

相槌に個性を持たせたテキスト対話システム

上野 洋^{1,a)} 井上 雅史^{1,b)}

概要: 人間同士が対話を楽しむ際には、相手の個性が重視される。対話システムが話しかけたいと思わせる魅力的な個性を持つことは、対話システムの利用可能性を高めうる。しかし、テキスト対話システムにおいては、語彙や文体といった発話の表層表現にみられる個性をコーパスベースで実現する手法は未発達である。そこで本研究では、テキスト対話システムにおいて、コーパス分析結果を反映した個性を実装した。人間同士の対話の円滑さを成り立たせる要素の一つが相槌であり、対話システムが、ユーザの発話に対して自然な相槌を返すことは、自然な対話を実現する上でも重要と考えられる。システムに聞き役として相槌を生成させる対話実験を実施し、個性の観点からシステムの印象評価を行った。

1. 序論

人間と機械が自然言語を通してコミュニケーションを取り合う対話システムの研究が進み、ユーザと対話を交わしながら指示に沿ってメールや検索などを行う秘書機能アプリケーションのような実用化も始まっている。このようなアプリケーションの動作は、ユーザが一つの発話を行うたびに一つの返答を返す形式が主流であり、人間同士の自然な対話とは異なっている。リアルタイムで円滑なコミュニケーションを実現する上で考慮すべき要素の一つである相槌が、適切に用いられていないことも、人間同士の対話と異なる点である。ユーザの発話に対してシステムが聞き役として自然な相槌を返すことができれば、自然な対話の実現に、一歩近づくことができると考えられる。

また、今後の対話システムに求められる機能として、話しかけたいと思わせる魅力的なキャラクターの重要性が挙げられている [1]。様々なキャラクターを生成するために、システムの発話の仕方に個性として差を付ける研究が行われている。発話音声の変化で個性を表現するアプローチと比較して、語彙や文体と言った発話の言語表現を利用する研究は多くなされていない [2]。また対話システムへの個性付与の研究は、人手やルールベースによる手法が多く、コーパスベース等の統計的な手法は少ない。

対話システムへの相槌実装・対話システムへの個性付与のいずれにも期待されることに、システムとの対話を飽きさせず継続的に対話を行うよう促すことがある。対話システムとの会話中に、利用者が何を話せばいいのか困る場面

が見られることも少なくなく、個性付与を相槌によって行うことは対話システムを話しやすくする上で重要なアプローチとなり得る。そこで本研究では聞き役として相槌を生成する対話システムを構築し、相槌の仕方において個性を実装する。言語表現で個性を明確にするために、音声対話システムでなくテキスト対話システムを用い、個性はコーパスの統計量を基に作成する。

2. 研究背景

2.1 相槌表現

相槌とは会話における聞き手の主要な行為の一つであり、対話の円滑な進行に大きな役割を果たす。例えば以下の会話における B の「はい」「会議室」が相槌となる。A の説明が一通り終了するまで B が発言する必要はないが、A の話の合間に B がこれらの相槌を打っている。これによって、B が話を聞こうとしていることや、B は会議の場所が「会議室」だということを確かに理解していることを A が知ることができ、A は安心して説明を先に進めることができる。

A : 明日の会議は 10 時からで、

B : はい

A : 場所は 3 階の会議室

B : 会議室

A : で実施します。

本研究では、堀口が相槌の機能・使用者・出現位置・表現形式についてまとめたものに [3]、相槌表現に関する項目を加えて、以下のように相槌を定義する。

- 相槌の機能は、聞いているということを伝える、わかったということを伝える、話の進行を助けるなどで

¹ 山形大学大学院理工学研究科

^{a)} tmk56575@st.yamagata-u.ac.jp

^{b)} mi@yz.yamagata-u.ac.jp

ある (機能)

- 相槌を行使するのは、聞き手である (使用者)
- 相槌が出現する位置は、話し手の発話権の中である (出現位置)
- 相槌は表 1 にまとめられた分類の中で、その他 (O) を除く 6 種類のいずれかに属する (表現形式)

表現形式の分類は、高梨らによるもの [4] を参考にし、表 1 のようにまとめた。以後表 1 の各分類をそれぞれ B 分類・E 分類・L 分類・R 分類・C 分類・A 分類・O 分類と呼ぶこととする。ただし O 分類の相槌は本研究で使用しない。

出現位置については、メイナードによると日本の日常会話において相槌が登場しやすいタイミングは文末のポーズ付近・終助詞や間投助詞の付近・話し手の頭の動き付近である [5] とされ、テキスト上で相槌の生成タイミングを検討する場合には品詞情報が重要であると言える。

2.2 個性

サピアによると、言葉が完全な文として形成されるまでの過程は、音声・音声学的要素・発音・語彙・文体に分けられ、それぞれが話者の特性によって組み合わされて、言葉の個性が生まれるとしている [6]。サピアは語彙は極めて重要だとしている。例えば、「堅苦しい単語ばかり使うからきっと学者の人だ」といった印象がこれにあたる。本研究では語彙・文体のレベルに焦点を当てて個性をコーパスから取り出し、テキスト上で個性を生成することを目指す。

3. 関連研究

相槌システムに関する関連研究として、神谷らの相槌位置の検出の試みがある [7]。音声と書き起こしテキストから相槌の位置を探してアノテーションを行い、相槌コーパスを作成し、直近の形態素の品詞や位置・発話速度・ピッチ等の素性からサポートベクターマシンを用いて相槌位置の検出器を学習した。この研究では本研究と異なり個性は考慮せず、生成される相槌は「はい」のみとしている。

個性付与に関する関連研究として、発話をすべて人手で作成することによって個人的な発話を実現した Nass らの実験 [8] や、心理学や言語学の知見を活かしてルールベースで効率的な個性の出し方を規定し、使用するルールやパラメータを調整することで個性を実現した Mairesse らの発話生成器 [9]、個性を生成しない通常の対話システムで生成された発話に対し、N-gram 分布を利用し統計的な手法で一部を加工する水上らの手法 [2] などがある。

相槌への個性付与に関する関連研究として、目黒らの研究では、聞き役としての対話を対話行為の分類毎に分析し、隠れマルコフモデルによって聞き役対話の個性を明らかにしようとした [10]。聞き役対話システムの開放性が高いよ

うに見せるには、自分の発話ターンをできるだけ長く持とうとし、逆に開放性が低いように見せるには頻繁に相手に発話ターンを譲るようにするとよい、とされている。また、相槌とよく似た機能を持つ「うなずき」の個性をシステムで再現しようとした試みとして濱元らの研究では日本人とアメリカ人の相槌の頻度を比較し (日本人の方が 2 倍以上高頻度) その頻度差を反映させたうなずきシステムを実装し、うなずきが多いほどユーザの話しやすさが増すことを実験で確かめた [11]。

4. 相槌生起状況調査

効率的に相槌で個性を表現するシステムを構築するためには、実際の対話において相槌の使い方にどのように個人差があるのかを把握する必要がある。対話を書き起こしたコーパスに対してアノテーションを行い、相槌の生起状況を調査した。

4.1 対象コーパス

特に相槌の個性について、特定地域の住民の相槌使用状況を詳細に調査した結果、個人差や男女差・年齢差があることが明らかにされている [12]。本研究でも話者の個人差や男女差・年齢差の情報を含む対話コーパスから、相槌の個性を抽出する。調査に使用するコーパスとして、名大会話コーパス*1 を使用した。このコーパスの書き起こしテキストの例を図 1 に示す。「うん」「わー」といった括弧で囲まれている部分が聞き手の相槌である。相槌がラベル付けされているため、対話システムで使用する相槌候補リスト (相槌コーパス) を作成することに適している。また、性別・年齢・出身地・参加した対話一覧といった参加者の情報が細かく記されており、個性を実現するために参加者を属性で分けることが可能である。ただし図 1 にあるように、相槌を発した話者が書き起こしテキストに明記されていない。そこで、話者を特定するために、参加者が 2 名の対話データのみを使用した。本研究で使用した部分の基本データを以下に記す。

データ形式 : 対話の書き起こしテキスト

収録期間 : 2001/10/23~2003/2/17

参加者数 : 138 名

(10 代前半~90 代前半の男性 20 名, 女性 118 名)

タスク : 話者 2 名による 30 分~60 分程度の自由雑談

対話数 : 96 回

合計時間 : 71 時間 50 分

発話数 : 93869 (コーパスの行数)

相槌数 : 28517 (括弧で区切られた領域の数)

*1 <https://dbms.ninjal.ac.jp/nuc/index.php?mode=viewnuc>

表 1 相槌表現の分類と表現例 [4]

相槌の分類	分類の説明	相槌の表現例
応答系感動詞 (B)	情報の受容や承認を示す感動詞	はい, うん, ええ
感情表出系感動詞 (E)	驚きや気づきなど, 心的状態の変化を示す感動詞	え, ふーん, へえ
語彙的応答 (L)	同意を示す慣用的表現による反応	なるほど, そうそう, だよ
繰返し (R)	ほかの話者の先行発話の繰返し	あのツアーで (ツアーで) 行かせてもらった。
補完 (C)	他者の発話の要素を予測し補う発話	靴にわらを, わらの縄を (ゆわいつけて) 結びつけて,
評価応答 (A)	他者の発話への評価的語彙を用いた反応	おもしろいな, こわ
その他 (O)	上記に分類できない相槌, 非相槌等	本当?, ***, <笑い>

F024: そうか, こたつか。
Aさん, これ, 置いてっちゃったよ。
<笑い>これは単4なんだっけ?
(うん) コーヒー入れよう。
F140: さっきから, なんか, ブラックもう何杯も飲んでる。
<笑い>
F024: <笑い>胃に悪いよ。
F140: でも, あんま濃くしてない。
わー, お湯がないかもよー。
(わー) それか, 先いーれよ。

図 1 名大会話コーパスの書き起こしテキスト例

4.2 話者毎の偏り調査

28517 個の相槌に対し, 表 1 に従って相槌を分類する簡易的なアノテーション作業を行った。それぞれの相槌に対して以下の 6 つのタグを付与した。

Aizuchi: 表 1 の中から当てはまる分類を一つ選択

Person: 話者毎に割り当てられた識別コード

(男性は M+数字 3 桁, 女性は F+数字 3 桁)

File: 対話の書き起こしファイル名 (data+数字 3 桁)

Time: 相槌が発せられた対話の収録時間 (分単位)

Age: 発話者の年代 (その年代の前半か後半も記される)

Region: 発話者の出身都道府県

このデータに対して, 相槌対話システムとして個性を発揮できそうな要素がないか分析した。

4.2.1 調査 1: 相槌の使用頻度

コーパス中における各個人の相槌の使用頻度を, 相槌分類毎にまとめた。個人別に相槌頻度を分類毎にまとめた図を図 2 に示す。最も相槌頻度が多い話者は 1 分あたり 14 回, 最も少ない話者は 1 分あたり 0.2 回であり, 最大で約 70 倍の頻度差が存在する。濱元らの実験では, うなずきの頻度に 2 倍以上の差が付いている 2 種類の個性を比較した所, ユーザの反応に有意な差が生じていた [11] ことから相槌の頻度に大きな差がある場合, 対話システム上で実装することでユーザが個性の差を感じる事が期待できる。

4.2.2 調査 2: 相槌分類毎の使用割合

分類毎の使用割合は, 全体的には B 分類の相槌が大半を占め, E 分類・L 分類と続き, その他の分類は 1%~2% の利用率に過ぎない。図 2 より, 使用割合についても個人別では大きな差が存在することがわかる。B 分類の相槌を最も多用する話者が大半を占めているが, E 分類の相槌の使

表 2 5 名以下の話者が複数回使用した名詞, 動詞, 形容詞を除く単語

あらー	かなり	あの一	なきや	え
超	いえいえ	まで	実は	
っつか	よし	ありがとう	えーっ	
べ	じ	なあ	よく	
ふうん	かしら	ある	いやいや	

用率が 6 割を上回っている話者や C 分類の使用率が 1 割を超える話者も見られる。相槌は分類毎に表現が大きく異なるため, このような分類毎の使用割合が異なる話者を取り上げて比較すれば個性の差を実感できると考えられる。

4.2.3 調査 3: 相槌に使用する単語

サピアが言語の個性において語彙が重要だと述べた [6] ように, 相槌においても使用する語彙の違いによって個性が表現できると考えられる。そこで使用する話者が少ない, 特徴的な単語を抽出することで個性が表現できないか調査した。ただし, 文脈が変化すると生起率が大きく変化する単語の多い名詞・動詞・形容詞といった品詞の単語や, 日常生活で使用する機会が著しく少ない単語は, 個人を特徴付けることのできる単語でなかったとしてもコーパス中で殆ど登場しないことが考えられる。そこで「名詞・動詞・形容詞を除く」「使用回数が 1 度のみの単語は除く」という条件下で使用話者数が 5 名以下の単語をまとめたものを表 2 に示す。若者の言葉遣いをイメージさせる「超」「っつか」, 女性の言葉遣いをイメージさせる「あらー」「かしら」など, 個性的な相槌を生成する上で効果が期待できる単語が多く見られる。このような単語が見られることから, 使用する話者が少ない単語を含む相槌を優先して抽出することで個性が表現できると思われる。

5. 個性生成手法

調査 1・調査 2・調査 3 を基に設計した提案手法をそれぞれ手法 1・手法 2・手法 3 とし, 以下にその詳細を述べる。

5.1 手法 1: 書き起こしテキストからの相槌生成率学習

調査 1 では相槌の頻度に大きな差があることが分かり, 頻度差を対話システム上で実装することで, ユーザが個性の差を感じる事が期待できると考えられる。そこで元のコーパスにおける書き起こしテキストから, 相槌の生成されやすい位置を学習する手法を提案する。実際の対話から

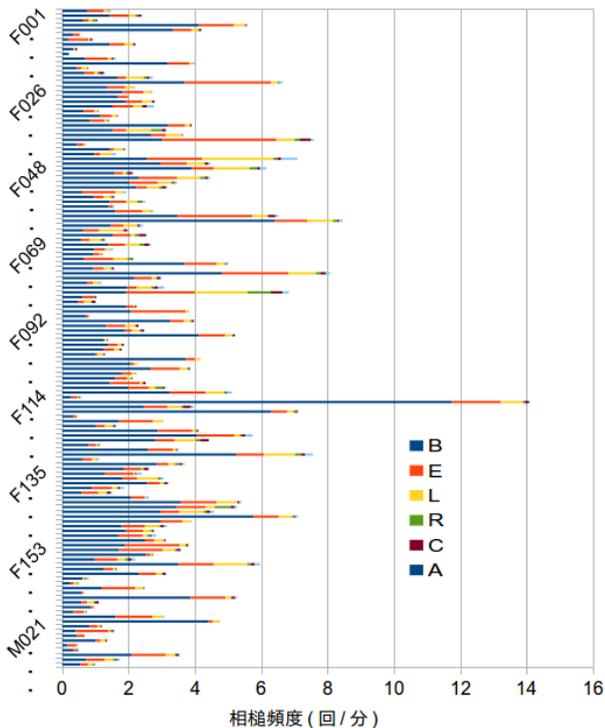


図 2 個人別の相槌分類毎の使用頻度比較 (色分けは表 1 に従う)

相槌生成率の学習を行うことで、自然な相槌生成位置を保ちながらも相槌頻度を反映させることができると考えられる。

2.1 で述べたように相槌の生成タイミングには直前の単語の品詞が重要であることがわかっているため、学習は品詞列ユニグラムを用いて行う。以下の対話例を用いて説明する。(相槌 B: うん) が聞き手の相槌であり、それ以外が話し手の発話である。

A さん、これ、置いてっちゃったよ。

これは単 4 なんだったけ?(相槌 B: うん)

コーヒー入れよう。

この文章を形態素解析器 MeCab^{*2} を用いて品詞列 (記号を除く) に変換し、相槌を分類毎に表記すると以下のようになる。

名詞 名詞 名詞 動詞 助詞 動詞 動詞 助動詞 助詞
名詞 助詞 接頭詞 名詞 助動詞 名詞 助動詞 助詞 相槌 B
名詞 動詞 助動詞

この対話の中に助詞は 4 つあるが、その後ろに相槌が存在する物は 1 つである。この対話から学習を行うと、助詞の後にこの聞き手が B 分類の相槌を打つ確率は 1/4 と学習する。

5.2 手法 2: 分類毎の相槌使用割合の強調

調査 2 では分類毎の相槌使用割合が異なる話者が存在することが分かり、使用割合差を対話システム上で実装する

ことで、ユーザが個性の差を実感できると考えられる。前節の手法 1 で分類毎に相槌を分けて学習を行っているため、分類毎の使用割合も反映できるが、対話システムとの短い対話の中で実感できるほどの比率の違いにはなりにくい。例えば最も A 分類の相槌が多い個人の割合は 0.0976 と全体平均の 0.015 と比べて 6 倍以上と非常に高い。しかし全員に高い値を示している B 分類の割合が 0.634 となっており、この状態で対話を行えば A 分類の相槌の頻度が高いことは目立たない。このような個性を実装するために、分類毎の相槌使用割合の強調を行う手法を提案する。

この手法には偏差を用いる。偏差 d はある数値集合 (要素数 n) に属するある数値 x_i とその数値集合の平均値との差であり、式 (1) で表される。

$$d = x_i - \frac{\sum_{k=1}^n x_k}{n} \quad (1)$$

偏差を乗算的に増加させることで、全体の割合の順序や平均値や合計値を変化させずに分散値のみを増加させ、平均からの差を強調することができる。この乗算的增加の際に用いる係数 α をパラメータとして設定し、偏差が α 倍になるように相槌使用割合を変化させる。ある話者 $i(i \in \text{People})$ における、ある相槌分類 $j(j \in \{B, E, L, R, C, A\})$ の使用割合 p_{ij} を式 (2) によって p'_{ij} に変換する。

$$p'_{ij} = p_{ij} + (\alpha - 1) \left(p_{ij} - \frac{\sum_k^{\text{People}} p_{kj}}{|\text{People}|} \right) \quad (2)$$

変換の過程で生成率が 0 を下回った場合はその値を 0 とみなし、下回った分の生成率は他の分類の生成率として按分する。本研究における評価実験では $\alpha = 3$ とした。 $\alpha = 3$ のときには、前述の話者の A 分類の相槌使用割合 0.0976 (平均 0.015) は 0.2528 に変換される。

5.3 手法 3: 単語 idf 値を利用した相槌選択

調査 3 では使用する話者が少ない単語を含む相槌が存在することを確かめた。使用話者が少ない単語を優先して抽出することで個性が表現できると考えられる。そこで、使用する話者が少ない単語を含む相槌を優先して抽出するために、使用する話者の少なさの指標として単語 idf (Inverse Document Frequency) 値を用いる、相槌選択手法を提案する。

単語 idf 値はドキュメント群 D 中のドキュメントにおいて単語 w が含まれる頻度の逆数で定義されており、総ドキュメント数を $|D|$ ・単語 w が含まれるドキュメント数を D_w で表すと、単語 idf 値は式 (3) によって定義される。

$$\text{idf} = \frac{|D|}{D_w} \quad (1 \leq \text{idf} \leq |D|) \quad (3)$$

素の逆数によって定義すると値の変化が急激になってしまうので、対数を取って使用する場合も多い。その場合は

*2 <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html> 式 (4) によって定義される。

$$\text{idf} = \log \left(\frac{|D|}{D_w} \right) + 1 \quad (1 \leq \text{idf} \leq \log(|D|)) \quad (4)$$

本研究では個性を強く表現するために式 (3) の定義を用いる。そして、コーパスの中から確率的に相槌を選択する処理において、最も単語 idf 値が高い単語によって確率に重み付けを行った上で相槌選択を行う。例えば「うん」は元コーパス中の 138 名全員が使用している単語であり、単語 idf 値は 1.00 である。対して「ええ」は 138 名中 28 名のみが使用している単語であり、単語 idf 値は 4.93 である。この場合「ええ」を含む相槌は相槌「うん」よりも 4.93 倍出現しやすくする。

5.4 聞き役対話システム

5.5 システム概要

手法 1・2・3 をユーザが発話する内容に対して相槌を打つ聞き役対話システムとして実装した。ただしユーザが発話を完了した文字列に対して相槌を生成するのではなく、ユーザが発話入力欄に発話を入力している最中に相槌を行う。この相槌はユーザの発話入力中に、発話入力欄の内容に依存する形で必要に応じて生成・消滅し、ユーザの発話が円滑になるよう補助する。この仕様にした理由は、2.1 で述べたように相槌の主な出現位置は話し手の発話権中であると定義したためである。テキストによる対話システム上で発話権を保持しているタイミングに相当するのが、発話完了時ではなく発話の入力中であると考えられる。本システムは聞き役対話システムとして上述の相槌を打つ機能のみを持ち、ユーザの発話に対して能動的に回答する機能は持たない。

本対話システムはコーパスベースで動作する。コーパスは 4 で相槌調査に使用した名大会話コーパスを用いる。名大会話コーパスの話者 ID を起動時に指定することで、コーパスから自動で該当話者の相槌や言語モデルを抽出し、話者の個性にシステムが適応する。

表 1 の相槌分類のうち、B 分類・E 分類・L 分類はそのまま使用する。ユーザの発話を先読みする必要がある C 分類の相槌および相槌とみなさない O 分類の相槌は実装しない。R 分類の相槌は文脈によって発話内容が変化するため、コーパス中の R 分類の相槌に対して置換可能箇所を人手で設定し、実際の対話では直前に出現した重要単語に置換されて出力される仕様とした。重要単語の抽出には Docomo Developer Support の発話理解 API^{*3} を使用した。

5.6 相槌生成処理の流れ

ユーザの入力によって発話入力欄が変化したことを検知すると相槌生成処理を行う。相槌生成処理のフローチャートを図 3 に示す。以下でこのフローチャートについて説明を行う。

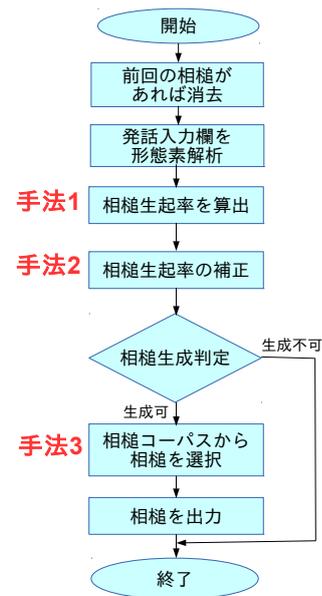


図 3 相槌生成処理のフローチャート

まず発話入力欄の内容を MeCab を用いて形態素解析し、発話入力欄の品詞情報を取得する。5.1 の手法 1 によって生成された相槌生起率学習データに基づき、この品詞情報から分類毎の相槌生起率を算出する。5.2 の手法 2 を用いてこの相槌生起率を補正し、この生起率を使用して相槌生成可否の判定を行う。この際に生成する相槌の分類も決定される。相槌を生成しないと判定された場合は処理終了となる。相槌を生成することが決まった場合は該当する相槌分類の相槌コーパスを呼び出し、5.3 の手法 3 によって重み付けされた確率に基づいて相槌を選択する。選択された相槌を画面に表示して処理終了となる。表示された相槌は次回の相槌生成処理の冒頭で消去される。

5.7 対話インターフェース

システムのスクリーンショットを図 4 に示す。このシステムは (1) から (7) の 7 パーツで構成されている。インターフェースの大枠は関野らが作成した WOZ 法における対話を補助するシステム [13] を元としている。以下に各パーツの説明を記す。

(1) ユーザ発話ログ表示部

ユーザの発話を記録するテキストボックス。「(対話番号)+U」を付与して対話のログを表示する。システムが発話する度に改行されることで位置を揃える。

(2) 相槌表示部

相槌が表示される領域。領域上のランダムな位置に相槌が表示される。システム発話と同一人物が発話しているように見せるため、(4) システム発話ログ表示部と共に薄く着色されている。

^{*3} https://dev.smt.docomo.ne.jp/?p=docs.api.page&api_docs_id=85 (3) 相槌

生成された相槌は (2) 相槌表示部内にラベルとして表示される。発話入力欄の内容が変化すると消滅する。

(4) システム発話ログ表示部

ユーザの発話を記録するテキストボックス。「(対話番号)+S」を付与して対話のログを表示する。システムが発話する度に改行されることで位置を揃える。本システムは能動的に発話する機能を持たないが、起動時の挨拶や連絡事項を伝達する場合にシステムが発話することがある。

(5) 発話入力欄

ユーザが発話を記入する入力欄。この場所に格納されている内容に基づいて相槌が生成される。

(6) 発話送信ボタン

発話入力欄に記入された文字列をシステムに送信するボタン。送信された内容は (1) ユーザ発話ログ表示部に記録される。(5) 発話入力欄で Enter キーを押下しても同様の処理が行われる。

(7) ログ出力ボタン

対話終了時に対話ログをファイルに出力するボタン。発話の内容・各発話が発せられた時刻・各相槌が発せられた時点の発話入力欄の内容がそれぞれ保存される。

6. 評価実験

実装した対話システムによって個性が表現されているか、また個性によって話しやすさに変化があるかを検証するために評価実験を行った。

6.1 使用データ

実験では名大会話コーパス中の話者 138 名の中からシステムで個性を發揮することが容易だと考えた個人として、若者を連想させる言葉遣いが強く出ている 10 代後半の女性「F046」・丁寧で女性らしい言葉遣いが強く出ている 60 代前半の女性「F082」・若い男性らしい言葉遣いが出ている 20 代後半の男性「M013」の 3 人を選出した。本実験ではこの 3 人の個性に対応したシステム 3 種類と、ベースライン 2 種類の計 5 種類を比較する。1 つ目のベースラインは全く相槌を打たず、黙って被験者の話を聞き続けるシステムであり「None」と呼ぶ。2 つ目のベースラインは名大会話コーパス 138 名の平均を取ったデータに適応するシステムであり「Average」と呼ぶ。

6.2 評価基準

本研究で実装した相槌によって話しやすくなっているかを、被験者に 5 段階で判定してもらうことによって検証する。また、コーパスの個性を再現できているかを 2 種類の方法で検証する。1 つ目の方法は対話したシステムに性別と年齢の情報を付与するならどうするかを選択式で答えさせる方法であり、実際にシステムの元となった話者の属性

表 3 実験で選択可能としたテーマ一覧

芸能人	ニュース	家族
ペット	TV 番組	恋愛
食べ物	旅行	友人・知人
本・漫画	ギャンプル	サークル
映画	ゲーム	学業
バイト	車・バイク	健康
ファッション	料理	歴史
音楽	夢	科学
スポーツ	地域	IT
勉強	天気	住居

と比較する。2 つ目の方法は名大会話コーパス中の F046・F082・M013 の 3 名の書き起こしテキストを被験者に見せて、3 名の中から直前に対話を行ったシステムの元となった個人を特定させる方法である。正しい話者を選択する被験者が多ければ相槌の個性を再現し他の話者と差別化することができていたと判断できる。

6.3 実験方法

実験方法は濱元らの研究の実験 2 の手順 [11] を参考にした。実験参加者には「5 分を目安に均等に語るができる」という基準で表 3 に示した 30 種類のテーマの中から 5 つのテーマを選択してもらい、各テーマについて 5 種類のシステムと対話を行ってもらう。対話はテーマについて被験者がシステムに対して語りかけ、システムが相槌を打つ形式で行われる。詳細な実験手順を以下に示す。

- 1) 被験者に実験の趣旨を説明し、同意書にサインを貰う
- 2) 実験手順を説明
- 3) 対話テーマを 30 テーマの中から 5 テーマ選択
- 4) 対話画面・対話における注意事項について説明
- 5) 被験者は選択したテーマでシステムと対話を行う
システムは 6.1 の 5 種類からランダム
- 6) 1 テーマ毎に話し終えた時点でユーザに「以上です」と入力してもらう
- 7) システムの印象について調査用紙を記入してもらう
- 8) 5)~7) を 5 回繰返し、全てのシステム・テーマで対話

7. 実験結果

話しやすさと相槌頻度の関係を図 5 に示す。話しやすさの平均点数は 5 段階をそれぞれ 2 点,1 点,0 点,-1 点,-2 点で計算している。システムの性別を推定させる設問の集計結果を表 4 に示す。平均 F 値は 0.65 である。また、全ての話者で正解の回答数が不正解の回答数を上回った。システムの年齢を推定させる設問の集計結果を表 5 に示す。ある年代前半の話者はその年代と直前の年代・ある年代後半の話者はその年代と直後の年代を正解として太文字で示した。対話を行ったシステムの元となった個人を特定させる設問の、適合率,再現率,F 値を表 6 に示す。平均 F 値は

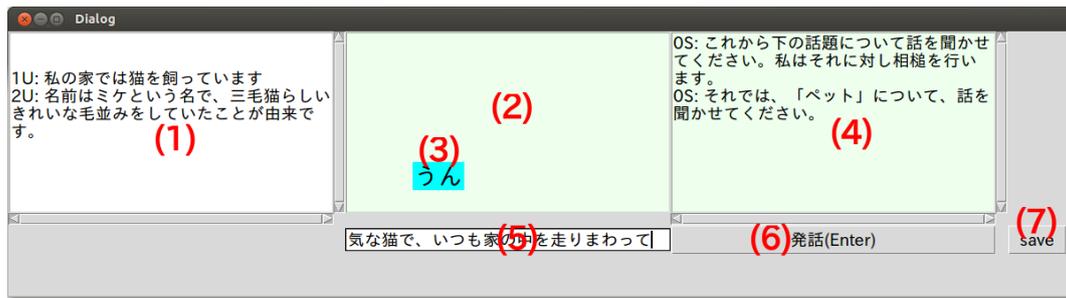


図 4 対話画面のスクリーンショット

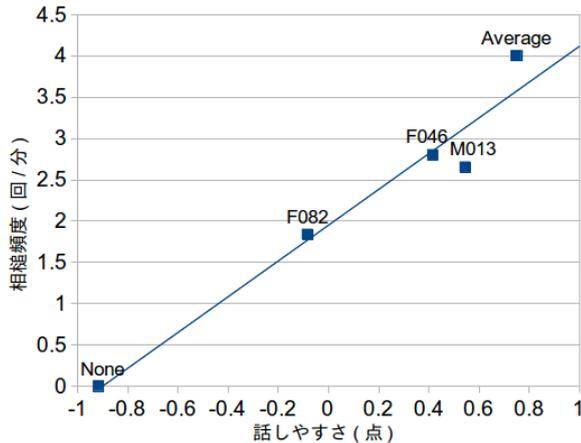


図 5 話しやすさとシステムの相槌頻度の比較

表 4 システムの性別を推定する設問の集計結果

	適合率	再現率	F 値
男性	0.70	0.50	0.58
女性	0.65	0.81	0.72

表 5 システムの年代を集計させる設問の集計結果

	F046	F082	M013
10代	2	3	1
20代	8	2	4
30代	1	3	5
40代	1	3	0
50代	0	1	0
60代	0	0	1
70代	0	0	1
80代	0	0	0
90代	0	0	0

表 6 特定個人を推定する設問の集計結果

	適合率	再現率	F 値
F046	0.58	0.41	0.48
F082	0.33	0.57	0.42
M013	0.08	0.08	0.08

0.33 である。

また、ある被験者がそれぞれのシステムと対話した内容を対話例として付録に示す。

8. 考察

図 5 より、相槌の頻度と話しやすさには強い相関関係 (相関係数 0.98) があることがわかる。これは濱元らの実験結果 [11] とも合致しており、相槌によって対話を円滑にすることには成功したと考えられる。表 4 で平均 F 値が 0.65 に達したこと・また性別の判定では全システムに対して正解の性別が多数を占めたことより、相槌によって話者の性別を再現する試みでは高い再現度を実現できたと考えられる。年代については表 5 より、F046 と M013 の年代は表現できているが、F082 の年代の表現はできていない。F046・M013 は若者を連想させる言葉遣いが強く出ていることから実験に選ばれた話者であるため再現しやすかったが、その 2 名ほど年代の特徴が顕著でなかった F082 については、年代を代表することまではできなかったと考えられる。特定個人を推定する設問では平均 F 値が 0.33 と低水準であり、特定の個人を再現するまでの個性の表現はできなかったと思われる。

個人に適応したシステムの話しやすさが、ベースラインとして設定した Average よりも低い値を示す結果となった。本研究では短時間の実験で個性を実感できるようにするため、分類毎の使用割合の強調・単語 idf 値の低い相槌選択といった、元の話者の相槌をそのまま再現せず特徴を強調する手法を用いている。これによってわかりやすく相槌の生成の仕方に差を付けることはできたが、「うん」等の純粋に相手の発話を促す機能のみを持つ相槌が減少し、情報量が大きく重く感じる相槌が増える現象が発生した。また、個性的な単語を持つ特定の相槌ばかりが何度も出現することが被験者に悪い印象を与えることがあったと自由記述による感想からわかっている。個性を強く表現したことによるこれらの現象が主な原因となり、Average に比べて個人に適応したシステムの評価が低くなったと考えられる。

9. 結論

本研究では個性的な相槌を生成する聞き役対話システムを実装することを目的とした。コーパスを用いて相槌の生起状況を調査したところ、相槌の頻度・相槌分類の使用割

合・使用する単語の3点に個性生成が期待できるとの調査結果を得た。調査結果3点をそれぞれ踏まえて3種類の個性生成手法を提案して聞き役対話システムに実装した。調査実験の結果としてシステムの相槌頻度が高くなるほど被験者の話しやすさが増すことを確認でき、システムの元となった話者の性別や年代を反映して被験者に実感させることが一定程度できていたことが確認できた。

本研究で実装した対話システムは聞き役としての機能に特化した対話システムであり、ユーザの発話に能動的に回答する機能は搭載されていない。既存の対話システムに対して今回の相槌を出力する機能を付与することは、対話システムと相槌機能の間で情報をやりとりする手法を決定できれば、比較的容易に実現可能と考えられる。今後は、発話を生成する対話システムの中における、相槌の個性生成機能の効果を検証したい。

参考文献

- [1] 河原達也：音声対話システムの進化と淘汰，言語・音声理解と対話処理研究会，Vol. 67, pp. 7-12 (2013).
- [2] 水上雅博, Neubig, G., Sakti, S., 戸田智基, 中村 哲：言語的個人性変換における言語モデルの適応と分析，人工知能学会第72回言語・音声理解と対話処理研究会 (SIG-SLUD) (2014).
- [3] 堀口純子：日本語教育と会話分析，くろしお出版 (1997).
- [4] 高梨克也, 伝 康晴：音声コミュニケーション研究のための分析単位，多人数インタラクションの分析手法，オーム社，pp. 19-64 (2009).
- [5] 泉子・K・メイナード：会話分析，くろしお出版 (1993).
- [6] エドワード・サビア：パーソナリティ特性としてのことば，言語・文化・パーソナリティ：サビア言語文化論集，北星堂書店，pp. 183-198 (1983).
- [7] 松原茂樹, 大野誠寛, 神谷優貴：音声対話コーパスに基づくあいづち生成タイミングの検出とその評価，言語処理学会年次大会発表論文集，Vol. 17, pp. 103-106 (2011).
- [8] Nass, C., Moon, Y., Fogg, B. J., Reeves, B. and Dryer, C.: Can Computer Personalities Be Human Personalities?, *Conference Companion on Human Factors in Computing Systems*, CHI '95, Vol. 43, ACM, pp. 228-229 (1995).
- [9] Mairesse, F. and Walker, M.: Personage: Personality generation for dialogue, *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 496-503 (2007).
- [10] 目黒豊美, 東中竜一郎, 堂坂浩二, 南 泰浩, 磯崎秀樹：聞き役対話システムの構築を目的とした聞き役対話の分析，情報処理学会研究報告. 自然言語処理研究会報告, No. 10 (2009).
- [11] 濱元貴範, 鈴木洋太, 神田智子：対話エージェントによるうなずきの日米文化間比較，情報処理学会研究報告. HCI, ヒューマンコンピュータインタラクション研究会報告, No. 10 (2011).
- [12] 黒崎良昭：談話進行上の相づちの運用と機能：兵庫県滝野方言について，*國語學*, Vol. 150, pp. 15-28 (1987).
- [13] 関野嵩浩：可搬コーパスを用いた対話シミュレーション支援，情報処理学会全国大会 (2013).

付録. 対話例

6.1 で示した5種類のシステムと、ある被験者がそれぞれ対話した際の対話内容を図6・図7・図8・図9・図10に示す。[]によって囲まれた部分がシステムの相槌であり、それ以外の部分は被験者の発話である。

私はトランペットを演奏します。
トランペットにはかなり長い間触ってきました。
具体的には今年で吹き始めてから12年が立ちました。
9歳からこの楽器に触れています。
昔はバイオリンもやっていましたが、母親にガミガミ言われながらなので練習に気持ちが入らずモチベーションの維持も難しかったです。
話が脱線してしまいましたが、
昔から音楽に関わらせてくれた両親には感謝しています。
以上です

図6 None に設定されたシステムと被験者との対話例

私の家では昔から犬を [うん] 飼っているんですよ [うん]。
私が生まれた直前に飼われ始めた犬は [うん、うん]
ロッキーという名前です [うん] やんちゃな犬でした [すごい] よ。
私が生まれた時から一緒だった [うん] その犬も、
残念ながら18歳で死んでしまいました。 [そうだよね]
今は、妹の [あーそうなんだー] わがままに [うん<笑い>]
押される形で [うんうん] 私の人生の [ほんと、ほんと、そうそう] 中では
2番目にあたる犬を実家で飼っていますよ。
こっちは犬はまだ1歳だから当然 [うーん]
死ぬことはないと思いますけど (笑)。
実家に [うん] 帰る機会が少なく、
あんまり会うことが [うん] できないから [うん]
少し寂しいかな [うん] ーって [うん] 思うことがあるよ。 [うん]
以上です

図7 Average に設定されたシステムと被験者との対話例

私が好きな本は基本的に [そうなんなんだ] 漫画ですが、 [だからなー]
一定数の本を読むことを心がけています [ほーんとだよ]。
ここで言う本は、軍事系が主ですが、
小説や新書なども [あ、そっかそっか]
読みます [あー、そっか、そっか][なーがーい]。
軍事系の本は値段が非常に [そうー] 高く設定されている
ことが多く [うーん。へー]、 [設定] この手の本を買うのには
多少財布と相談しなければなりません。 [あつ、そっか]
以上です [ねー、あーあ]

図8 F046 に設定されたシステムと被験者との対話例

自分は昔からゲームが大好きなんですよ。
最近やっているゲームは FPS というジャンルなんですけど、
自分の腕前が全てのゲームだから
多少シビアに映るかも。[ええ、ええ][あーすごいですねえ]
特に日本人にはあんまり受けがよくないかなー [ええ、ええ、ええ]。
だって [ええ、ええ]、日本人って、
俺最強ってしたい人が多いんですもん [ああ、そうですね]。
かけた [ああ、そうですね] 時間に [そうそう、あのとき] 対する
確実なリターンを欲しがっちゃうんですよね。[ええ、ええ]
だからかもしれないけれど、日本では RPG の方が
人気になってしまう気がしてちょっとさみしい、そう思います。
以上です

図 9 F082 に設定されたシステムと被験者との対話例

自分の参加しているサークルは [ふーん、まあね]
吹奏楽団と言って読んで字のごとく、
吹奏楽を好きな人で集まって楽しもう
という感じの [ほうほうほうほう] サークルです。
このサークル、各キャンパスで [まあな] 独立しているわけではなく
全てひとまとめで一つのサークルを作っています。
だから、吹奏楽団の人数は [ほうほうほうほう]
100人を [ほうほうほうほう] 軽く超えているんですよ。
この工学部の [おー、そうですね] サークルだけは特殊で、
近くの [ほうほうほうほう] 短期大学と
合同で [おー、そうですね] 活動しています。
なので、工学部内だけでも [ほうほうほうほう] 人数は
正規のサークルで一位を誇る大人数になっています。
以上です

図 10 M013 に設定されたシステムと被験者との対話例