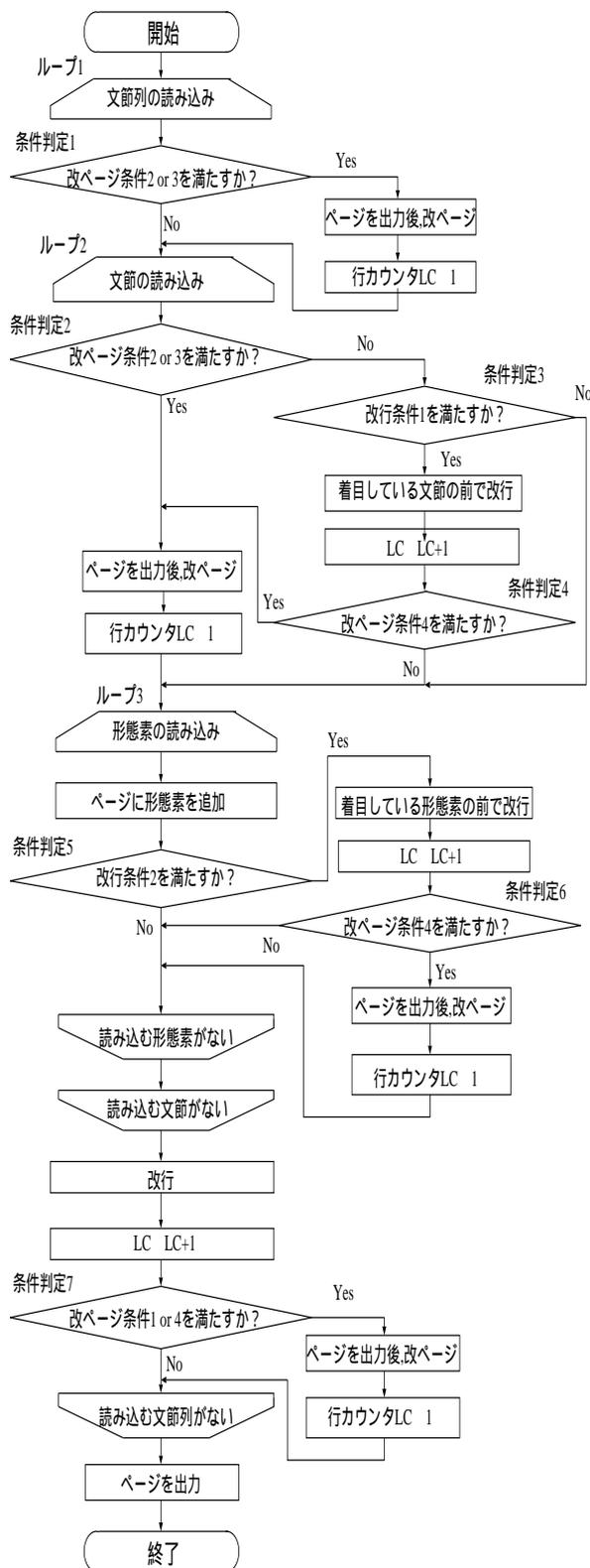


図2: 改行・改ページ挿入処理の流れ図

入力テキストから生成した文節列に対して、図2の流れ図にしたがって改行と改ページを行う。図2の処理で参照する改行・改ページ条件は次の通りである。なお、行カウンタ  $LC$  は現在のページ(吹き出し)において現在着目している行が何行目であることを示す。

改ページ条件 1: 節境界が検出された場合

改ページ条件 2: 行カウンタ  $LC$  が 3 であり、かつ、現在のページに文節列を追加すると文字数が 7 文字以上になる場合

改ページ条件 3:  $LC$  が 2 であり、かつ、追加しようとする文節列が行数を 3 行以上使用する場合

改ページ条件 4:  $LC$  が 4 の場合

改行条件 1: 現在の行が空でなく、かつ、現在着目している文節の文字数と現在の行の文字数との和が 7 文字以上になる場合

改行条件 2: 次に読み込む形態素が存在し、かつ、それを追加すると行の文字数が 7 文字以上になる場合

以下で、改行と改ページを行う手順について述べる。まず、文節列を読み込み、文節列が現在のページに入るかどうかを調べるため、改ページ条件 2 もしくは 3 を満たすかどうかのチェックを行う(条件判定 1)。このとき、文節列を追加するのに必要な行数は、文節境界の数、文節に含まれる文字数を手がかりに得る。

次に、現在着目している文節列の前から順に文節を読み込む。読み込んだ文節をページに追加すると改ページ条件 2 もしくは 3 を満たす可能性があるため、改ページ条件 2 と 3 のチェックを行う(条件判定 2)。改ページ条件 2 もしくは 3 を満たさない場合、読み込んだ文節を現在の行に追加できるかを調べるために、改行条件 1 のチェックを行う(条件判定 3)。改行条件 1 を満たす場合、 $LC$  が 4 となる可能性があるため、改ページ条件 4 のチェックを行う(条件判定 4)。

その後、現在着目している文節の前から順に形態素を読み込み、その形態素をページに追加する。追加した形態素の次の形態素が現在の行に追加できるかを調べるため、改行条件 2 のチェックを行う(条件判定 5)。このとき、改行条件 2 を満たした場合、改ページ条件 4 のチェックを行う(条件判定 6)。これは、現在着目している文節の文字数が 19 文字以上の場合への対応

である。読み込む形態素がなくなるまでループ 3 の処理を行う。読み込む文節がなくなるまでループ 2 の処理を行い、読み込む文節がなくなれば改行を行う。

改行を行ったことで  $LC$  が 4 となる可能性とループ 2 を抜けることで節境界が検出される可能性があるため、改ページ条件 1 と 4 のチェックを行う(条件判定 7)。読み込む文節列がなくなるまでループ 1 の処理を行う。

図 2 の手順に従って、図 1 の文節列を吹き出しに表示すると図 3 のような結果が得られる。

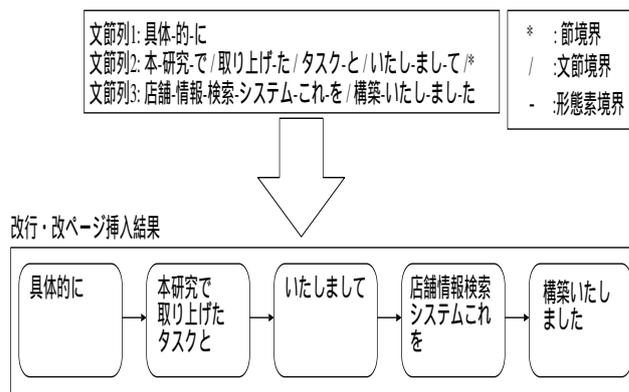


図 3: 改行・改ページ挿入結果

## 4 評価実験

### 4.1 実験方法

提案手法の有効性の検証には、日本語話し言葉コーパス [6] の講演データを用いた。この講演データから 14 講演を選び、さらに各講演から 10 文ずつ選び、計 140 文に対して提案手法を適用した。係り受け解析には CaboCha<sup>1</sup>を、形態素解析には ChaSen<sup>2</sup>を使用した。実験で使用する講演データでは節境界は人手ですでに付与されている。

提案手法の評価を行うために、読みやすさを考慮しながら改行・改ページ位置を付与した正解データを著者 1 名が作成した。この正解データと提案手法による改行位置、改ページ位置とを比較して、再現率・適合率・F 値を求めた。また、比較のために、6 文字ごとに改行を行う既存の表示方式についても同様の評価を行った。

<sup>1</sup><http://chasen.org/taku/software/cabocha>

<sup>2</sup><http://chasen.naist.jp/hiki/ChaSen>

表 1:改行位置についての再現率・適合率・F 値

	既存の表示方式	提案手法の表示方式
再現率	19%	78%
適合率	23%	70%
F 値	21%	74%

表 2:改ページ位置についての再現率・適合率・F 値

	既存の表示方式	提案手法の表示方式
再現率	30%	60%
適合率	37%	45%
F 値	33%	51%

## 4.2 実験結果と考察

改行位置についての評価結果を表 1 に示す。提案手法の F 値は 74%，既存の表示方式の F 値は 21%であり，提案手法のほうが高い数値を示している。このことより，提案手法による改行位置のほうが，読みやすい改行位置に近いといえる。

改ページ位置についての評価結果を表 2 に示す。提案手法の F 値は 51%，既存の表示方式の F 値は 33%であり，提案手法のほうが高い数値を示している。しかし，提案手法での改ページ位置の F 値は，改行位置の F 値に比べると低い。

提案手法の問題点として，図 3 に示した吹き出し「具体的に」のようにテキストが 1 行しか表示されないということが挙げられる。このように 1 行しかない吹き出しは，提案手法で生成された吹き出し 743 ページのうち 155 ページで見られた。このような吹き出しは「～という」「～ていう」「～して」などが節境界になっている箇所が生じやすいことがわかった。例えば，図 4 の場合「解くという」の後に「やり方」を追加することができるが「解くという」の後は節境界になるため改ページが行われる。提案手法では，現在のところ，節境界で無条件に改ページを行っているが，節境界の種類によっては，改ページを行わないほうがよいものもあると考えられる。今後，節境界を分類し，改ページを行わないほうが良い節境界を改ページ位置候補から除外することで，さらなる読みやすさの向上が期待できるのではないかと考えられる。

節境界での改ページ

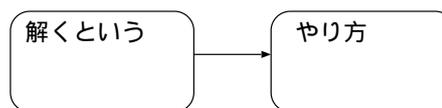


図 4: 提案手法の問題点

## 5 おわりに

本研究では，表示できる文字数に対する制約が強い吹き出し型字幕表示システムを対象として，読みやすい字幕を生成する手法を提案した。実験において正解データに対する再現率・適合率・F 値を求めた結果，既存の表示方式より提案手法のほうが高い値が得られることを示した。したがって，提案手法は，既存の表示方式よりも読みやすい字幕を生成できるといえる。

今後の課題として，改ページの場合の再現率・適合率・F 値は改行の場合に比べて高くないことから，改ページを行う位置の検討が挙げられる。

## 参考文献

- [1] 藤井絢子, 南條浩輝, 吉見毅彦. 会議の情報保障を目的とした吹き出し型字幕提示方式の検討. 情報処理学会研究報告, SLP75-14, pp.75-82, 2009.
- [2] Ayako Fujii, Hiroaki Nanjo, and Takehiko Yoshimi. Speech balloon captioning system for meetings based on automatic speech recognition. *Proc. 13th International Conference on Human-Computer Interaction (HCHI 2009)*, pp. 313–317, 2009.
- [3] 久木一平, 南條浩輝, 吉見毅彦. 会議の情報保障における吹き出し型字幕の有用性の調査. 情報処理学会創立 50 周年記念全国大会, 5U-4, 2010.
- [4] 中野聡子, 牧原功, 金澤貴之, 中野泰志, 新井哲也, 黒木速人, 井野秀一, 伊福部達. 音声認識技術を用いた聴覚障害者向け字幕提示システムの課題 話言葉の性質が字幕の読みに与える影響. 電子情報通信学会論文誌 (D), vol.J90-D,no.3,pp.808-814, 2007.
- [5] 村田匡輝, 大野誠寛, 松原茂樹. 講演テキストにおける読みやすさを考慮した改行位置同定. 情報処理学会研究報告, NL188, pp.37-44, 2008.
- [6] 前川喜久雄. 言語研究における自発音声. 日本音響学会研究発表会講演論文集 (春季), pp.19-22, 2001.