

2012/7/19

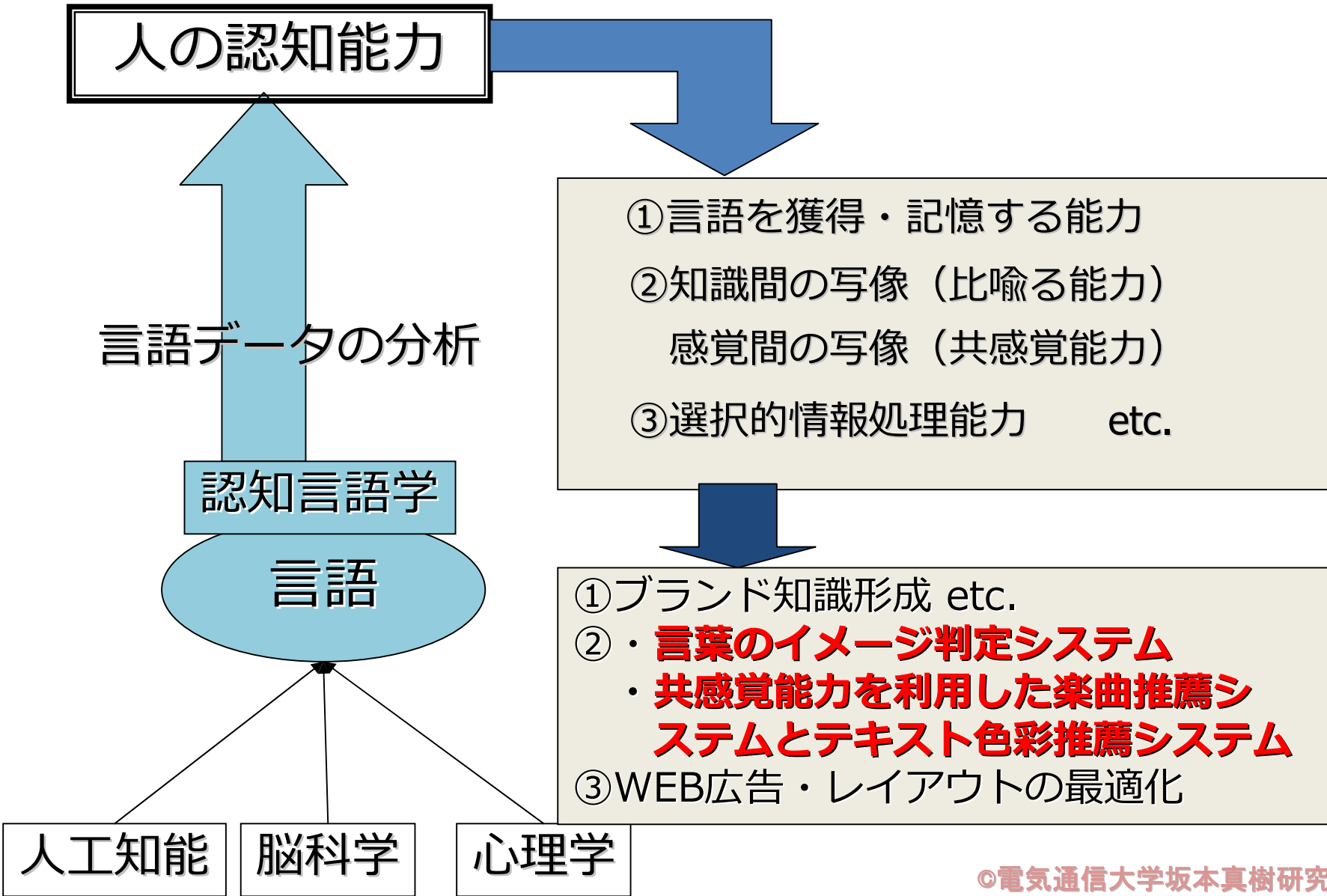


人が言葉からイメージすることを 計算してできるおもしろシステム

国立大学法人 電気通信大学
大学院 情報理工学研究科 総合情報学専攻
准教授 坂本 真樹

1

坂本研究室の特色





研究事例 1 共感覚能力を利用した 言葉のイメージ判定システム

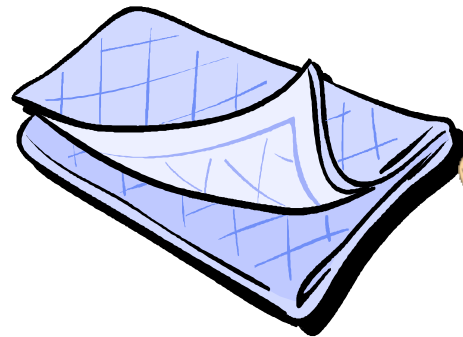
**言葉の音から五感イメージが喚起される人の特性を
利用した言葉のイメージの判定システムです！**

3 入力された言葉が人に喚起するイメージを、
言葉を構成する個々の音韻から定量的に推定します

研究背景

人は様々な質感・経験をオノマトペ（擬音語・擬態語）で表す

手触りの質



さらさら

ざらざら

痛みの質



頭が
ズキズキ痛い

喉が
ヒリヒリ痛い

人の感性と直接的に結びつく言葉に含まれる情報を客観的に定量評価するシステムを作成

研究背景

オノマトペは人の五感にうったえかける感性的な言語表現

⇒ 文芸作品、商品名、商品の宣伝広告に不可欠

例) 「崖の上のポニョ」「きらり」

例) 「ソフラン」「フレア」

「ぷっちんプリン」

例) 「ふわふわ」「ふんわり」

高橋しん「きみのカケラ」(小学館) より



表したいイメージに適合した
新奇性のある表現の創作支援も可能

オノマトペの音象徴性

○ 音象徴的意味

- 一般的に言語のもつ音と意味の関係は必然的なものではない：任意的・慣習的・恣意的（言語の恣意性）
- しかしオノマトペ表現で、音と意味との間になんらかの関連が見られる場合がある（音象徴）
- オノマトペのもつ音象徴的意味（Hamano, 1986）
 - 日本語オノマトペの音象徴性を体系的にまとめた
 - オノマトペにおける特定の音または音の組み合わせは、語中の箇所によって特有の意味をもつ



オノマトペを構成する音から、基本的な意味が予測可能

オノマトペの音象徴性

オノマトペの音と意味の結びつき (Sound Symbolism) 言語学・心理学における学術的知見 (Hamano, 1986)

音韻	意味
母音	1モーラ目の語基
/i/	線, 一直線に延びたもの, 光 (光線)
/a/	平らさ, 広がり, 大きい表面, 派手さ
子音	
/p/	ぴんと張ったもの, 水しぶき, 表面, 突然性, 力強さ
/h/	やわらかさ, 不確定, たよりなさ, 弱さ, 繊細な優雅さ
/k/	金属のような硬い表面との接触
/t/	表面の張りが無い状態, 打撃 (木材, 床, 地面)
/y/	ゆったりした動き, あてにならない動き

2012/7/19

©電気通信大学坂本真樹研究室

(一部を抜粋)

オノマトペの形態と意味

- オノマトペの形態は多様にみえるが、そのほとんどが数音からなる基本形の組み合わせで構成されている
- そこに「り」、促音「っ」、撥音「ん」、母音長音化「ー」、反復いずれかが加わることがある：それぞれ固有の意味

CV	
CV + CV	
CV + CV'	
⋮	
(C: 子音, V: 母音)	

「り」	ゆったりした動き, 動作の完了
促音 (語末)	スピード感, 急に終わる様子
促音 (語中)	強調
撥音 (語末)	共鳴 (擬音的なニュアンス)
母音の長音化	長い音, 強調
反復	音や動作の継続・繰り返し

意味

オノマトペの音と意味の結びつきの定量化

評価尺度	音韻							
	カ行	サ行	タ行	ナ行	ハ行	マ行	ラ行	ワ行
きれいなーきたない	0	0.34	0.16	0.08	-0.02	0.15	1.56	-0.64
滑らかなーざらざらした	0	-0.13	0.27	0.76	0.19	0.56	0.81	0.17
明るいー暗い	0	0.05	0.09	-0.21	-0.12	0.14	1.21	0.35
きめの細かいー粗い	0	0.23	0.10	0.19	-0.01	-0.03	0.27	-0.94
澄んだー濁った	0	0.17	0.00	-0.45	-0.09	-0.48	0.77	-1.10
快いー不快な	0	0.76	0.52	0.80	0.31	0.70	1.66	0.38
かたいーやわらかい	0	-0.82	-0.68	-2.59	-1.26	-1.83	-1.26	-1.90
とげとげしいー丸みのある	0	-0.30	-0.68	-2.10	-0.78	-1.75	-1.40	-1.51
	⋮							

(藤沢ら2006より一部を抜粋)

システムのアルゴリズム

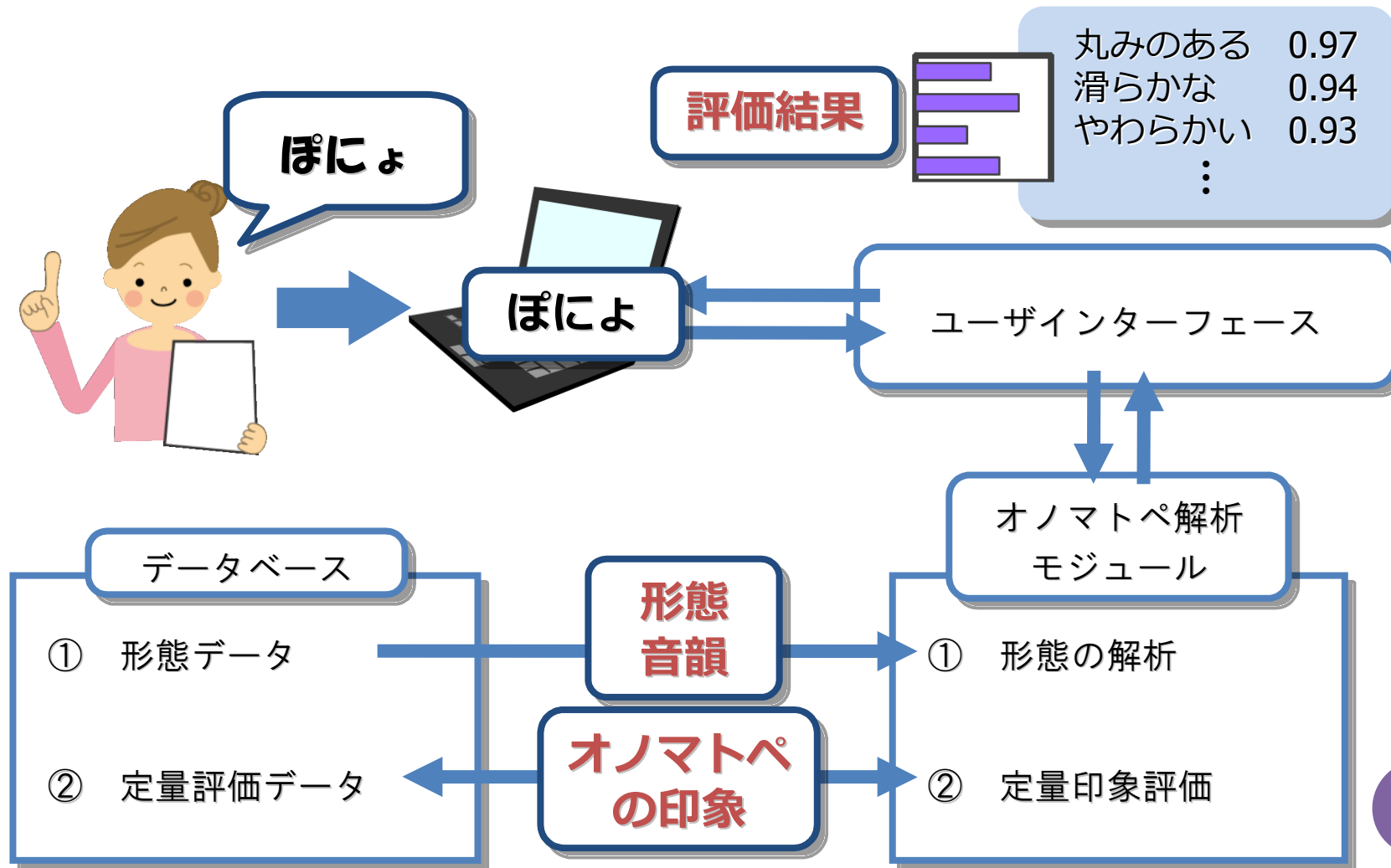
○ 本システムへの適用

- 藤沢らの印象評価実験では、2モーラの擬音語を対象
 - よって実験および分析より得られた音韻特性のカテゴリ数量は、2モーラオノマトペ表現の印象評価値に対応
- 本システムでは任意の長さのオノマトペ表現を扱いたい
 - システム内で入力されたオノマトペ表現の音節数（モーラ数）をカウント
 - 評価値の重みが2モーラ擬音語表現の場合と等しくなるよう藤沢らのモデルを正規化
 - 下式のように印象評価値を補正：

$$\hat{Y} = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + Const.}{Mora} \times 2$$

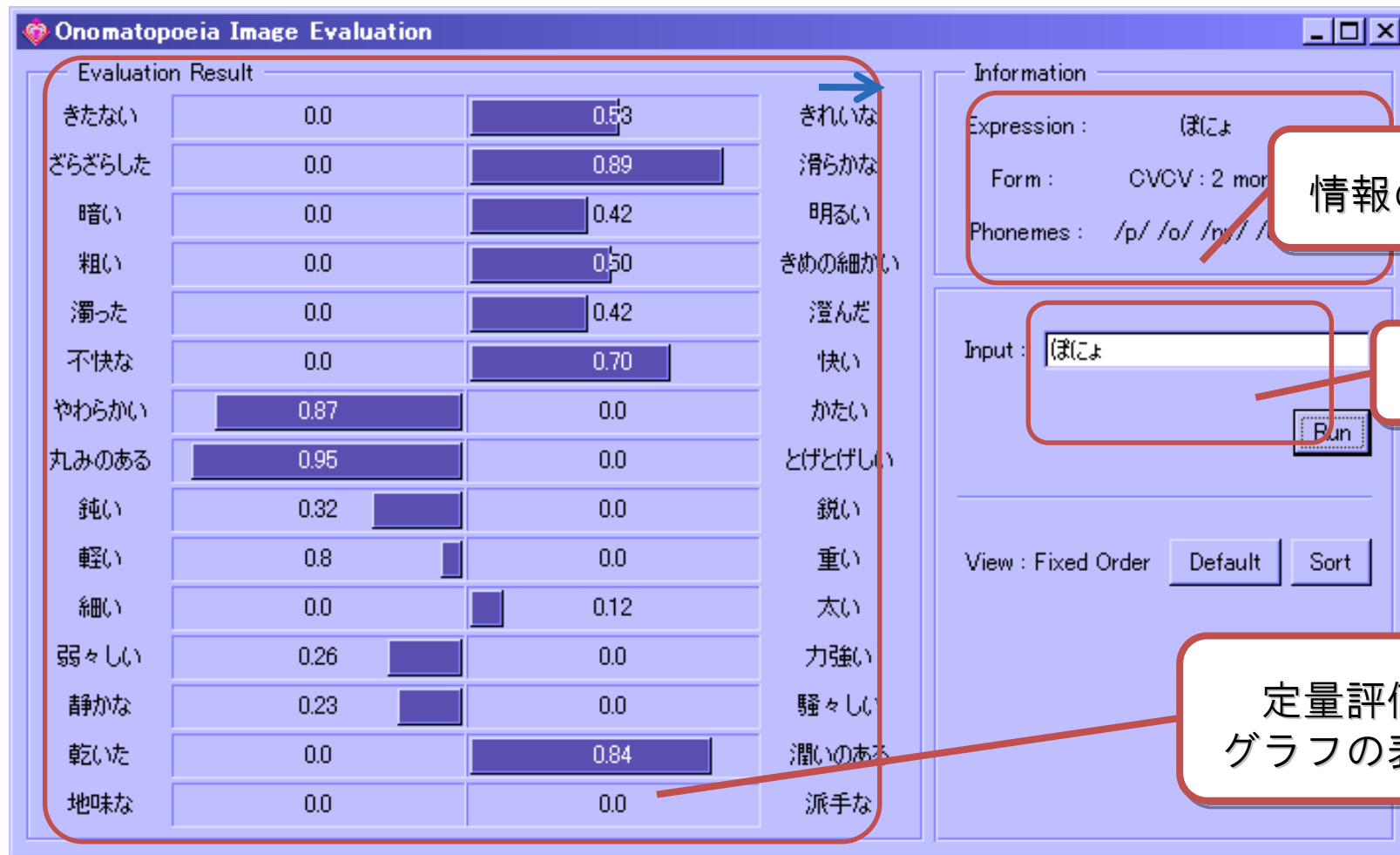
- この補正されたモデルに基づいて、印象評価処理を行う
 - 15組の形容詞評価尺度対に対応する15個の印象評価値を得る

システムの構成



システムのインタフェース

坂本・清水：特願2009-102796



2012/7/19

©電気通信大学坂本真樹研究室

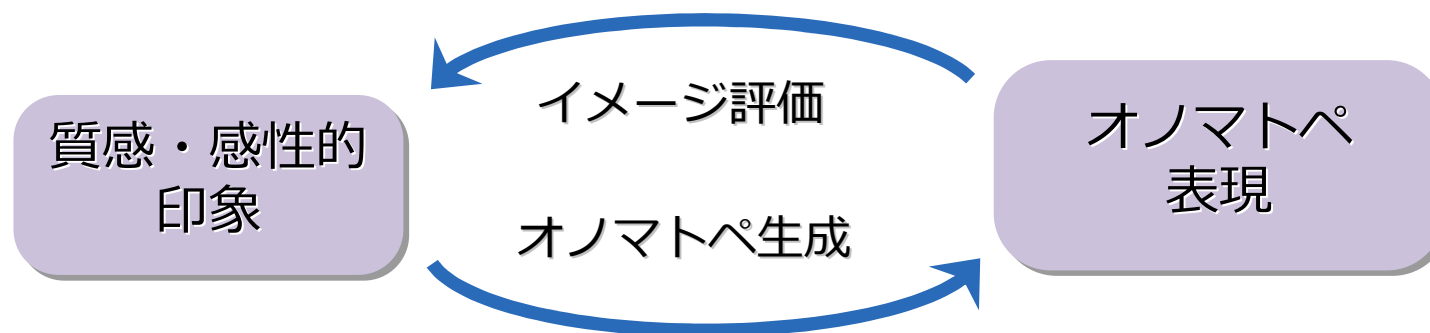
言葉を生成するシステムも作りました

オノマトペ自動生成システム

出願番号：特願2011-168539(出願日2011年8月1日)

出願人：国立大学法人電気通信大学、発明者：坂本真樹、清水祐一郎

- 坂本・清水（2009）の印象評価手法を応用
- ユーザの印象評価を入力として、オノマトペ表現を生成するオノマトペ生成システムを開発



オノマトペ生成システムのインタフェース

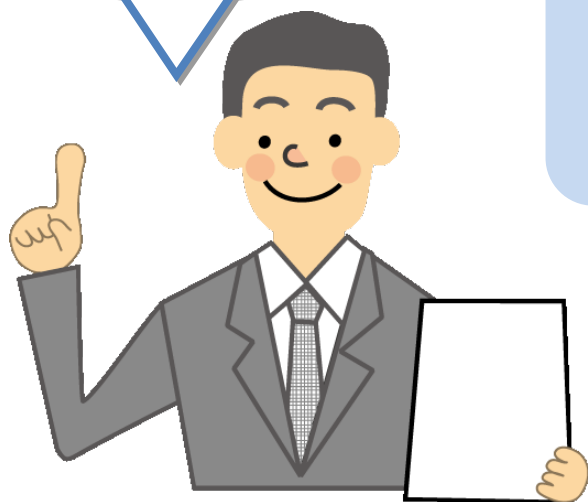
The screenshot shows the 'Onomatopoeia System' interface. It has two main sections: 'Evaluation Input' and 'Result'. The 'Evaluation Input' section contains 18 sliders, each with a label on the left and right. A red box highlights the 'きたない' (dirty) slider. The 'Result' section contains a table of generated onomatopoeia words and their scores. A red box highlights the table. Three callout boxes with red arrows point to the interface: 'システムの切り替え' (System Switch) points to the top right, '評価値の入力' (Evaluation Value Input) points to the sliders, and '生成結果の表示' (Display of Generation Results) points to the result table.

Onomatopoeia	Score
びよんこよ	0.471
みよー	0.462294
もりよー	0.462291
りよーも	0.462291
みよんみよ	0.462148
みゆるー	0.454
りよーむり	0.454
こよんみよ	0.448
みよんこよ	0.448
みよーろ	0.447676
みよろん	0.447676
みよんろ	0.447676
もんりよ	0.447676
りよー	0.447
ろーみよ	0.447
みよろっ	0.444
りよっも	0.444
びよーこよ	0.443383
びよこよー	0.443383
こよりよー	0.443230

想定される用途①命名システムとして

子供の名前、会社名、商品の特徴が効果的に伝わり、新奇性のある商品名などの創作

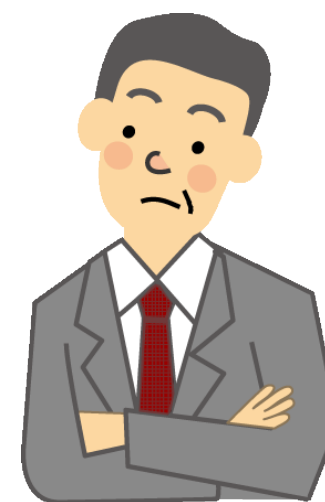
新ブランド
シャララで
いきましょう！



軽い： 0.86
爽やかな： 0.71
...



もっと高級感重視の
名前がいいかな...



15

想定される用途②

感性・質感評価、素材提案システムとして

質感や感性が重要な商品の説明・宣伝広告

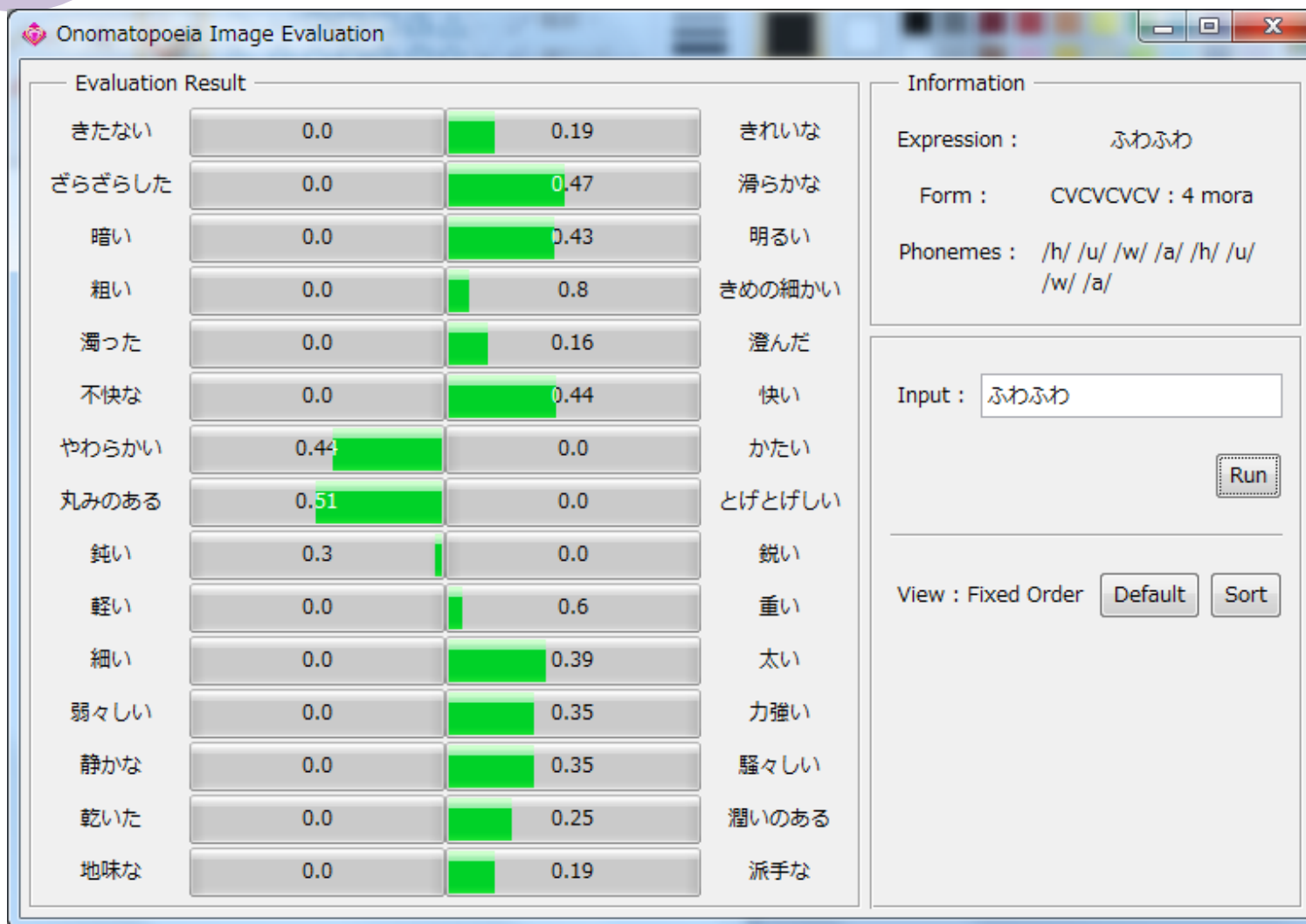
「この商品はさらりとしながらふわふわした素材です」

やわらかい： 0.92
軽い： 0.82

ユーザの所望する素材の提案可能



質感を表すオノマトペ「ふわふわ」



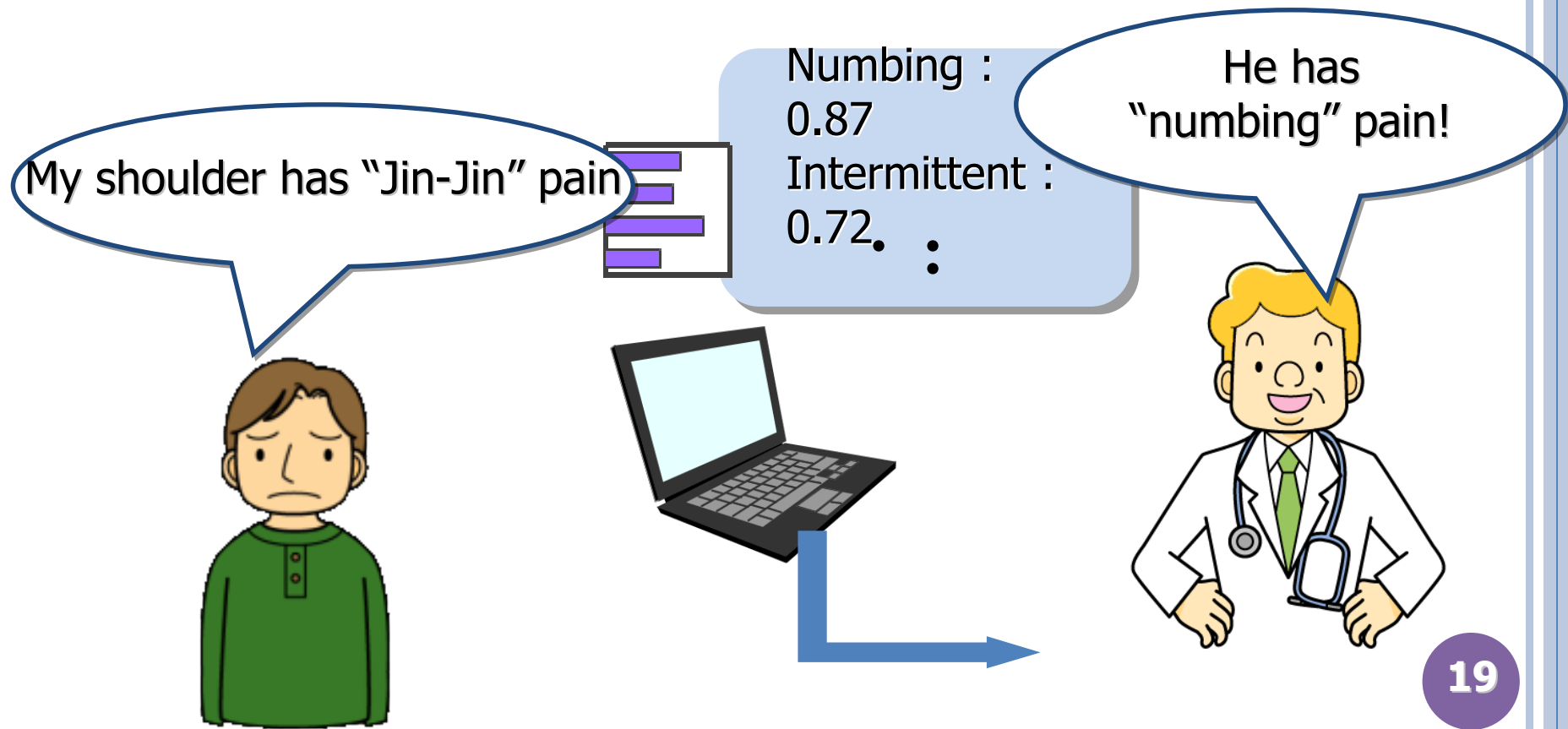
質感を表すオノマトペ「もふもふ」



想定される用途③

海外での問診支援システムとして

- 「頭がずきずき痛い」という患者の痛みの質を定量化、評価尺度表示を多言語化し、国内外での問診を支援



2012/7/19

©電気通信大学坂本真樹研究室



研究事例2 テキスト情報に適合した色彩を提案するシステム

**テキスト情報からイメージされる色彩を
自動で選定・提案するシステムです！**

単語と色彩の認知的連想関係を利用
テキスト内容との結びつきが強い色彩を選定
テキスト情報の自動彩色技術として幅広い応用が可能
(電子書籍・ブログ・プレゼンテーションツール)

従来の選定手法と本研究の位置づけ

- ◆ テキストと色彩の結びつきに着目した先行研究
→ 想定されるテキスト色彩選定手法

手法1：外的色彩情報の利用

- 同じページ内にある画像など、テキスト以外の色彩情報を利用
(村山ら, 2008)

手法2：色彩語・感性語の利用

- テキストの色彩語や感性語に対応した色彩情報を利用
(服部ら・2008, 中山ら・2004)

- ◆ システムの特徴：テキストと色彩の結びつき
⇒ 単語から想起される色彩情報

従来手法との比較

- ✓ テキストと色彩の結びつき：弱(手法1)⇒強
- ✓ 扱えるテキストの種類：少(手法2)⇒多
- ✓ 提案色彩バリエーション：少(手法2)⇒多

本研究の基盤となるシステム

- 歌詞と色彩の相関に着目した楽曲推薦システム
(坂本・仲村・内海, 2010),(仲村・川西・坂本, 2011)
- ✓ ユーザが入力した色彩に適した楽曲を検索
— 基本原理：楽曲毎に『楽曲に適した色彩情報』を与える

【楽曲の色彩選定手法】

1. 楽曲の歌詞を解析
2. 『単語から想起される色彩情報』を基に歌詞全体の色彩
を算出

本研究：“歌詞” = “テキスト”

『単語から想起される色彩情報』を利用した

テキストの色彩提案システム

システムの概要

システムの特徴

- ✓ 入力テキスト情報のイメージに合った色彩提案
- ✓ 単語から想起される色彩情報を基に色彩を選定
- ✓ 選定色彩イメージを円グラフ・棒グラフで視覚的に提示

◆ 当初の使用色彩：35色

➤ Microsoft Word2003の標準カラーパレットから採用



現在はより妥当な基準で
45色に更新



システムの基本原理

- ◆ 『単語の色彩ベクトル』：単語から想起される色彩情報
- ✓ 35色の各色彩が想起される確率を値とするベクトル

例) 単語A：『赤が想起される』と回答した被験者20名,
全被験者120名の場合, $20 \div 120 = 0.17 \rightarrow$ 赤の想起確率

色彩ベクトル = (黒の想起確率, 茶の想起確率, ..., 赤の想起確率,
..., 青の想起確率, ..., 緑の想起確率, ..., 白の想起確率)

- ◆ 『単語の影響度』：色彩との結びつきの強さ
- ✓ ある単語から色彩が想起されやすいと回答した被験者の割合

例) “雨”を含むテキストで“雨”を回答した被験者10名,
全被験者20名の場合, $10 \div 20 = 0.5 \rightarrow$ “雨”の影響度 (= 回答率)

- ◆ テキストの色彩ベクトルの推定

テキスト内の単語

⇒ 色彩ベクトル × 影響度 × 頻度 の重心ベクトル

実験の回答用紙サンプル

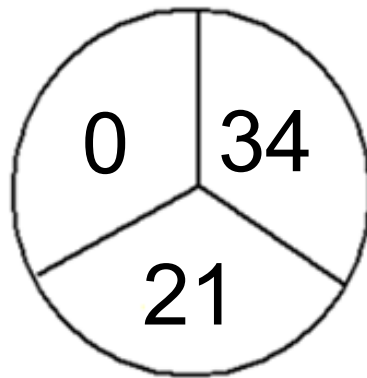
○ 1人30文章 — テキスト提示順は4パターン設定

問 1: 次の記事を読んで、記事から連想される色彩を〈別紙 1〉のサンプルカラーから最低 1色・最大 3色選んで、【色彩回答】の円形セルに番号を記入してください。

記事

午前5時過ぎから餅つきが始まり、白いタオルを頭に巻いた修行僧3人が、タイミングを合わせながら次々ときねを振り下ろした。この日ついた餅は餅米約33升分。直径約40センチにもなる3段重ねの鏡餅は、30日に本堂などに飾られる。1月7日に鏡開きをして、8日に開かれる黄檗ほていまつりでぜんざいとして振る舞う予定だという。

【色彩回答】

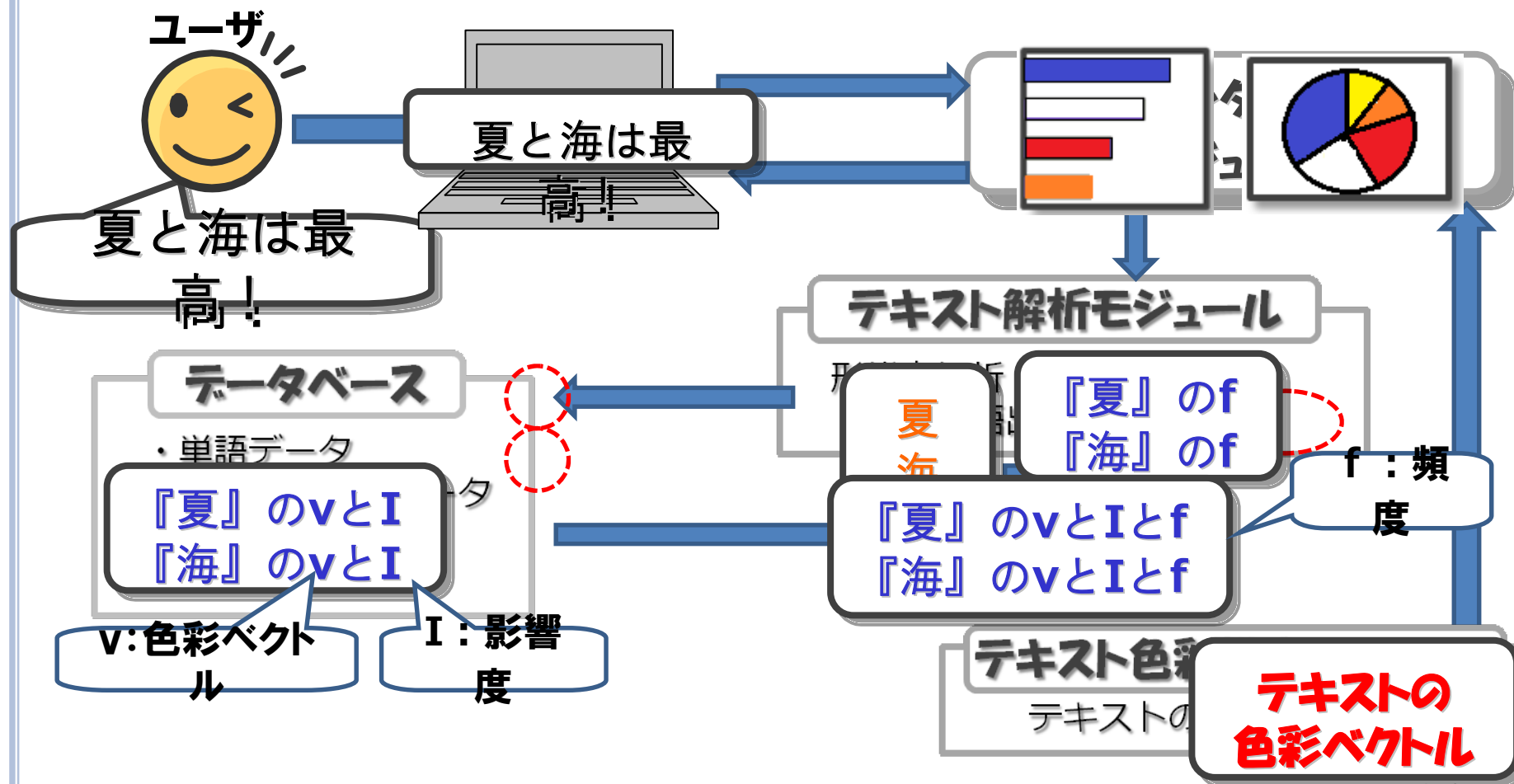


問 2: 問 1で色彩を選択した際、色彩の連想に影響を与えた単語全てに○を付けてください。

◎電気通信大学坂本真樹研究室

システムの構成

- ユーザーインターフェースモジュール・テキスト解析モジュール・テキスト色彩推定モジュール・データベースの4つで構成



システムの実装

▶ 試作システム(Javaを使用)

▼テキストを入力してください

カニなら岩美町の「たつみ・親がに御膳（ごぜん）」に感服させられた。カニ膳の大定番である松葉ではなく親ガニ、それも7枚半！使うのだ。もうこの時点でカニ膳の政権交代だが、食した知人女性がこう熱弁した。「内子、外子、身と一緒に食べたら、も〜幸せ〜」。カニに仕分けは無用！

サンプルテキスト
▼以下の中から選択 ▼

追加したテキスト
▼以下の中から選択 ▼

テキスト入力

テキストを追加 **判定** リセット

イメージカラー：

推定色彩のグラフ表示

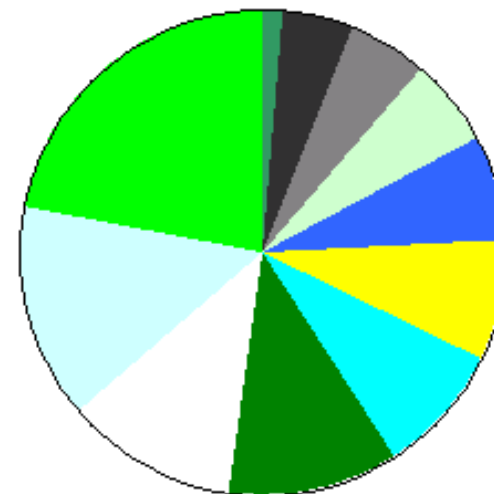
システムの提案色彩例（1）

- 「エコポイント・エコカー」に関するテキスト (217字)

プリミティブワード：

— **エコ**, **省エネ**, **環境**, **ポイント**, **車** の5語

未知語：60語

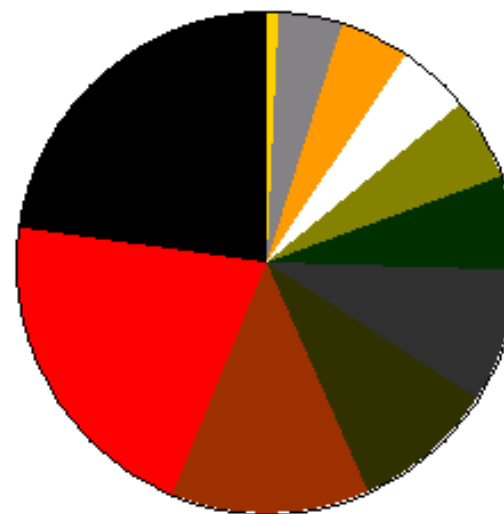


- ◆ 「イラク情勢・核開発」に関するテキスト (174字)

プリミティブワード：

— **イラン**, **イラク**, **破壊**, **核**, **戦争**,
自由, **攻撃**, **野心**, **夜** の9語

未知語：36語

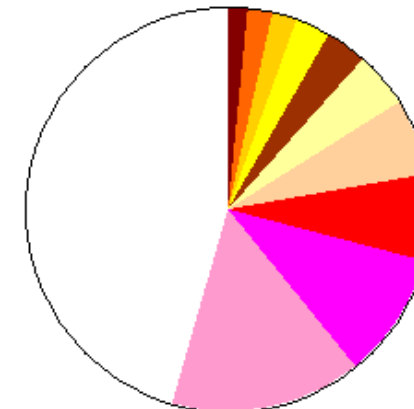
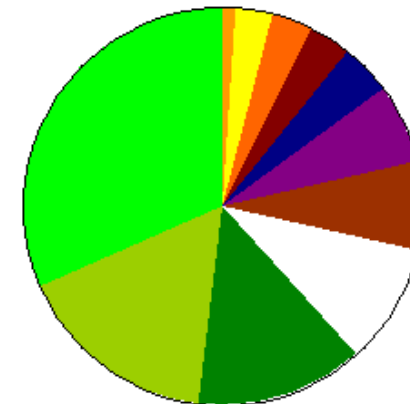
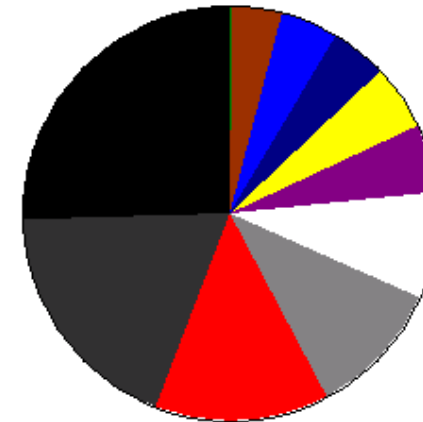


システムの提案色彩例（2）

- 「ネットいじめ」に関するテキスト(77字)
プリミティブワード：いじめ, 嫌がらせ, 悪口,
ネット, パソコン, インターネットの6語
未知語：14語

「キャベツ・カレー」に関するテキスト(191字)
プリミティブワード：キャベツ, カレー,
ブルーベリー, 春, みどり, 野菜の6語
未知語：28語

- ◆ 「芸能人の結婚報道」に関するテキスト(147字)
プリミティブワード：ウェディングドレス・結婚・
可愛い, メロメロの4語
未知語：23語





ご清聴ありがとうございます

ございました！

システムのデモを
楽しんでくださいね♪

