

自然言語対話処理のための聞き返し発話生成手法

家守正人, 目良和也, 相沢輝昭

広島市立大学情報科学部

1 はじめに

人とコンピュータのインタフェースシステムに音声入力が使われはじめるようになってきたが, まだまだ正確に認識されているとはいいがたい。一部の音声認識システムは聞き返し機能を持っている。しかし, 常に同じ形式の文で聞き返されると, 人間は違和感を覚えるものである。

そこで本研究では, 人間がマイクなどを通して音声入力を行うとき, その言葉や話した内容に応じて多様な聞き返し表現を生成する手法を提案する。まず相手発話の誤りタイプから聞き返しパターンを決定する。次にパターンごとに用意している聞き返しの表現を選択する。また, 分からない言葉についてはその箇所を伝えて, もう一度正しい言葉を入力することを促す。

本稿では, 相手の発話内容に含まれる誤りタイプの分類を述べ, それに応じた聞き返しパターンと聞き返し表現について述べる。そして聞き返し発話の生成手順を示した後, 実行例と評価結果を述べる。

2 発話内容の誤りタイプの分類

相手の発話内容が聞き取れない状況を実際の会話コーパスから収集するのは困難である。そこで音声認識ツールを使って発話し, 認識した文字列が発声した内容と異なるとき, それを誤り文のサンプルとした。

文に誤りが生じた場合, 単語の区切りや語形が崩れるため, 形態素の並びが不自然になる。そのため, 誤りを含んだ文を判断する方法として, 本研究では形態素の並びに注目した。そして不自然になった形態素の並びをもとに誤りのタイプを以下の4種類に分類した。

- 連用形の並びによる誤り
- 終助詞の並びによる誤り
- 文の先頭に助詞がくる品詞の誤り
- 未知語に対する誤り

まず, 「連用形の並びによる誤り」について説明する。連用形の次にくる語は必ず用言であるので「述語連用形+体言」という形態素の並びがあれば, 誤りとみなす。誤り箇所は述語連用形から次の助詞の手前までとする。「終助詞の並びによる誤り」は, 終助詞が文末にくる助詞であることを考慮する。また, 「～なのよ。」の「のよ」ように, 終助詞「の」「よ」が続くこともある。そこで終助詞が文末か他の終助詞の

直前にある場合以外に誤り文と判断する。誤り箇所は対象の終助詞から終助詞以外の助詞の手前までとする。「文の先頭に助詞がくる品詞の誤り」は, 文の先頭にはほとんどの品詞が入り得るが, 誤りとみなされるもの一つとして, 先頭に格助詞がくる場合がある。一般の文の場合, 助動詞も文の先頭になることはない。しかし, 本研究では会話を対象にしているので, 相手の発言に対する同意「だね。」のような表現もあり得る。よって格助詞の誤りのみを扱い, 格助詞の前に何も来ない場合, 誤り文とみなす。誤り箇所は先頭から次の助詞の手前までとする。

上記の3つの誤り以外に, 「未知語に対する誤り」がある。形態素解析では解析できなかった語を「未知語」とするので, 本研究でもこれを未知語として扱う。未知語になるものはたいてい, 「ゃ」「ゅ」「よ」などの拗音や, 「Φ」などの特殊語である。誤り箇所の検出について未知語のある語から助詞までを行うと, 本来聞き返したい誤りの範囲が足りないことがおこる。よって未知語が出てきた場合, 検出の範囲を限定しない。

3 聞き返し発話の分類

3.1 聞き返しパターン

広島市内の公立大学生17名(男性14名, 女性3名)に質問紙調査を行った結果をもとに誤りのタイプに応じた聞き返しのパターンを調査した。結果の分類に際して, 尾崎の日本語学習者による聞き返しの表現の分類を参考にした[1]。尾崎は聞き返しの表現を動詞型(例: 分かりません, もう一度言ってください), 名詞丁寧型(例: 今ですか?), 名詞普通型(例: 今?), 不完全型(趣味って?), 間投詞型(え? は?)の5つに分類している。質問紙調査と尾崎の聞き返しの表現を照らし合わせてシステムに適応しやすいように分類し, 聞き返しのパターンを以下の4つに分類した。

- 単純な聞き返し
- 文全体の聞き返し
- 誤り箇所に対する聞き返し
- 言い換え

本実験では複雑化を避けるために「名詞丁寧型」「名詞普通型」「不完全型」の3つの表現を1つにまとめ, 「誤り箇所に対する聞き返し」とした。

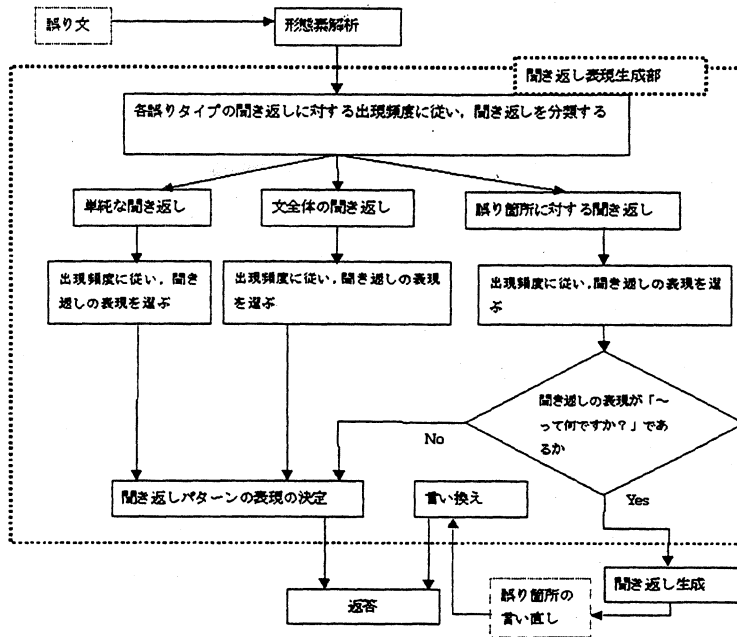


図1：聞き返し発話生成手順

単純な聞き返しとは「え?」「は?」などの感動詞により、素早く相手の反応を求める聞き返しである。文全体の聞き返しとは相手の発話がほとんど聞き取れず、繰り返して話すことを要求する聞き返しである。誤り箇所に対する聞き返しとは、相手の発話の中で大まかには聞き取れるものの、一部の誤り箇所を確認するための聞き返しである。言い換えは、誤り箇所に対する聞き返しを行った後、相手が言い直した言葉を確認するという意味で、聞き返しのパターンとして考慮に入れた。

表1に各誤りタイプの聞き返しパターンの出現頻度を示す。ただし「言い換え」は誤った箇所を相手が入力し直したときのみ用いるのでこの表から除いた。

表1：聞き返しパターンの出現頻度

聞き返しパターン	単純な聞き返し	文全体の聞き返し	誤り箇所に対する聞き返し
通用形の並び	43.1%	13.8%	43.1%
終助詞の並び	33.3%	0%	66.7%
先頭に助詞	63.4%	0%	36.6%
未知語	0%	100%	0%

3.2 聞き返し表現

聞き返しパターンごとに用いられる聞き返しの表現を分類し、その出現頻度を求めた。表2にそ

れらの出現頻度を示す。

また、単純な聞き返しのうち、「はあ?」「へ?」「はい?」は、相手に失礼な態度をとっているとされるため、出現頻度の高かった「えっ?」に含める。

表2：聞き返しの表現の出現頻度

単純な聞き返し		文全体の聞き返し		誤り箇所に対する聞き返し	
表現	出現頻度	表現	出現頻度	表現	出現頻度
えっ?	92%	意味がわかりません。	40%	～?	33.3%
え、何?	4%	今、何ていいましたか	40%	～って?	36.7%
ん?	4%	よくわかりません	20%	～ですか?	13.4%
				～って何ですか?	16.6%

4 聞き返し発話生成

4.1 聞き返しパターンの生成手順

図1は聞き返し発話生成手順である。入力された誤り文から形態素解析結果に基づいて、誤りを検出し、その誤りタイプを調べる。次に4種類の誤りタイプ各々に対して、出現頻度に従った聞き返しパターンを決定する。そして決定した各パターンにおける聞き返しの表現を、出現確率に従って選ぶ。ただし「誤り箇所に対する聞き返し」にのみ、相手に誤り箇所を言い直してもらう表現があるとき、「言い換え」を行う。図中の点線部分

実際の聞き返し生成部分である。

4.2 実行例

聞き返し文生成の例を次に示す。質問文「何か集中できることはありますか」に対する返答として、「趣味は三味線です」と発話したにもかかわらず、システムは「し実は三味線です」と受け取ったとする。この場合、まず受け取った文に対して形態素解析を行う(図2)。構文の誤りがない場合または聞き返しが3回以上繰り返されている場合はあいづち生成システム[2,3]へ移る。この例では、形態素解析の結果、「する」の連用形「し」に關する並びの誤りがある。そこでこの「し」を誤り箇所の特徴とし、次の助詞「は」の一つ前の語までを誤り箇所の終わりとする。よって誤り箇所は「し実」となる。

し	実	は	三味線	です
動詞連用形	名詞	係助詞	名詞	助動詞

図2：形態素解析結果

誤りのタイプに応じた聞き返しのパターン頻度(表1)に基づいて聞き返しパターンを選択する。乱数出力を生起確率に当てはめると、今回は誤り箇所に対する聞き返しが選ばれる。聞き返し表現も同様に出現頻度に従って選択する(表2)。そして乱数により「し実ですか?」が選ばれる。これを聞き返しの表現とし、返答する。

5 実験、評価

返答発話がどの程度自然な聞き返しを行われているか質問紙調査で評価した結果、コンピュータの聞き返しを人間とみなしたものは45.8%であった。誤り箇所に対する聞き返しの表現を文字列で表した場合、会話としてぎこちなさを感じたためと思われる。

表3は誤りタイプとそれに対する聞き返しパターンの組み合わせで、システムの結果を人間の聞き返しとみなした割合を示している。5件法で不適正を-2に、適正を2に数値化し、どちらでもないものは0として回答された評価値の平均を調べた。また、評価値の平均を最高値で割ったもの(適正度)も評価基準に用いた。人間としてみなされた割合が高く、適正度が大きい場合、自然な聞き返しを行っていると考えられる。反対に人間としてみなされた割合が低く、適正度が小さい場合聞き返しとしての役割を果たしていないことになる。例えば、連用形の並びによる誤りに対して単純な聞き返しをする場合、システムの聞き返しを人間の聞き返しと思ったのは73.68%、適正度は0.36となる。これは、連用形の並びに対して誤った場合、単純に聞き返すものは人間に近く、聞き返しかたも自然であることをあらわしている。

一方、連用形の並びによる誤りに対して、誤り箇所の聞き返しを行ったものは適正度が-0.11となっている。このとき、自然な聞き返しができていないことになる。単純な聞き返しにおいては人

間とみなされる割合が高く、また、このパターンの聞き返しの出現頻度が高いため(表1)、自然な聞き返しを行っているように見える。

表3：誤りタイプに応じた聞き返しパターンの適正度

誤りタイプ	聞き返しパターン	人間とみなされたもの	評価値平均	聞き返しとしての適正度 (評価値平均/最高値)
連用形の並び	単純	73.68%	+14	0.36
	誤り箇所	42.5%	-12	-0.11
	全文	63%	+12	0.32
先頭に助詞がくる	誤り箇所	26.31%	-15	-0.02
終助詞の並び	単純	89%	+1	0.02
	誤り箇所	21%	+13	0.34
	未知語	全文	47.3%	+8

6 まとめ

本論文では自然対話において、単調な繰り返しでしかない聞き返しに、様々な表現を追加するための手法を提案した。この手法では、多様な表現を得るために、音声発話に生じやすい誤りを聞き間違いとして扱い、誤りのタイプを分類した。そして、誤りタイプに応じて聞き返しをパターン別にわけ、さらに聞き返しの表現を分類した。

聞き返しのパターンは「単純な聞き返し」「誤り箇所に対する聞き返し」「文全体の聞き返し」の3つに分けた。聞き返しのパターンを決定するにあたっては、質問紙調査の結果から出現頻度を割り当てた。各パターンの聞き返し表現は同じ質問紙調査の聞き返し結果より3、4種類得られた。誤り箇所に対する聞き返しの場合、表現方法によっては相手に別の表現で言い直してもらうように要請することがある。相手が言い直しをしたとき、誤りを正しい文に言い換えて聞き返す動作を付加する。これを「言い換え」とした。こうして構築した聞き返しシステムについて、システム評価に対する質問紙調査を行ったところ、対話をする際に自然な聞き返しをしていることがわかった。しかし言い換えについては評価が低く、問題点が多かった。

今後の課題としては、誤り箇所を聞き返す表現に対する評価が低かったので、より多くのサンプルを採取する必要がある。また、質問データや、相手発話の内容を保存しておくことで、聞き返しをすると共に相手に話題を提供し、入力装置に音声インターフェースを採用することで本格的な対話処理システムを構築していく。

参考文献

- [1] 尾崎明人：「聞き返し」のストラテジーと日本語教育，pp. 251-263『日本語研究と日本語教育』名古屋大学出版会(1992)
- [2] 末廣朋子：自然対話におけるあいづち生成システム，広島市立大学情報科学部平成12年度卒業論文(2001)
- [3] 佐藤康将，目良和也，市村匠，山下利之，相沢輝昭，吉田勝美：肯定/否定意図を検出するチャットシステムのためのあいづち生成手法，第17回ファジィシステムシンポジウム講演論文集，pp. 513-516(2001)