

## Hitch Haiku ～コンピュータによる俳句創作支援システム～

Hitch-Haiku: Japanese Haiku Poem Creation Support System by Computer

土佐尚子<sup>†</sup>, 尾原秀登<sup>†</sup>, 美濃導彦<sup>†</sup>, 松岡正剛<sup>††</sup>  
Naoko Tosa<sup>†</sup>, Hideto Obara<sup>†</sup>, Michihiko Minoh<sup>†</sup> and Seigo Matsuoka<sup>††</sup>

**Abstract** The authors maintain that there are unique and common forms of cultures that humans set aside such as the behavior and grammar within each culture. We suggest a computer model of this and a method of interactive expression and experiencing cultural understanding using IT called "cultural computing". This paper describes Hitch Haiku, a system based on cultural computing that interactively aids users in generating haiku, the world's shortest poems with imagery-maximizing mechanisms. First, "kire-ji" and particles are added to the word (s) input by a user to make a five or seven-character phrase, then phrases including terminology related to the user's input are located in a phrase database holding examples of haiku from the literary calendar, ensuring the cultural validity of the haiku. These phrases are then "hitched" together to generate a haiku. Although the haiku generated by this system have periodically resulted in flawed haiku, the ability to generate haiku that support the expansion of users' minds has been confirmed through assessment experimentation.

キーワード：カルチュラルコンピューティング、俳句生成、アート&テクノロジー、編集工学

## 1. ま え が き

従来、文化とコンピュータの関係は、失われていく文化をコンピュータでアーカイブ<sup>1)</sup>として保存したり修復<sup>2)</sup>をしたり、または歴史をシミュレーション<sup>3)</sup>するなどの視点でしか考えられてきていない。また歴史的文化理解は、本を読んだり、博物館で文物を見てその文化を理解するしか方法がなかったが、異なる文化の場合はなかなか理解がたい。そこで筆者らは、新たに「カルチュラルコンピューティング」の方法を提案する。「カルチュラルコンピューティング」とは、人間が歴史の中で培った各文化の中で行為や文法などの形で蓄えてきたものには文化に固有のまたは文化に共通の形式があると考え、これらを対象に文化のしくみを抽出し、コンピュータでモデル化を行い、情報技術を用いてインタラクティブ表現する文化理解体験方法を概念とする。

特に筆者らは、日本文化のコンピューティングに注目し、①気象・自然風土から培われる思想や美意識、②日本文化

2007年7月6日受付, 2007年10月8日再受付, 2007年11月6日採録

<sup>†</sup> 京都大学 学術情報メディアセンター

(〒606-8501 京都市左京区吉田二本松町, TEL 075-753-9081)

<sup>††</sup> 編集工学研究所

(〒107-0052 港区赤坂7-6-64, TEL 03-3568-2100)

<sup>†</sup> Academic Center for Computing and Media Studies, Kyoto University

(Yoshida Nihonmatsu-cho, Sakyo-ku, Kyoto 606-8501, Japan)

<sup>††</sup> Editorial Engineering Laboratory

(7-6-64, Akasaka, Minato-ku, Tokyo, 107-0052, Japan)

とアジア文化との関係性、③神仏習合を根底とした文化構造、④和歌、俳諧や能などの日本語独特の特性、⑤日本の意匠(紋・識・色・型・能・歌舞伎など)を対象に研究している。文化のしくみを抽出する際には、著者らの専門である編集工学<sup>4)</sup>、アート&テクノロジー<sup>5)</sup>的見地から、研究を進めている。

このような「カルチュラルコンピューティング」的方法を用いることにより、今まで定量化できなかった主観・感性・民族性・物語性といった人々に内在する文化の本質をモデル化・構造化する。そして、現在のコンピュータのコミュニケーション能力を拡張し、文化に基づく感情・意識・記憶の違いを反映し人々に内属する民族性まで深くゆすぶるコミュニケーションを目標としている。これらが実現すれば、言語・音声・映像を通したカルチュラルなコミュニケーション・メディアで社会的実用な情報表現システムが幅広い分野で可能になるであろう。

「カルチュラルコンピューティング」では、文化そのものを扱うため、特定の文化を取り上げ、それを基礎としたシステム作りをする必要がある。われわれは日本文化を取り上げ、禅文化体験システム「ZENetic Computer」<sup>6)</sup>システムの研究開発を行い、国内外の博物館で展示を行った実績を経験している。本論文では、世界最短の詩でありながら、豊かなイメージを喚起する“しくみ”を持つ「俳句」を取り上げる。コンピュータを使った俳句関連の研究<sup>7)</sup>としては、吉岡らが季語の推定を行う研究を行い、さらに俳句を投稿

しユーザ同士が鑑賞できるシステムを構築している<sup>8)</sup>。しかし、吉岡らの研究は、俳句の解析が主であり、システム自体は俳句の創作は行っていない。これに対し筆者らは、俳句の“しくみ”に注目し、俳句が持つ「季語」辞書である歳時記と用例俳句、筆者の一人である松岡正剛氏が一夜一冊の本についての書評を千夜にわたって書いた「千夜千冊」<sup>9)</sup>の文章を用いて、俳句を生成する手法を提案する。松岡・土佐らは、類語分類表などの言葉のネットワークを動的に検索することで、入力に対する応答を構築する手法を用いて、連想的に英語や漢字の言葉の繋がりを表示する、“i.plot”<sup>11)</sup>という作品を制作している。発想がそのまま生きるような面白い俳句を生成するため、フレーズの生成においても同じ入力に対して一通りしか出力しないのではなく、i.plotのような動的な検索手法を用いて、一つの入力に対する複数の応答を構築し俳句を生成する。このような俳句の醍醐味を味わいながら作れる俳句作成支援システムの設計・構築・評価が本論文の内容である。

本論文では、まず、Hitch Haikuのコンセプトとインタラクション過程を述べる。次に俳句生成処理、データベースの設計について記述し、最後に評価と今後の課題について述べる。

## 2. Hitch Haikuのコンセプト

俳句は、5・7・5の音数による言葉の韻律から構成される日本古来の詩であり、世界最短の詩である。その中には新年、春、夏、秋、冬のいずれかの季節を表す「季語」と「や」、「かな」、「けり」などのフレーズを接続する「切れ字」が含まれている。俳句は、室町時代に生まれた上の句と下の句を繋げていく付け合いによる連歌を起源としている。その後17世紀に、松尾芭蕉が連歌の最初の部分である発句を独立させて俳句にした。芭蕉には「奥の細道」という著名な俳句旅行記がある。芭蕉の俳句は「風雅」である。つまり、卑俗を離れ「微」にあらわれるところに「情」を感じる過程がそのまま「句」になっていく。

人が俳句を作る時、しばしば固定観念などに縛られ、いい言葉が思い浮かばず創作が進まないことがある。そんな時、従来は歳時記を見て連想を膨らませていくのであろうが、素人ではなかなかうまくいかない。そこで「Hitch Haiku」では、本の文脈からフレーズをHitchして、関連する歳時記に自動的につなぎ、俳句を生成する俳句創作支援システムを構築することを狙った。

本システムを用いて俳句を生成するプロセスは3章で詳しく述べるが、簡単に概要を説明しておく。まずユーザは、松岡正剛氏を中心として構築中のNICT図書街プロジェクト<sup>12)</sup>における古今東西の書を納めた図書街マップに示されたキーワードから、興味を持ったキーワードを選択する。その結果、千夜千冊の本のキーワードに対応するテキストが表示される。その中からユーザが俳句のベースとなる任意の言葉を選択すると、コンピュータが膨大な歳時記の俳

句などから構成される季語データベースを用いて、フレーズをつなぎ(Hitch)俳句を生成する。

## 3. インタラクション過程の詳細

- (1) 興味のあるフレーズの選択：松岡正剛氏によって作成された古今東西の本を納めた仮想図書街を訪れたユーザは、街を散策しながら同著者の「千夜千冊」の本にリンクされている好きなフレーズを選択する(図1)。
- (2) 書評の選択：(1)で選んだフレーズに対応した千夜千冊中の書評の本のタイトル・著者名と書評の内容が現れる(図2)。
- (3) 俳句にする言葉の選択：ユーザは書評のテキストからペンアイコンを使って、任意の言葉をマーキングする(図3)。
- (a) 入力レベルは2種類あり、一回めでマーキングした言葉は緑色でマーキングされ、この緑色のマーキングされた言葉の中から、今度は実際の俳句に使

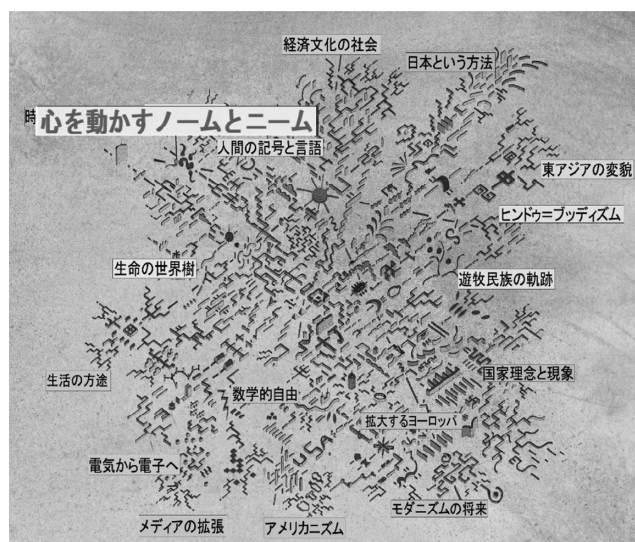


図1 図書街で選択したフレーズ  
Selected phase in the virtual library.

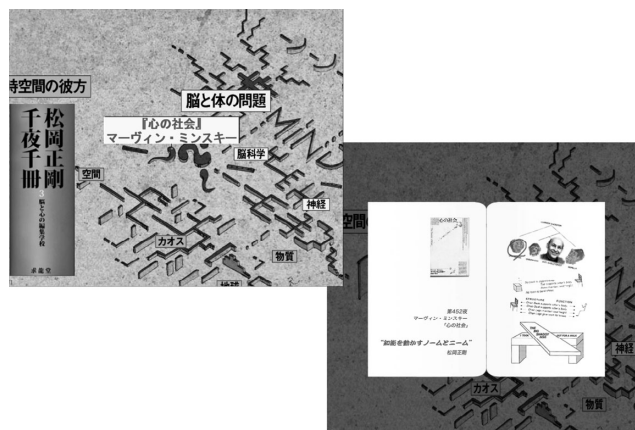


図2 選んだ本のタイトルと著者と載っている千夜千冊の巻  
Selected book, author and volume no. of 1,000 books 1,000 nights.

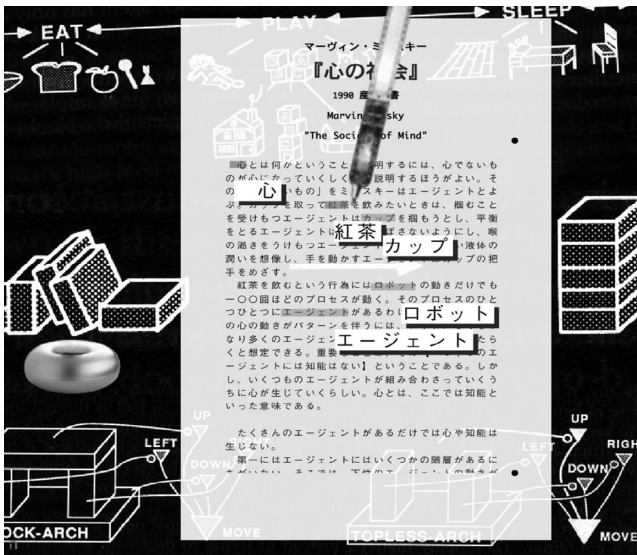


図3 テキストから言葉をマーキングして俳句の候補を挙げ「心」と「紅茶」を選択

Marking to text of books then select to phase "heart" and "tea" of Haiku Poem.

う言葉をつか二つマーキングすると、赤色でマーキングされる。システムは、緑色のマーキングされた言葉と、赤色のマーキングされた言葉のすべてから、関連のあるフレーズを選びだし俳句を作る。ただし、検索の重み付けは、赤色にマーキングされた言葉の方が高い。

- (b) マーキングの際に、誤入力・単語として成立しない文字列をマーキングした場合、消しゴム機能でマーキングを消し、再度マーキングできる。
- (c) 予備実験により、経験的にマーキングできる単語数の最小値は1、最大値は20と決定した。
- (4) 切れ字の接続：ユーザが入力した言葉を俳句らしくするために、「や」、「かな」、「けり」等の句末や句中に用いた切れの働きのある助詞・助動詞の切れ字が適切に選択され、語尾に接続される。
- (5) 季語の句の選択と俳句生成：俳句に文化的妥当性を与えるため、ユーザの入力した言葉から連想される季語を、俳句の辞書である歳時記と任意の俳人の俳句用例から作られたデータベースから検索し、より俳句らしい言葉が出てくるしくみになっている。
  - (a) その後、(3)で選んだ千夜千冊テキストから、選ばれた季語と関連のある言葉を選択し、本に関連した俳句を生成する。
  - (b) システムは、ユーザのマーキングした言葉から一意に決定されて、俳句が出てくるのではなく、関連する膨大な言葉の候補を生成し、候補の高い言葉の一つずつ、選択して表示する。
- (6) 俳句アニメーションの表示と効果音の選択：俳句の表示デザインは、任意の本のイメージによって決められている。文字のデザイン(日本語・英語の書体・レイアウト・表示アニメーション)と背景の映像と音



図4 創作された日本語・英語俳句とアニメーションの例  
Example of generated Hitch Haiku animation by Japanese and English.

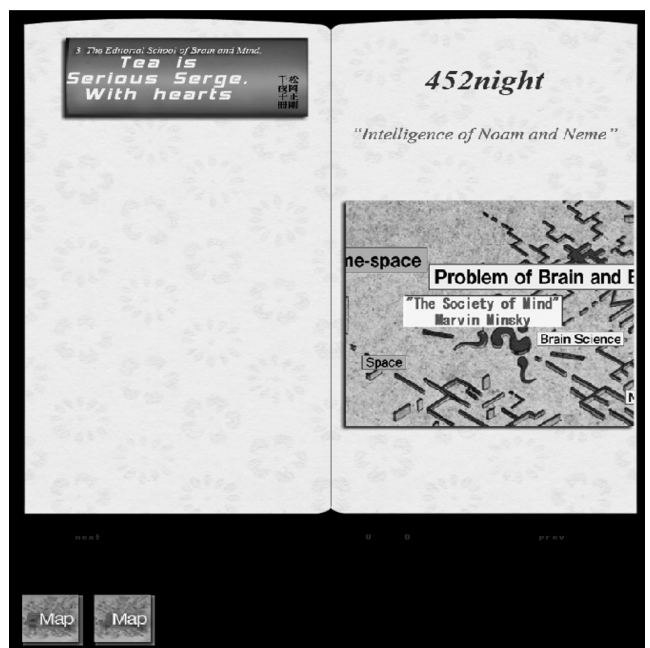


図5 俳句手帳に記録・学習し、DBに追加される  
Generated Haiku record to Haiku note also studied then add to DB.

楽が選択される(図4)。

- (a) 俳句表示のバリエーション機能として俳句表示右下の[next]をクリックすると、同じ言葉を使って次々と別の俳句を生成することができる。また、[EN]をクリックすると英語俳句に変わる。
- (7) 添削機能：ユーザが生成された俳句に対して上書きでき、より良いと感じるフレーズに修正することが可能である。
- (8) 俳句手帳機能：ユーザは作成した俳句を、俳句手帳に記録できる(図5)。
  - (a) 俳句手帳では、最初のページに千夜千冊の何夜であるか、本の内容を表現したフレーズ、本のタイトル、著書、本が存在する周辺の地図とともに表示される。
  - (b) 記録される俳句の内容に応じて、初期設定ではレ

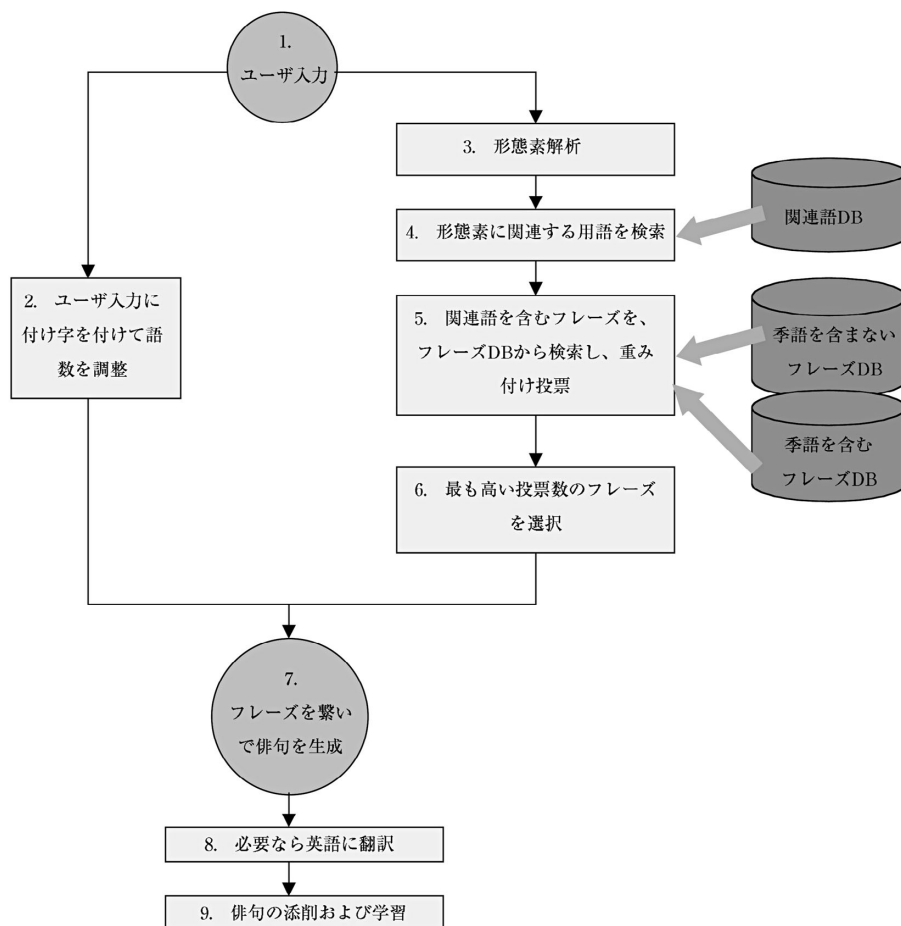


図6 俳句生成のためのフローチャート  
Flowchart of generating a haiku.

アウトを自動的に判断して配置している。後から、レイアウトを変更することもできる。

#### 4. 俳句の生成処理

ここでは、前章で述べたインタラクションのシステムの処理プロセスを述べる。俳句生成の全プロセスを図6に示す。俳句の生成においては、形態素解析、歳時記などによる関連語の検索などの主要な処理は日本語で行い、必要に応じて英語で処理を行う。以下、順を追って説明する。

- (1) まず、システムはユーザのマーキングした言葉の入力を受け付ける。
- (2) システムは、ユーザが入力した言葉が、5字または7字のフレーズになるように処理する。「ぞ」、「が」のように切れ字や助詞をユーザ入力の語尾に、あるいは「げに」のように副詞をユーザ入力の頭に付け字を、ユーザ入力の語尾または語頭の活用形に応じて付与している。
- (3) 同時に、システムはユーザが選び出した言葉に対して、形態素解析ソフトのMeCab<sup>13</sup>により、形態素に分け、名詞・動詞・形容詞・形容動詞等を抜き出す。
- (4) 抜き出した形態素に関連する言葉を、以下の7個の関連語のデータベースから検索する。類語DB、俳句類語DB、季語DB、慣用句DB、格フレームDB、擬態語

DBそして学習DBである。各々のデータベースは、分類名一つに、複数の用語を持つ。なお、ある言葉Sに関連する語とは、各々のデータベース内のSと同分類名で、S以外の用語とする。

- (5) 生成する俳句中のフレーズにすでに季語が含まれる場合は、季語を含まない俳句フレーズデータベースから、含まれない場合は季語フレーズデータベースから、関連語を含むフレーズを探す。フレーズが複数見つかった際は、類語による関連語を含むフレーズは重み3、俳句類語による関連語を含むフレーズは重み3、季語類語による関連語を含むフレーズは重み3、慣用句類語による関連語を含むフレーズは重み3、擬態語による関連語を含むフレーズは重み3、学習による関連語を含むフレーズは重み5、格フレームによる関連語を含むフレーズは重み1として、重み付け投票を行う。なお、これらの重みは、どのデータベースの関連語も偏りなく使われるように、事前の予備実験により調整している。
- (6) 最も投票数が多いフレーズを、ユーザ入力に関連するフレーズとして選び出す。
- (7) 付け字を付けられたフレーズまたはユーザ入力に関連するフレーズを、順に5, 7, 5字になるように繋いで俳句を生成する。

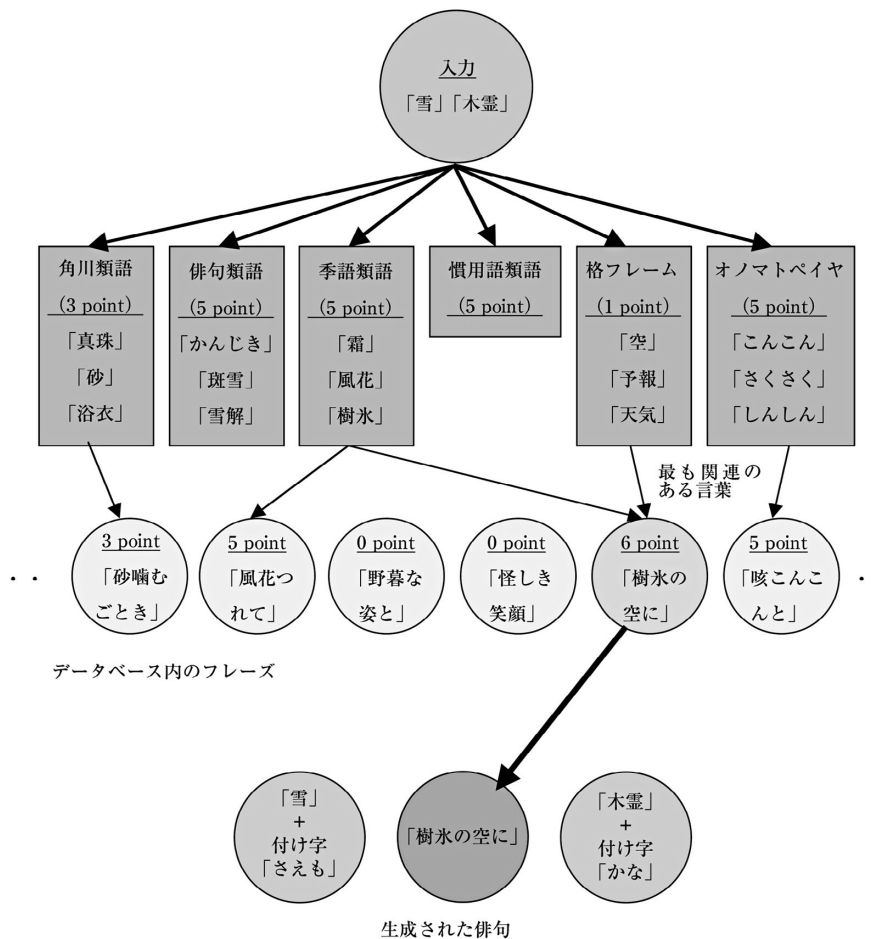


図7 最も関連のある言葉を選ぶ過程  
Selection of the most relative phrase.

- (8) 英語の俳句を生成する際は、以下の処理を行う。本システムで生成される俳句は古語を多く含み、例えば、「来ぬ」は「Has Come」と訳されねばならないが、通常の翻訳ソフトを用いると「Does Not Come」と誤って訳される。そこで、古語を現代口語に変換した後、日英翻訳を行うこととし、「古語－現代口語辞書」、「擬態語－英訳辞書」などを特別に作成している。日英翻訳のソフトには、さまざまな言語・コミュニティの辞書を所持することができ、さらに最適な辞書を選んで使用することができるため、NICTの言語グリッド<sup>14)</sup>を用いている。現在は、日英翻訳しか行っておらず、言語グリッドの利点あまり活かされていないが、今後多言語の翻訳を行う際に有用である。
- (9) より良い俳句生成のための学習機能：ユーザは生成されたフレーズを添削する際、システムはこの結果を用いて学習を行う。学習においては、添削されたフレーズは、まず俳句フレーズのデータベースに登録される。さらに、ユーザ入力と添削されたフレーズは関連があると仮定し、ユーザ入力と添削されたフレーズの形態素の関連性を関連語のデータベースへと登録する。図7は、ユーザが「雪」、「木霊」を俳句に使う言葉として入力した際の俳句生成処理の例

である。

- (a) まず雪、木霊に付け字をつけることで、生成される俳句のフレーズの文字数が5字または7字になるように調整する。
- (b) そして雪と木霊に関連する言葉をデータベース中より検索し、例えば、俳句類語の場合、かんじき、斑雪、雪解、などの関連語を見つける。
- (c) 雪という冬の季語が含まれるため、季語を含まない俳句フレーズデータベースから関連語を含むフレーズを検索する。
- (d) その結果、関連語である樹氷と空を含む「樹氷の空に」の投票数が6で最も高くなり、このフレーズをユーザ入力に最も関連のあるフレーズとして生成する俳句に用いる。
- (e) その結果「雪さえも 樹氷の空に 木霊かな」という俳句が生成される。
- (f) 英語の俳句を生成する際は、この段階で日本語の俳句をEven Snow/The Sky of Rime on Trees/Reverberateと翻訳する。
- (g) ここで、中の句を「樹氷の帽子」と添削すると、この添削されたフレーズが俳句フレーズのデータベースに登録される。さらに、このフレーズに形態

素解析が行われ、「樹水」と「帽子」が得られる。その後、分類名「雪」が「樹水」, 「帽子」に, 「木霊」が同じく「樹水」, 「帽子」に関連することが, 学習データベースに登録される。

## 5. データベース

システムを構築するにあたって, 以下のようなデータベースを用意した。まず, 関連語のデータベースは, 類語, 俳句類語, 季語, 慣用句, 格フレーム, 擬態語, 学習データベースの6種類ある。各々のデータベースのカラムは, 分類名と用語を持つ。また, 俳句フレーズのデータベースは, 季語を含むものと含まないものが2種類ある。季語を含まない俳句フレーズデータベースのカラムは, フレーズのみを持ち, 季語を含む俳句フレーズデータベースのカラムは, 季節とフレーズを持つ。

類語とは, 意味の似通った語を意味ごとに分類したものであり, 類語データベースは, 国立国語研究所の類語辞典<sup>18)</sup>中の分類名と用語を, それぞれデータベース内の分類名と用語として挿入することにより構築した。レコード数は約32,000であり, 例えば, 分類名「量の程度・限度」は, 用語「極端」, 「きわまり」, 「極度」, 「絶頂」, 「満足」などを持つ。

俳句類語とは, 俳句用語辞典中の見出しと見出しの説明の形態素には関連があると仮定し, これらの形態素を見出しごとに分類したものとした。俳句類語データベースは, これらの形態素を用語として, 見出しを分類名として挿入することで構築した。レコード数は約2,500であり, 例えば, 分類名「愛」は, 用語「求愛」, 「慈愛」, 「相愛」, 「溺愛」, 「偏愛」などを持つ。

季語とは, フレーズの季節を決める言葉であり, 季語データベースは, 季語をカラムの用語として, その季語の季節と分類名をカラムの分類名として挿入することにより構築した。レコード数は約13,000であり, 例えば, 分類名「春の天文」は, 用語「春雨」, 「初雷」, 「斑雪山」, 「東風」, 「陽炎」などを持つ。

慣用句とは, 長い文が一つの意味を表すようになって固定したものである。慣用句データベースは, 慣用句辞典の見出しを分類名として, 見出しの説明内の個々の形態素を用語として説明文の個々の形態素をカラムの用語として, 挿入することで構築した。レコード数は約1,300であり, 例えば, 分類名「泣く」は, 用語「泣きべそ」, 「涙に暮れる」, 「不覚の涙」, 「目頭」などを持つ。

格フレームとは, 見出しとその見出しに関係する名詞・動詞を整理したものである。格フレームデータベースは, 見出しに関係する名詞・動詞をカラムの用語として, 見出しをカラムの分類名として, 挿入することで構築した。レコード数は約31,000であり, 例えば, 分類名「ほほえむ」は, 用語「あなた」, 「人」, 「瞳」, 「顔」, 「女神」などを持つ。

擬態語とは, 擬音・擬声のことである。擬態語データベースは, 俳句擬音擬態語辞典の見出しである擬態語を分類

名として, 見出しの説明の個々の形態素をカラムの用語として, 挿入することで構築した。レコード数は約8,800であり, 例えば, 分類名「うんうん」は, 用語「りきむ」, 「承知」, 「了解」, 「同意」, 「一声」などを持つ。

学習データベースでは, 生成した俳句のユーザによる修正した学習データを格納することできる。学習データベースは, 俳句生成における, ユーザ入力をカラムの分類名として, 修正されたフレーズの個々の形態素をカラムの用語として, 挿入することで構築した。レコード数は, 学習すればするほど増える。

俳句フレーズのデータベースは, 季語を含むものは歳時記から約13,000フレーズ季語を含まないものは俳句類語辞典, 俳句慣用句辞典, 擬音擬態語辞典から約25,000フレーズ抜粋することで構築している。

## 6. 評価

本システムの評価を行った。2007年8月米国サンディエゴで行われたACM SIGGRAPH<sup>15)</sup> Art Showでシステムのデモを行い, 200人の来場者にシステム評価に関するアンケートに回答してもらった。日本人には, 日本語での俳句生成を, それ以外の人には, 英語俳句を評価してもらった。日本語俳句と英語俳句を一緒にアンケートをとったが, 日英の俳句の評価の差は特になかった。日本語俳句ができが悪い場合は英語俳句も悪く, 日本語俳句の生成がよい場合は英語も良いからである。

アンケート回答結果は, 以下の通りである。

- (1) 国籍: 回答者の半数は米国人, 次に日本人, カナダ人, 韓国人, 英国, メキシコと続き, 約20カ国の方々から, 国際的に回答を頂いている(図8)。
- (2) 年齢: 約40%が20代, 次に多いのは, 30代, 40代で, これだけで, 89%を占める。全体的には, 10代~60代以上まで参加している(図9)。

以下の7つの質問を行った。

- (1) 日本の俳句に興味がありますか?
- (2) ヒッチ俳句には, 興味がありますか?
- (3) 貴方は, ヒッチ俳句がセンスのある良い俳句を生成したと思いますか?
- (4) 貴方は, 自分自身で俳句を作った感じがしましたか?
- (5) 俳句生成と背景のアニメーションは, 合っていましたか?
- (6) 貴方がヒッチ俳句システムを使用した後, 俳句をもっと深く理解したいと思いましたか?
- (7) 貴方はヒッチ俳句システムを欲しいと思いますか?

質問のグラフ(図10)の結果から, 70%以上のユーザがシステムに関心を持っており, 約60%のユーザもセンスある俳句を生成したとしており, 比較的満足していると言える。

そして, 40%のユーザは確実に, Neutralまで入れると約80%のユーザが, 俳句を自分で生成した気持ちになってい

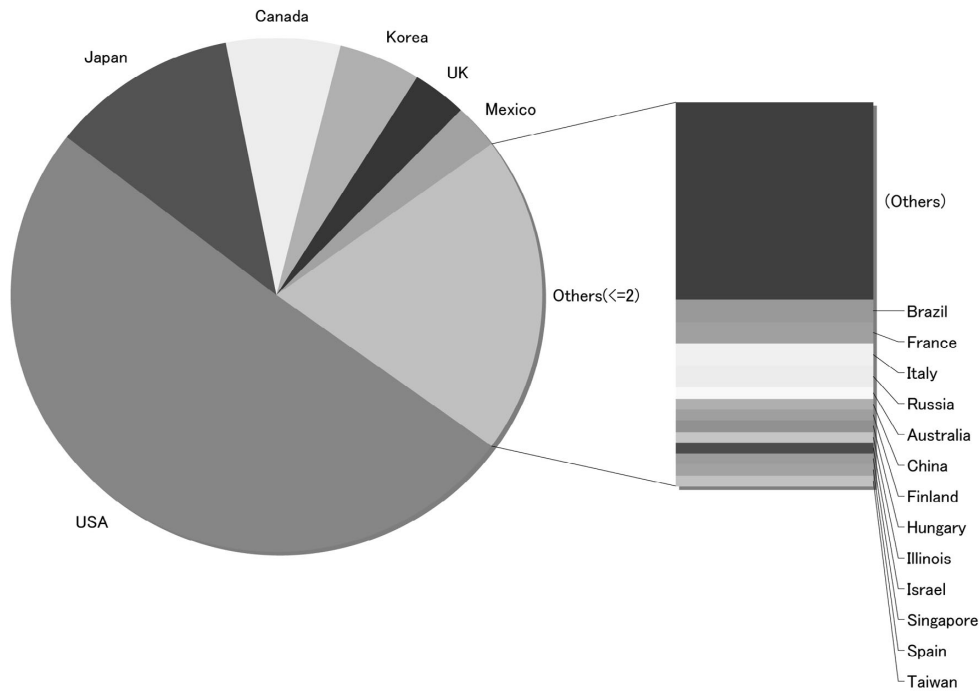


図8 評価実験に参加した約200人の国籍別グラフ  
Nationality of participants who tried Hitch Haiku.

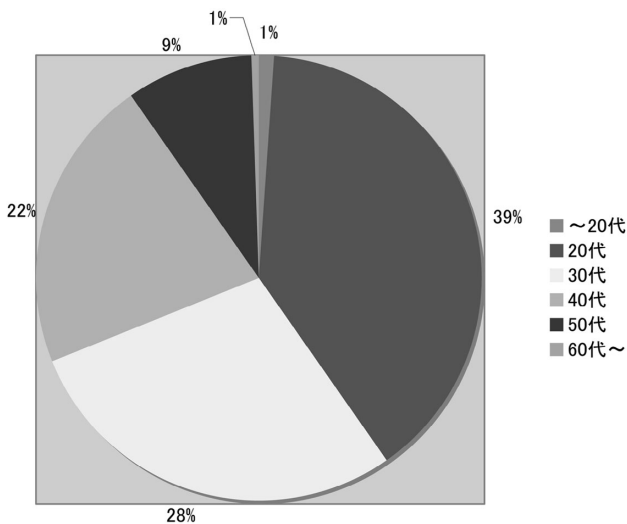


図9 評価実験に参加した約200人の世代別グラフ  
Age of participants who tried Hitch Haiku.

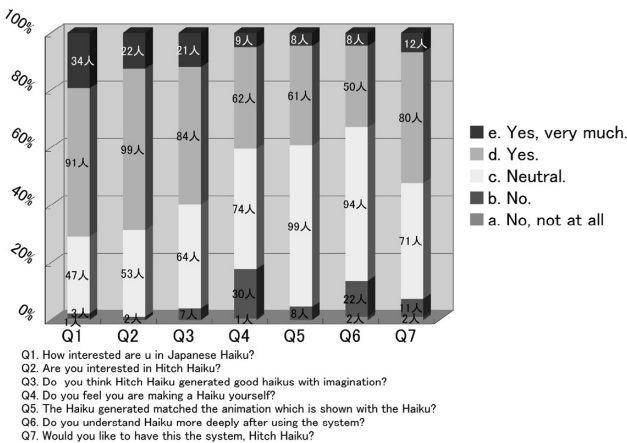


図10 アンケートの質問  
Questions.

る。この点は、システムが一つだけの俳句を生成するだけでなく、同じユーザが選んだ言葉を使って、連想的にシステムが沢山の俳句を生成し、ユーザがその中から、好きな俳句を選ぶことができる機能によるところが大きい。俳句生成と本に関するアニメーションの評価は、約40%のユーザが好意的な評価をしている。そして、システム体験後約30%以上のユーザが俳句に対する関心を深めている。最後のシステムを欲しいかどうかという質問は、約半数がYesである。

次に国別・世代別の比較分析した。質問2に関しては、国別にみると日本・韓国などのアジア圏が俳句システムに対する関心が高く、世代別では30代以上に関心が高い。

質問3に関しては、日本人は生成された俳句に対する見方が厳しい。日本人は、システムの言語表現に関する繊細な対応、つまり「侘び」、「寂び」のような表現のボキャブラリーが求めていると言えよう。世代別では、30代以上に好評である。

質問4の支援システムとしての評価で目立ったところは、国別・世代別で約半数が、Yesであり、米国人、30代にややNoが見られる。それに対して質問5のアニメーションと俳句のマッチングでは、国別・世代別でも一桁しかNoがない。

質問6の俳句へのより深い理解は、国別・世代別でも8割強が、まあまあ深まったとしている。

質問7のシステムが欲しいかどうかは、国別・世代別でも、約半数がyesで愛着を持っているという結果が出ている。その他に、ユーザからの「コメント・リクエスト」という項目では、沢山のコメントを頂いたので、傾向ごとに分類した。

### 【文化的視点】

「私は俳句が好きになった」, 「このシステムは私に日本文化に愛着を持たせた」, 「従来のストーリーテリング(本のテキスト)と新しいメディアのストーリーテリング(ヒッチ俳句システム)の間のインタラクションがとても面白い」, 「本の文章から言葉をとって俳句を作る試みは面白い」, 「生成されたヒッチ俳句のサプライズが好き」など, デジタルテキストを使った新しい文字文化を見出した。

### 【技術的視点】

観客は, ユニークなインタフェースの構造と, 俳句生成のメカニズムにとっても関心を持ち, このシステムの技術の探求をしていた。これは, 言葉のボキャブラリー表現の豊かなことから, 言語の勉強に最適なツールになるという意見もある。

### 【クリエイティビティの視点】

最も古典的な知であるアナログの書籍からデジタル俳句を作成する, とてもクリエイティブな試みとして, 従来の俳句を超える先端的なコンセプトが理解された。「深いレベルの知的好奇心を刺激する」, 「印象的でとても直感的なシステムで沢山の研究が詰まったプロジェクト」, 「沢山の可能性と美学をもったアプリケーション」というコメントをいただいている。

### 【エンタティメント的視点】

観客は, とても愛嬌のあるワンダフルなシステムとして, 言葉を選択して, コンピュータがどんな連想をするのかを見るのを楽しみ, またゲームみたいなエンタティメントとして面白く楽しんだようである。

### 【問題点】

一番の問題は, 芸術と感情を自動生成するのは難しい試みである。また, アイディアは面白いが古語の英訳表現が難しいので, 俳句の英訳の研究はもっとするべきという指摘を受けた。ユーザがもっと自由に俳句生成をコントロールできれば, 詩的なセンテンスが出てくるや, 音楽を変えたら, もっと良くなるというアドバイスがある。

### 【リクエスト】

「本の文章から俳句を生成するのではなく, 自分で言葉を入れて俳句を作りたい」, 「英語だけでなく, 他の言語でも俳句を作れるようにしてほしい」, 「インターネット上で公開してほしい」以上のような傾向のコメントが寄せられている。

これらのコメントから分析・考察すると, 基本的にヒッチ俳句システムというアイディアは受け入れられている。入力方法としては, 本のテキストから入力する方法だけではなく, ユーザ自身の言葉から俳句を生成するシステムの可能性もあると言える。また, 日本語の俳句を英訳する際, 基本的に「古語」, 「現代日本語(擬態語・慣用句含む)」, 「英語」という三つのデータベースを行き来するため, どうしても英語表現でうまく言い回しが難しい。また筆者らが日本人のため, 英語の詩のシラブルのルールなどに暗く,

詩的な英語俳句の感覚がつかみにくかった。今後, 英語のネイティブスピーカの詩人と共同研究することも必要であろう。また, 俳句生成には多大な関心が寄せられており, このようなメカニズムを持ったシステムが少なかったことがあげられる。

## 7. む す び

本研究では, 「カルチュラルコンピューティング」の研究の一例として, 日本文化の俳句の型に注目し, そのしくみをモデル化し, ユーザが選択した単語から俳句を自動的に生成する俳句生成支援システムについて述べた。本システムは, 米国サンディエゴで2007年8月に行われたSIGGRAPH Art Showで最初の公開展示を行い, 好評を博した。最近では, ブログ・携帯メディアが若者の関心を集めており, デジタルメディア上での言葉の表現に関心が集まっている背景もあり, 多くの人々の関心を集めることができた。また, 本システムにおけるコンピュータを使った俳句生成は, 従来の俳句作りを踏まえた上でその現代的な一つの発展した形を提示しており, 海外の人々に対しても日本文化の理解に役立ったと考える。本論文で提案した俳句生成システムは, 人間が文章を作る際に, 豊かなボキャブラリーを提供するための連想辞書のような役割を果たすため, 一般的には人間のイマジネーションを広げるシステムとして考えることも可能である。

「カルチュラルコンピューティング」という観点からの本システムに残された課題としては, まず, コンピュータがユーザに提供する俳句をより俳句的表現に近づけるための方法を模索することがあげられる。また, マルチメディア俳句研究として, 生成された俳句とマッチした映像や音楽を生成する方法の研究が必要である。さらに, 日本語の俳句を英語の俳句に英訳するための特殊辞書の精密化があげられるであろう。それらを通して, 人々の記憶や意識を, その奥にある民族性まで深くゆさぶる印象的なデジタルコンテンツとして洗練させていきたい。

最後に, 本研究は, 筆者の一人である編集工学研究所長の松岡正剛氏と慶應大学の金子郁容教授が進めているNICT図書街プロジェクトの一環であり, サポート頂いたNICTに感謝します。またコンピュータによる俳句の英訳に関しては, NICT言語グリッドプロジェクトリーダーであり, 京都大学情報学研究科の石田亨教授に感謝致します。英語俳句表現でコラボレーションをした国立シンガポール大学のエイドリアン・チョオク准教授, 同大学研究員のニュートン・フェルランド氏, 京都大学大学院情報学研究科修士課程のジェイ・コフマン氏に感謝いたします。最後に, この研究に数々の助言を与えて頂いた国立シンガポール大学の中津良平教授に感謝を致します。

## 【文 献】

- 1) Digital History, <http://www.digitalhistory.uh.edu/>



- 2) 兵馬俑の修復, [http://j.peopledaily.com.cn/2002/02/04/jp20020204\\_13973.html](http://j.peopledaily.com.cn/2002/02/04/jp20020204_13973.html)
- 3) K. Ikeuchi: "The Great Buddha Project", IEEE ISMAR03, Tokyo, Japan (Nov. 2003)
- 4) 松岡正剛: "知の編集工学", 朝日新聞社 (2001)
- 5) N. Tosa: "Affective Minds", Elsevier, 19, pp.183-201 (2000)
- 6) N. Tosa and S. Matsuoka: "ZENetic Computer: Exploring Japanese Culture", MIT Press, Leonardo, 39, 3, pp.205-211 (2006)
- 7) 鈴木雅実, 皆川直凡, 山本健一, 吉田香, 吉岡亮衛: "俳句への学際的アプローチ—クロスメディア俳句研究フォーラムの発足に寄せて—", 情報処理学研報, 2003, 58, pp.1-8 (2003)
- 8) 吉岡亮衛: "季語データベースの構築と「俳句投句鑑賞システム」の概要", 情報処理学研報, 2006, 71, pp.25-32 (2006)
- 9) 吉岡亮衛: "季語データベースの構築と俳句の季語の自動判定の試み", 情報処理学研報, 2000, 100, pp.57-64 (2000)
- 10) 松岡正剛: "千夜千冊", 編集工学研究所, <http://www.isis.ne.jp/mnn/senya/senya.html>.
- 11) N. Tosa and S. Matsuoka: "i.plot ACM SIGGRAPH Emerging Technologies", Proc. The 32nd International Conference on Computer Graphics and Interactive Techniques, p.154 (2005)
- 12) NICT図書街プロジェクト, <http://kc.nict.go.jp/project3/index.html>
- 13) T. Nakagawa, T. Kudo and Y. Matsumoto: "Support Vector Machineを用いた形態素解析と修正学習法の提案", 情報処理学論誌, 44, 5, pp.1354-1367 (May 2003)
- 14) T. Ishida: "Language Grid: An Infrastructure for Intercultural Collaboration", IEEE/IPSJ Symposium on Applications and the Internet, pp.96-100 (2006)
- 15) <http://www.siggraph.org/>



**土佐 尚子** 武蔵野美術大学講師, ATR, MIT Center for Advanced Visual Studiesフェローを経て, 京都大学学術情報メディアセンター教授. ニューヨーク近代美術館・メトロポリタン美術館等に招待展示. 作品は国立国際美術館等6カ所の美術館に収蔵. IEEEマルチメディア'96最優秀論文賞. メディアアーティスト. 工学博士.



**尾原 秀登** 2004年, 京都大学工学部情報学科卒業. 2006年, 同大学大学院情報学研究科知能情報学専攻修士課程修了. 現在, 同大学大学院博士課程在学中. インタラクティブメディア, インタラクティブアートに関する研究に従事.



**美濃 導彦** 1978年, 京都大学工学部情報工学科卒業. 1983年, 同大学大学院博士課程修了. 同年, 工学部助手. 1989年, 京都大学工学部附属高度情報開発実験施設助教授. 1995年, 同教授. 1997年, 京都大学総合情報メディアセンター教授. 2002年, 京都大学学術情報メディアセンター教授. 2006年, 同センター長. 画像処理, 人工知能, 知的コミュニケーション関係の研究に従事. 工学博士.



**松岡 正剛** 早稲田大学卒業後, 東京大学客員教授, 帝塚山学院大学教授を経て, 編集工学研究所所長, インス編集学校校長. 情報文化と情報技術をつなぐ研究開発に従事. 日本文化研究の第一人者. 主な著書に, 『日本という方法』(NHKブックス), 『空海の夢』, 『17歳のための世界と日本の見方』(春秋社), 『松岡正剛千夜千冊』(求龍堂)ほか多数.