

6H-1

日・英電話取り次ぎ会話の音声対話処理

小島順治、

トニー・コックス

NTT ヒューマンインタフェース研究所 British Telecom Research Laboratories

1. はじめに

筆者らは、自然な音声入力を対象とした音声会話システムの実現を目指して音声対話処理の研究を行っている。電話の取り次ぎ会話のような極めて限定されたタスクにおいては、音声認識技術を用いずにある程度会話内容が推定でき、会話の制御が可能であることが既に明らかになっている。⁽¹⁾⁽²⁾ 今回、更に英語と日本語の電話取り次ぎ会話について会話の特徴を比較すると共に、不在時の自動応答システムに適用する対話処理法を検討したので報告する。

2. 音声収録

まず、日本語、英語による人間同士の電話の取り次ぎ会話を同一条件で収録した。収録にあたっては、実際の電話を用いずにスタジオでそのような場面を設定して行った。話者については、性別、年齢、職業等が偏らないよう配慮した。また、相手や状況により異なることを想定してオフィスにける場合と、親しい友人の家庭にける場合、及び各々の在、不在の場合を収録した。収録条件を表1に示す。

表1 会話音声の収録条件

項目	内 容
発呼者	日本人：20人 英語を母国語とする人：20人 (各々男女各10)
被呼者	日本人：2人 英語を母国語とする人：2人 (各々男女各1)
会話環境	電話を想定(簡易なシナリオ有り) 互いの顔は見えない 会話の開始：被呼者からの発声
シナリオ	オフィスにける場合：2種 家庭(友人)にける場合：2種 (各々在、不在)
収録環境	場所：録音スタジオ 入力・受聴：ヘッドセット使用 (ヘッドホン、接話マイク)

3. 会話推移の分析

収録した音声を聴取・文書化し、会話内容に着目して分析を行った。日本語の電話取り次ぎ会話として、通話を望む相手が在室の場合の例を表2、英語の場合の同様の例を表3に示す。

表2 日本語電話取り次ぎ会話の例

話 者	会 話 例
被呼者	はい、NTT研究所です
発呼者	えっと鈴木ですが、田中さんいますか
被呼者	あっ、少々お待ち下さい

表3 英語電話取り次ぎ会話の例

話 者	会 話 例
被呼者	Hello. NTT Laboratory. Can I help you?
発呼者	Er I'd like to speak to Mr. Smith.
被呼者	Sure, he is here. One moment please.

電話取り次ぎ会話においては、例のように日本語、英語とも発呼者、被呼者の会話内容がパターン化している。そこで、被呼者発声後の発呼者の会話推移に着目して内容をカテゴリー別に分類することとした。⁽²⁾ カテゴリー種別を表4に示す。

表4 会話内容のカテゴリー種別

パターン	内 容
1	発呼者名のみが含まれる(C)
2	被呼者名のみが含まれる(R)
3	発呼者・被呼者名が含まれる(CR)
4	もしもまたは、Hello(M)
5	その他(挨拶、繰り返し等)(O)

表4のカテゴリーに基づいて、被呼者の第1声に続く発呼者の第1声について日本語、英語の会話内容を分析した後、カテゴリー別に分類し、比較した。結果を表5に示す。

表5 発呼者第1声の分析結果の比較

ボタン	日本語	英語
1 (C)	1 4 (17.5%)	
2 (R)	3 6 (45.0%)	6 6 (82.5%)
3 (CR)	1 3 (16.3%)	1 4 (17.5%)
4 (M)	3 (3.7%)	
5 (O)	1 4 (17.5%)	
会話数	8 0	8 0

比較の結果、英語の第1声の識別は日本語に比べ容易であるが、不在時の対応で重要な情報である発呼者の名前がわかりにくいことが分った。

4. 対話処理法

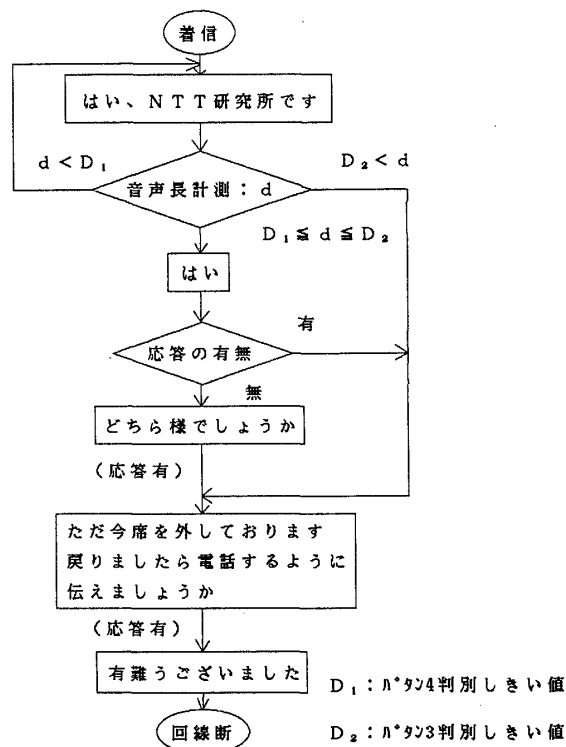
不在時に会話形式で自動応対を行う留守番対話システムを念頭に、話者の内容に応じて複数のメッセージの中から1つを選択出力する会話手順制御法の検討を行った。日本語の会話については、

(1) 音声長により、ボタン3、ボタン4、およびそれ以外の3種のグループに判別できる。

(2) ボタン1とボタン2は、相づち「はい」の後の音声の有無により判別できる。

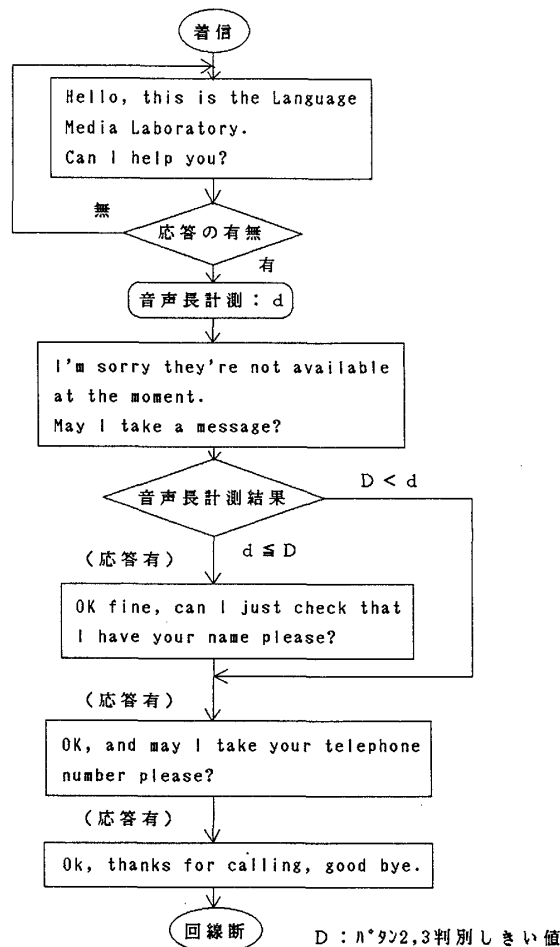
ことを利用する。(1)(2)

日本語の会話手順の検討結果を図1に示す。



英語の場合も同様に、ボタン2とボタン3が音声長により判別できることを利用する。英語の会話手

順の検討結果をを図2に示す。



現在、この処理を盛り込んだ音声会話システムをWS上で構築中である。

5. おわりに

日本語と英語の電話取り次ぎ会話を分析した結果、日、英では異なる対話処理が必要ことが分った。このことは、自動翻訳電話等の会話システムを構築する場合、習慣や常識等の差を考慮することがよりよいマン・マシンインターフェースを実現する上で重要であることを示している。今後は、さらに詳細な検討を行うと共に、システムの構築・評価を進める予定である。最後に日頃御指導戴く、当所川嶋部長、中津グループリーダーに深謝します。

[文献]

- (1) 五味他：「統計的知識を利用した音声対話処理」、情報処理学会第33回全国大会講演論文集、7N-4、1986
- (2) Kojima et al. : "A voice man-machine communication system based on statistical information received from telephone conversations", AVIOS'87, 1987