

物語生成システムにおける言語表現の諸相

大石 顕祐[†] 晴山 秀[‡] 小方 孝[‡]

岩手県立大学大学院ソフトウェア情報学研究科[†] 岩手県立大学ソフトウェア情報学部[‡]

1. はじめに

我々が構想している物語生成システムでは、物語の構成要素を物語内容 (story)・物語言説 (discourse)・物語表現 (representation) に分類している。物語内容は「何を語っているか」を意味し、登場人物や舞台等の設定・物語中で起こった出来事とその構成・主題等によって構成される。物語言説は「如何に語るか」の側面であり、視点・時間順序・距離・速度等の修辭的技法や叙法の技法により物語内容の語り方を意味する。そして物語表現は、言語・映像等表層的な媒体による表現そのものを意味する。

本稿では、物語表現における自然言語生成に焦点を当てる。具体的には、物語生成の観点から、自然言語生成の各レベルにおいて、物語テキストが命題内容の変化なしで、どのような要素によって変形されるかを考察した。また、この考察に基づき、特定の文法範疇に関する処理を行う試作を一部実装した。

以下、2 節では物語生成における言語表現生成の位置付けを示し、3 節では言語生成処理の概要と、必要とされる言語パラメータや語りの状況の概要を説明する。4 節では、この構想に基づき、アスペクチュアリティを対象とする動詞の接尾辞の変換システムについて説明する。

2. 研究の位置付け

図 1 に物語生成における言語表現の生成の位置付けを示す。

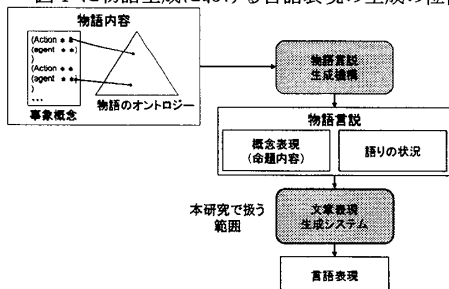


図 1 物語の言語表現の生成過程

物語内容として、行為 (action) を中心に、行為者 (agent)、被行為者 (counter-agent)、行為対象 (object) 等を含む深層格フレームで記述された事象概念と、関連情報を知識ベース (オントロジー) の形式で保存したものを想定する。物語言説生成機構は、物語言説を表現する情報として、命題内容と語りの状況を出力する。言説における命題内容とは、物語内容の情報を取捨選択や、記述順の時系列順から語りの順序への入れ替え等を行った情報であり、語りの状況とは語り手及び聴き手の時間的水準、性格、知識等、言説の主体に関する諸状況を意味する。

本研究で構想している文章表現生成システムは、命題内容と語りの状況を入力とし、物語の言語表現を出力する。

3. 文章表現生成システムの構想

3. 1. 文章表現生成システムの構成

語りの状況に応じて、命題内容を変えない範囲で異なる言語表現を多様・柔軟に生成することが目的である。システムは、

命題内容と語りの状況を入力とし、言語表現スクリプトを出力する言語表現スクリプト生成モジュールと、これを入力として言語表現を出力する言語表現モジュールから構成される (図 2)。

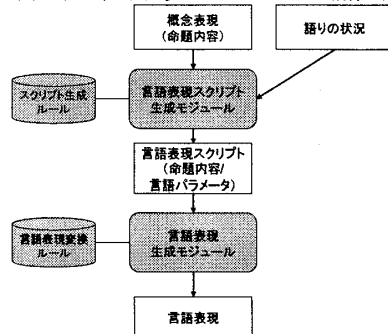


図 2 文章表現生成システムの構成

スクリプト生成ルールは、特定の命題内容と語りの状況の組み合わせに対して、命題内容の語彙・構文・文脈の各レベルの記述に対し言語表現の形式を制御するパラメータを追加することで、言語表現の具体的戦略を生成するルールである。この時追加されるパラメータを言語パラメータと呼ぶ。一方言語表現変換ルールは、特定の命題内容と言語パラメータの組み合わせから対応する言語表現を生成するルールである。

3. 2. 言語パラメータ

現在、スクリプト生成・言語表現変換の各ルールを設定するために、言語パラメータのリストアップを行っている。以下に言語パラメータのおおよその分類を示す。

①語彙生成に関連する言語パラメータ

語彙表現の生成に関連するパラメータは、行為内容及び格要素に付随し、内容語の語彙もしくは述語の接辞・副詞などの機能語の語彙を決定する。

機能語の内容を左右する機能・意味のカテゴリとして、テンポラリティ (出来事の時間的位置付け)、アスペクチュアリティ (動態的出来事の時間的展開の様態)、待遇表現がある [工藤 1995]。このうちアスペクチュアリティの表現を定義するパラメータには、行為の完了/継続、行為の時間展開、パーフェクト性、反復性といった表現の意味を制約するパラメータと、アスペクチュアリティの文法形式を制約するパラメータがある。

②構文生成に関連する言語パラメータ

単文・節の構文に関わるパラメータは単一の命題内容に付随し、主題・述語、語順を決定する。また、態・使役の交換、自他の動詞交換、授受動詞の交換等の言い換えのような、単文・節中の複数の語彙変化を伴う言い換えも構文レベルのパラメータが決定する。

③文脈的表現の生成に関連するパラメータ

2 つ以上の文にまたがる要素として、接続詞・接続助詞等の文・節間の関係を表す表現や、意味変化を伴わない省略や代名詞化を考慮する必要がある。文脈的表現の生成に関するパラメータは複数の命題内容を包括する形式で記述する必要がある。

④その他のパラメータ

本研究に関連する分野として言い換え処理が挙げられる。[Fujita 2008] による分類では、言い換え処理は、節間の言い換え、節内の言い換え、文法カテゴリを変える言い換え、主辞交

Aspects of the Linguistic Expressions in Narrative Generation System
[†]Kensuke Oishi: Graduate School of Software and Information Science, Iwate Prefectural University
[‡]Sakai Hareyama and Takashi Ogata: Faculty of Software and Information Science, Iwate Prefectural University

換、内容語の複合表現の言い換え、機能語/モダリティの言い換え、内容語句の言い換え、慣用表現の言い換えに大別されている。言語パラメータの構成のために参考になる。

また、フォントの大きさ・種類・色・文字と文字の間隔等、紙面のレイアウトに関する情報を扱うパラメータや、比喩的表現の使用、繰り返し、意図的な文法上の間違い等の修辞に関するパラメータについても今後考察する予定である。

3. 3. 語りの状況による制御

語りの状況としては強調・婉曲したい要素、人称、出来事に対する時間・空間的な注視点、語り手の時間的位置、語り手の語彙・文法知識の範囲、語り手の嗜好等が考えられる。これらによって適切な言語パラメータを選択することが、スクリプト生成ルールの機能である。(4.3 等で、特定の文法範疇を対象とした例を示す。)

4. アスペクチュアリティに基づく動詞の接尾辞の変換

文章生成システムにおける語彙、特に機能語の生成処理の試作として、アスペクチュアリティに基づく動詞の接尾辞の変換システムを一部(文法的変換処理の部分)実装した。

アスペクチュアリティの表現として、「スル-シテイル」という語尾変化を起こす文法カテゴリであるアスペクト、「シテアル」「シテオク」等の準アスペクトの他、「シヨウトスル」「スルトコロダ」のように組立形式で表されるもの、「シハジメル」「シツツケル」のような派生動詞による表現、動詞自体の範疇性、副詞による表現等が挙げられる。

4. 1. 試作システム

アスペクチュアリティの表現のうち、動詞の活用と接辞の接続によって表すことのできる表現を対象としている。活用語尾の変形は語幹によって決定されるため、'e'と'i'で終る二種類の母音動詞、s, k, g, m, n, b, t, r, w で終る九種類の子音動詞と変則的な変形パターンであるタ系語幹と前述したアスペクチュアリティの表現の対応関係を考察し、三段階の処理と語幹毎に違う4つの変形処理パターンに分類した。この分類に基づき、入力された動詞文の述語に対して指定されたアスペクチュアリティの表現を接続した文を生成する機能を実装した。

4. 2. 入出力

入力として、変換対象の動詞文と、変換する接尾辞の形式を指定する。出力として、入力文の動詞の接尾辞のアスペクチュアリティの表現が変換されたものが出力される(図3)。

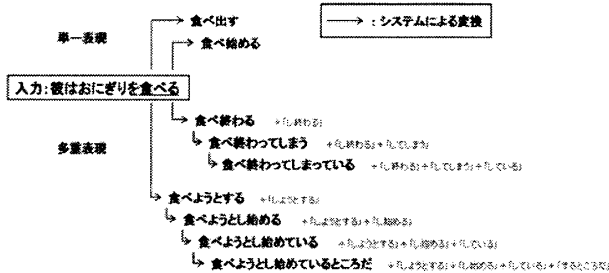


図3 入出力の例

本システムのテスト結果として、309種類の動詞と各接尾辞の組み合わせの入力データ1203通りに対して、構文解析器が解析不能の動詞が含まれた6通りの場合を除く1196通りの文章の接辞変形が可能となった。

4. 3. 課題-語りの状況変化に伴う接尾辞変換-

試作システムでは変換する接尾辞を直接指定しているが、今後、語りの状況によって言語パラメータを変更し、パラメータに応じた接尾辞を指定することで、語りの状況に応じた表現を出力する機能の実装を予定している。このモデルでは、語りの状況として語り手自身の時間的な立場が時制のパラメータを、語り手が出来事の時間展開のどの部分に注視しているかが出来事の時間展開のパラメータを変化させる。ここでは、時制のパラメータは「過去-現在-未来」、出来事の時間展開のパラメータは「動作開始前-動作開始時-動作途中-動作終了時-動作終了後」のいずれかの値を取るものとする。

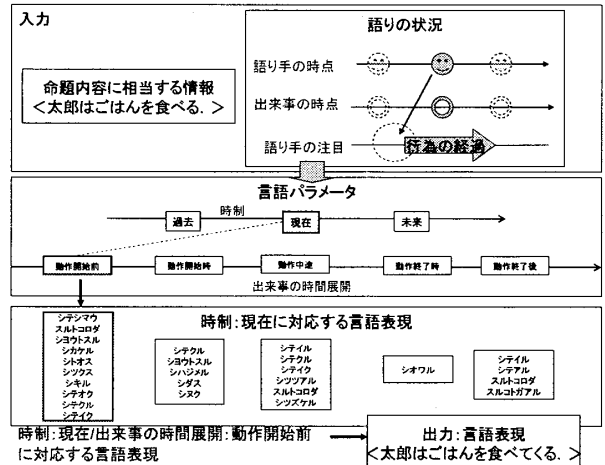


図4 語りの状況変化に伴う接尾辞変換の例

図4の例では、語り手・出来事の時間的位置が一致している時制が現在になる、語り手が出来事の動作前に注目している時出来事の時間展開が動作前になるという、2つのスクリプト生成ルールが成立している。

また、実装済みの接尾辞変形機能は言語表現生成モジュールに相当し、パラメータが現在・動作開始前に対応する接尾辞を付与するという言語表現変換ルールが成立している。

5. おわりに

本研究では、物語生成システムの一機能としての言語生成の問題を考察し、物語テキストが命題内容の変化なしにどのような要素によって変形されるかを考察し、システム構想を提案した。このシステムは、命題内容と語りの状況から使用する言語パラメータを決定して言語表現スクリプト、さらに言語表現を生成する。またこの具体化としてアスペクチュアリティを対象とする試作を示した。

自然言語生成システムの基本的な設計では、システムは文章の内容と構造を決定する文章プランナ、これによって選ばれた内容と構造をどの単語・統語構造を用いて伝えるのかを決定するマイクロプランナ、これによる抽象表現を実際のテキストに写像する表層生成器の3つのモジュールから成る[Ehud 2000]。文章プランナは物語生成システムにおける物語内容・物語言説の生成に対応するが、その処理の内容は物語特有のものであり大きく異なる。またマイクロプランナと表層生成器が文章表現生成に対応するが、語り手の状況と具体的な言語表現とを関連付ける処理は本研究に特有のものである。

今後は、本稿で提案した概念レベルの構想を具体化するための知識内容の検討と整理を進め、実装の範囲を拡大して行く予定である。

参考文献

- [Ehud 2000]Ehud Reither and Robert Date: Building Natural Language Generation Systems, Cambridge University Press, 2000.
- [Fujita 2008]Atushi Fujita: 言い換えのあれこれ, <http://paraphrasing.org/paraphrase.html> 最終アクセス:2009, 1, 13.
- [工藤 1995]工藤 真由美: アスペクト・テンス体系とテキスト—現代日本語の時間の表現—, ひつじ書房, 1995.