

未知語対応した情報検索用音声対話インタフェースの検討

井上剛[†] 沖本純幸[†] 續木貴史[†] 九津見洋[†]
松下電器産業(株)先端技術研究所[†]

1. はじめに

近年のデジタルコンテンツの増加とそのデータベース化に伴い、大量のコンテンツから所望のコンテンツを検索する機能は益々重要に成りつつある。

情報検索機能のインタフェースとしては、キーボードやリモコンなどボタンによる入力も可能であるが、我々は誰もが簡単に入力できるインタフェースとして音声インタフェースを用いた対話型情報検索システムの実現を目指している。

音声対話型情報検索システムの問題点の一つとして、様々なシステムの想定外の発話が行われることがある[1]。特に電子番組表(Electronic Program Guide: EPG)を利用した番組検索のように、日々更新するデータベース検索に音声インタフェースを適用する場合、データベース検索に必要な認識対象語彙も日々変化するため、ユーザの辞書外語彙(Out of Vocabulary: OOV)発声による誤認識が大きな問題になることが予想される。実際、以前我々がEPGを対象とした番組検索時の発話傾向を調べた際も[2]、多くの辞書外発声が確認された。

このような辞書外語彙の発声に対して従来の多くのシステムは誤認識結果のまま対話を進行するか、再度入力を促す。しかし、それではユーザは自分が辞書外語彙を発声したことに気付かないため、何度も同じ発声を行うことになってしまう。この問題に対し我々は、新規語彙獲得[5]や不要語の無視[6]で利用されている未知語検出技術をシステム応答に反映させ、ユーザに辞書外語彙の発声が行われたことを伝えることが有効であると考えている。

本稿では、EPGデータを対象とした音声対話型番組検索システムを想定し、未知語検出に対応したシステム応答の有効性について検討を行った結果について報告する。

2. 未知語検出対応のシステム応答

EPGから取得した番組情報(通常当日~10日前後の情報)を検索するために、EPGに含まれる単語から音声認識辞書を作成した場合、ユーザは

その語彙を完全には把握できないため、辞書外語彙を発声する可能性がある。例えば出演者名を用いた番組検索の場合、検索対象外の人物名、すなわち辞書に含まれない人物名を発声することがある。

この問題の対処法として、EPGデータから作成した人物名辞書を用いて認識を行い、未知語検出を行って、未知語だと判定された場合、その結果のみを伝える(図1 Type1)方法と、EPGデータから作成した人物名辞書の他に、大語彙人物名辞書を用いて認識を行い、ユーザが発声する人物名に対して、その人物名は検索対象に存在しないことを伝える(図1 Type2)方法が考えられる。

しかし、Type2のシステムを実現するには、非常に大きな認識辞書が必要となるだけでなく、その辞書には、日々新たに現れる人物名の名前も追加する必要がある。さらに、そのような非常に大きな認識辞書に対してリアルタイムかつ精度良い認識率で認識を行うのは現状では極めて困難である。

そこで、我々は未知語検出結果のみを伝えるType1のシステム応答がどの程度有効であるか検討するために、出演者名による番組検索において、未知語検出の結果のみを伝えるシステム応答(Type1)と、ユーザの発声した出演者名を含むシステム応答(Type2)の2種類の応答を用意し、それぞれのシステム応答に対するユーザ発話や主観評価にどのような傾向があるか比較を行った。

Type1: 「その人が出ている番組はありません」

Type2: 「<人物名> が出ている番組はありません」

図1: 未知語対応したシステム応答

3. 実験方法

評価実験はユーザの反応を適切な割合で収集するために、検索語としての人物名は固定とし、システムの応答はWizard of OZ方式を用いた。実験システムはPC上で動作し、画面上にはアニメキャラクターのエージェントを表示する。ユーザの発声に対して実験者はシステム応答作成し、エージェントから音声(自社開発のテキスト合成エンジン)と吹き出しによりユーザに伝える。

被験者には検索語の人物名リストと簡単なマニュアルを提示し、下記の内容を伝えた。

Spoken Dialogue Interface with Out of Vocabulary Detection for Information Retrieval.

†Tsuyoshi Inoue, Yoshiyuki Okimoto, Takashi Tsuzuki, Hiroshi Kutsumi · Advanced Technology Research Laboratories Matsushita Electric Industrial Co.,Ltd.

- ・EPG という番組情報を持つエージェントに対して音声により出演者による番組検索を行う。
- ・音声認識は誤認識する可能性がある。
- ・検索対象の EPG データは日々変わるのでリスト上の人物名は検索対象 DB に無い可能性がある。

実験者は予め定められた計画表に基づきシステム応答を決定し出力する。システム応答には上記 Type1, Type2 および検索が成功したシステム応答の3種類を用いた。

実験は EPG を用いた番組検索を行ったことのない20代から40代の男女17名(男性8名、女性9名)を対象に下記の2種類の条件下で実験を行い、各システム応答文に対する7段階評価のアンケートによる主観評価と発話傾向を調べた。

実験1(被験者7名):

Type1の応答が一定の順序と割合(10名の人物名リスト中4名)で含まれるシステム応答を出力。

実験2(被験者10名):

Type1とType2の両方の応答が一定の順序と割合(20名の人物名リスト中Type1、Type2共に4名)で含まれるシステム応答を出力。

4. 実験結果と考察

実験1におけるシステム応答(Type1)に対する7段階評価のアンケートによる主観評価と言い直し回数の平均値を表1に、実験2におけるシステム応答(Type1とType2)に対する7段階評価のアンケートによる主観評価と言い直し回数の平均値を表2に示す。

4.1 実験1についての考察

表1よりType1の応答に対する主観評価値は4.6であり、「やや良い=5」に近い評価を得ることができた。また、平均言い直し回数は0.57回であり、2回に1回以上の割合で言い直していることが分かる。これは、Type1の応答では、ユーザの発話に対して発話内容のフィードバックが無い場合、ユーザが誤認識の可能性があるかと判断し、言い直しを行ったためである。この結果よりType1の応答は少々言い直しは存在するものの、音声インタフェースとして必ずしも不十分だとユーザは感じているわけではないということが分かった。

表1: 実験1における主観評価と平均言い直し回数

システム応答	主観評価	平均言い直し回数
Type1	4.6	0.57 [回]

4.2 実験2についての考察

表2より両応答における主観評価は、未知語検出結果のみ利用して検索結果を伝える応答である

Type1に対して、出演者名(発声内容)を含む応答であるType2の方が高い評価を得ており、平均言い直し回数も少ない。このことより、未知語対応のシステム応答として、ユーザの発話内容を返す応答は有効であることが分かる。

表2: 実験2における主観評価と平均言い直し回数

システム応答	主観評価	平均言い直し回数
Type1	2.6	0.85 [回]
Type2	6.1	0.05 [回]

4.3 実験1と実験2の結果比較

実験1と実験2において同じシステム応答を行っているにもかかわらず、実験2の方が評価は低く、平均言い直し回数も増える結果となった。この原因として、「より良い応答があることを知ったため相対的に悪い評価となった」「2種類の応答が混在するためにType1の応答が認識誤りかもしれないと感じるようになった」という2つがあると考えられる。

5. まとめと課題

日々データが更新され、認識対象語彙も変化する情報検索タスクである EPG を利用した番組検索システムを想定した未知語検出対応のシステム応答におけるユーザの発話傾向及び主観評価を調べた。

その結果、未知語検出結果のみ利用して検索結果を伝えるシステム応答は単独評価としては、評価値4.6を得ることができたが、ユーザの発声内容を含む応答と比較すると、低い評価を得る結果となった。

今回の実験ではユーザ発話を返す理想的な応答と単に未知語検出結果のみを伝える応答との2種類のみの比較を行ったが、今後は未知語検出技術を用いて実現可能な範囲でより理想的な応答を返すシステムの実現を目指す予定である。

参考文献

- [1] 安達, 駒谷, 河原, "音声対話情報検索システムにおける想定外の発話の分析とその対処", 人工知能学会研究会資料 SIG-SLUD-A001-2
- [2] 井上, 西崎, 小沼, 桑野, 脇田, "EPG 番組情報検索向け音声インタフェースにおける認識語彙選択手法の検討", 情報処理学会第65回全国大会, 2F-1
- [3] 小川, "ロボット対話における自然な新規語彙の獲得", 人工知能学会研究会資料, SIG-Challenge-0216-3
- [4] 甲斐, 中川, "冗長語・言い直し等を含む発話のための未知語処理を用いた音声認識システムの比較評価", 電子情報通信学会論文誌, D-Vol.J80-D-No.10 pp2615-2625 1997-10