

## 蓄積型音声対話システムにおける発話促進手法

西 宏之                      北井 幹雄

NTT ヒューマンインタフェース研究所

**あらまし**                      蓄積系音声対話サービスにおいて発呼者が音声メッセージを残すことなく切断するという問題点を改善するため、発呼者の発話を促進する音声対話手法について報告する。はじめに、蓄積系音声サービスのうち、もっとも普及している留守番電話を対象に、切断の原因となる要因の抽出と、それらを改善し録音率を向上させるため「対話の快適さ」、「対話の必要性」、「情報提供による興味の喚起」の3つの仮説を抽出した。次に、上記仮説のうち現状の技術で検証可能な「対話の快適さ」、「情報提供による興味の喚起」の2仮説を検証するための評価実験の条件を整理し、評価実験を行なった。その結果、従来型に比較して、「対話の快適さ」を提供する対話録音型が平均オピニオン評価値で約0.8ポイント、「情報提供による興味の喚起」を行なう情報提供型が約0.6ポイント向上することがわかった。最後に、今後の課題として、さらに録音率を向上させるため「対話の必要性」を満足するサービスについて提案する。

**和文キーワード**                      音声対話, 発話促進, 留守番電話, ボイスメール, 伝言サービス

## Utterance Promoting Method on Speech Dialog systems

H. NISHI                      M. KITAI

NTT Human Interface Laboratories.

1-2356, Take, Yokosuka-shi, Kanagawa 238-03, Japan

**Abstract**                      New dialog promoting methods for voice-storage dialog systems are discussed. This paper describes a some hypothesis that attempts to put at ease those callers who hesitate or are shy by designing an interactive dialog or that offers attractive information for the caller. The following two hypothesis are introduced.

- i) The less information per one utterance that is required by a system, the easier users feel.
- ii) The more attractive the information that is given after an utterance, the more the callers will leave messages.

The experimental conditions are explained to evaluate the previous hypothesis using an opinion score. The system consists of PC, telephone line interface board, and control software. For eliminating the effect of recognition accuracy, dialogs are designed without speech recognition. Finally, the experimental results are described which indicate the usefulness of each hypothesis, and show an increase of 0.6 -0.8 points in mean opinion score.

**英文 key words**                      speech dialogue, dialog system, answering machine, voice mail.

## 1 はじめに

音声メディアとした通信サービスの対話処理およびヒューマンインタフェースの検討を進めている [1] ~ [5]。音声メール、伝言サービス、留守番電話などの蓄積系サービスは、被呼者が不在の場合でも、音声による伝言が残せることから、多忙な現代社会における通信システムの有効な付加サービスとして注目される。

しかし、その反面、そのようなシステムが導入されていることを知らずに発信した場合、メッセージを残すことなく切断する発呼者が多いという問題点がある。

本論文は、36名の被験者を対象としたオピニオン評価実験結果を基に、現状のサービスに対してメッセージを残さずに切断する原因の分析と、その原因を払拭し、伝言の録音率を改善する手法について報告するものである。

はじめに、蓄積系音声サービスのうち、もっとも普及している留守番電話を対象に、切断の原因となる要素の抽出と、それらを改善し録音率を向上させるための仮説を抽出する。次に、仮説を検証するための評価実験の条件を整理し、実験結果を述べる。最後に、さらに録音率を向上させるため将来的に実現すべきサービス機能について提案する。

## 2 切断要因の分析

伝言を残さずに切断する発呼者の心理状態として下記のような状況が推測できる。

- a. 人が出ると思って電話したにもかかわらず、留守番電話に応答され驚いてしまう。
- b. 用件をうまく整理することができず、まごついているうちに思わず切ってしまう。
- c. 留守番電話では用事を済ませることができない。
- d. 機械に向かって話すのは不愉快である。
- e. 用事があった電話するのではなく、会話そのものが目的である。
- f. その他

a. および b. は留守番電話が応答することを知らずに電話する発呼者の心理から類推される。c. は留守番電話がシステム側のガイダンスメッセージと発呼者の伝言メッセージの一往復の処理であり、人間とシステム間のインタラクションが期待できないことから推測した。d. は、音声人間同士のみのコミュニケーション手段としてとらえ、マンマシンインタフェースとしては不自然さを覚えるユーザがいるであろうということから抽出した。e. は被呼者との会話を楽しむことを第一の目的とするユーザの心理を代弁したものである。

## 3 発話促進要因の仮説

次に上記切断要因を吟味し、その結果を基に、発声に対する動機を発呼者に与え、録音率を向上させる要因を考察する。

前節の a. は、近年不在時に留守番機能や音声メール装置への転送サービスを利用するユーザが徐々に増大するにつれ徐々に改善していくものと思われる。このようなユーザは、着信後応答メッセージを聴取した後、伝言メッセージの録音の段階で切断するケースが多いので、応答メッセージの工夫や対話処理により改善できる可能性が高い。b. の場合も同様である。このような場合、伝言を一度にまとめて発声させるのではなく、名前、被呼者側からかけ直すための電話番号、伝言など複数の短いメッセージに分割して発声させることにより、伝言を短時間に整理する心理的圧迫から解放されることが期待できる。即ち対話型の伝言録音による「対話の快適さ」の向上が録音率向上をもたらす第一の仮説である。図 1 に示す従来型の留守番電話の対話シーケンスと比較すると図 2 に示す対話型のシーケンスで発呼者の思考の負担が分散され軽減されることが推測できる。

c. はすでに述べたように留守番電話では人間とシステム間のインタラクションが期待できないことから推測される問題である。将来システムがインタラクティブな対話能力を持つことを想定すると、対話を行なうことが発呼者にとって問題解決の手段となる、即ち対話が必要であるならば、伝言を残すはずである。「対話の必要性」が第二の仮説である。川端 [6] が指摘しているように、システムが知的にならなければ、c. の意見をもつユーザに発話させることは困難であると思われる。しかし、将来的には、音声認識の高度化に加え、被呼者のスケジュール情報の知的な管理システムが統合されることにより、本人に代わってシステムが発呼者の用件を済ませてしまうサービスが期待できる。このようなサービスの実現のためには多くの困難な問題が山積しているが、音声処理および知識処理の将来の目標の一つを与えるものとして興味深い。

d. および e. については、被呼者側が機械である場合は本質的に解決不可能であるように思われる。しかし、d. および e. の要因を一種の壁とするならば、その壁を乗り越えるほどの魅力に対話サービスに盛り込むことができるならば、発呼者は録音する意思をもつことが期待できる。具体的には発呼者にとって極めて有益な情報を伝言メッセー

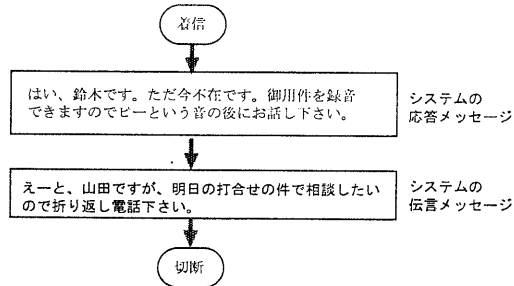


図 1: 従来の留守番電話の対話例

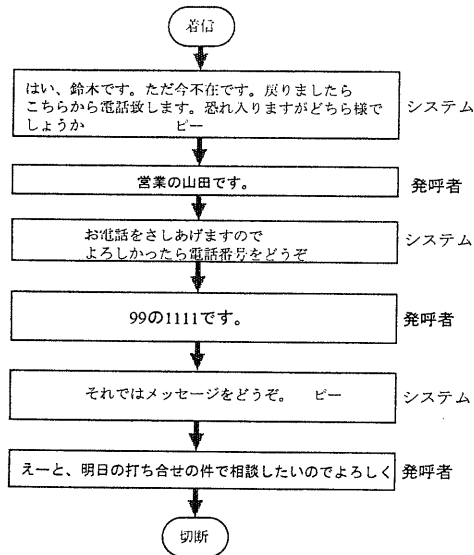


図 2: 対話型留守番電話の対話例

ジを残した発呼者のみに聴取させるという手法である。即ち、「情報提供による興味の喚起」が第三の仮説である。本仮説の例を図3に示す。

上記の三仮説を以下に整理する。

「対話の快適さ」 一度にすべての伝言メッセージを発声するのではなく、名前、電話番号など少量の情報に分けて発声することにより心理的負担を軽減し、対話の快適さを増大させる。また、将来的には相づちやBGMなどによりさらに快適さの増大を図る。

「対話の必要性」 発呼者にとって音声対話を行うことの必要性が満足されるよう、システムを高機能化する。この場合、一方的な伝言ではなく、知的かつインタラクティブな対話を成立させようシステム機能の実現が必要である。

「情報提供による興味の喚起」 伝言録音後に、発呼者にとって有益な情報をもつ第二のメッセージを送出する。

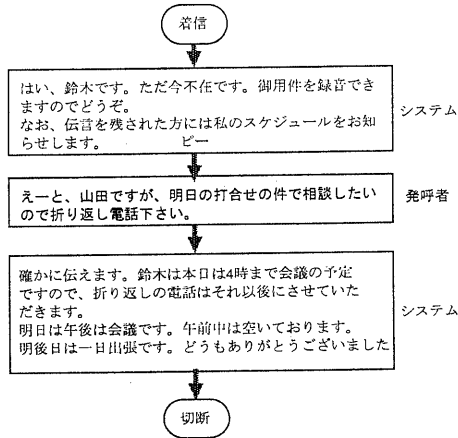


図 3: 情報提供により興味を喚起する対話例

#### 4 実験条件

前記仮説を検証するための評価実験を行うにあたり、表 1 の条件を設定した。なお、上記仮説のうち必要性を検証することは、自動的に動作するシステムを構築することが技術的に困難であることから今回の報告からは除外した。したがって、従来システムと比較して対話録音型と情報提供型がどのような印象を与えるかを評価した。

表 1: 実験条件

項目	実験条件
被験者数	合計 36 名 男性：19 名 女性：17 名 20 歳～39 歳：14 名 40 歳～59 歳：22 名
調査項目	(1) 従来型に対する評価 (2) 切断する際の理由 (3) 対話録音型に対する評価 (4) 情報提供型に対する評価
アンケート内容	アンケート用紙参照 (付録)

評価実験のためのシステムは図 4 に示すように電話回線ボードを装着したパーソナルコンピュータで実現した。評価試験の手順を下記に示す。

- (1) 事前アンケートの記入 予め、各被験者に A 4 一枚のアンケート用紙 (付録参照) を配布。アンケート用紙は上下に分かれており、上半分は事前アンケート、下半分は実際のアンケートである。事前アンケートには予め記入しておいてもらった。
- (2) システムへの電話 被験者は指定された番号に電話をする。
- (3) オペレータからの説明 オペレータが簡単な説明をした後、一旦電話を切る。
- (4) 各種対話方式の体験 再度被験者が電話すると留守番電話がでる。ガイダンスに従ってメッセージを録音する。

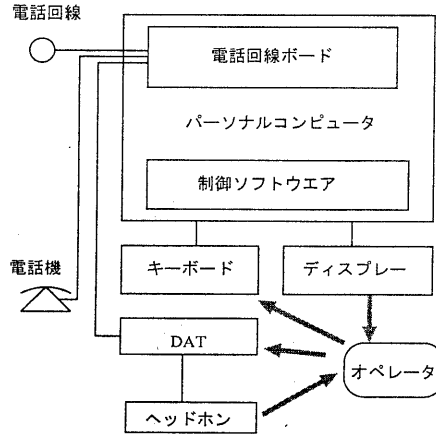


図 4: 評価システムの構成

- (5) 評価 電話を切りアンケート用紙に記入する。
- (6) 提出 アンケート用紙を郵送で提出する。

なお、システム側のガイダンスは図 1, 図 2, 図 3 のシステム側の応答メッセージと同様とした。被験者が発声する伝言メッセージの内容は、特に指定せず任意の内容とした。

## 5 実験結果

### 5.1 対話方式の比較結果

付録に示したアンケート項目のうち、事前アンケート 1, アンケート A, アンケート B について、「a. たいてい、何らかのメッセージを残している。」を 5 点、以下 b. を 4 点、c. を 3 点、d. を 2 点、e. を 1 点として評価点の分布を求めた結果を図 5 に示す。

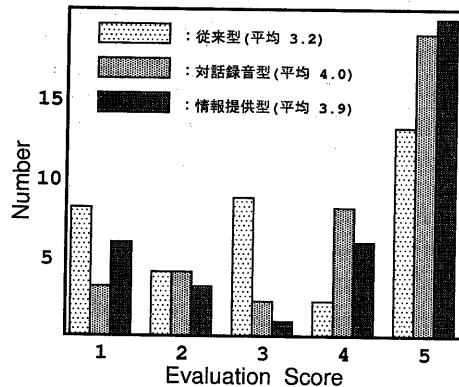


図 5: 各方式の平均評価点

平均評価点は、従来型が 3.2, 対話録音型が 4.0, 情報提供型が 3.9 であった。従って、従来型に比較して、対話録音型の向上分が約 0.8 ポイント、情報提供型の向上分が約 0.6 ポイントであることがわかる。また、それぞれについ

て、 $\chi^2$  二乗分布を用いて評価結果の独立性を検定したところ、危険率5%で分布の差が有意であった。これらの結果から、従来方式に比べメッセージ録音を対話型にする方式、発呼者に興味ある情報を提供する方式のいずれも対話のヒューマンインタフェースを改善できることがわかった。

次に、男女性、年代などによる評価の違いを検証する。男女性別評価結果を図6に示す。

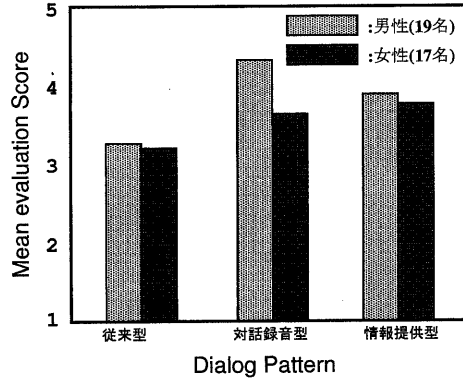


図 6: 男女性別評価結果

男性の場合、対話録音型の評価の向上が女性に比べて著しい。女性は、対話録音型に対する評価の向上が小さい。情報提供型は男女性とも一定の評価の向上を示しているが男女性の差は顕著でない。一方、図7に示すように、年代による大きな差は見られなかった。

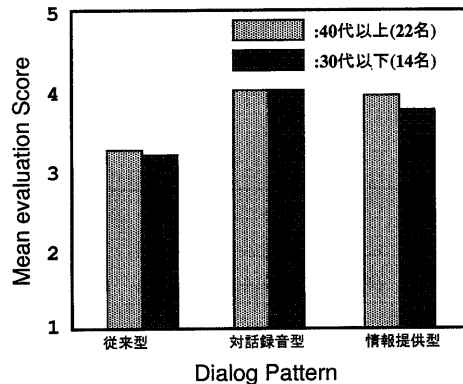


図 7: 年代別評価結果

最後に、伝言を録音せずに切断する理由を質問 (Q2) した結果を図8に示す。

切断の理由としては「c. 留守番電話では用事を済ませることができない」が圧倒的に多いことがわかる。留守番電話は一方的なメッセージの通知のみの処理しかできず、インタラクティブなコミュニケーションは、現状では不可能であるのでやむを得ない面もあるが、ユーザの率直な感想として受け止める必要がある。将来的には、音声認識の他、話者認識、知的なスケジュール管理、セキュリティ機能などを組合せ、発呼者とシステム間のインタラクティブな会話を提供するサービスが期待される。

他の切断理由はいずれもヒューマンインタフェースの問題であり、それらのうちの一部分が本実験の対話録音型、あるいは情報提供型のサービスで改善されることが証明されたと考えられる。

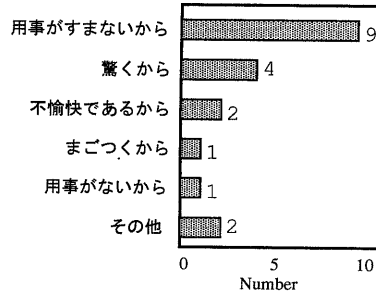


図 8: 切断の理由の調査結果

## 6 まとめと今後の検討課題

従来の一方的な伝言録音方式に比べ、対話型によるメッセージ単位を短くする方式、および発呼者にとって興味を喚起する情報を伝言録音後に聴取させる方式の採用により、メッセージ伝言のヒューマンインタフェースが向上され、平均オピニオン値で0.6～0.8程度改善することができた。今後は発呼者とシステムの知的かつインタラクティブな対話によりサービスの高度化の検討を進めていく。

謝辞 日頃御指導いただき、ヒューマンインタフェース研究所音声情報研究部北脇部長、管村グループリーダーに深謝いたします。

## 参考文献

- [1] 北井, 西: "電話取り次ぎシステムにおける実対話分析", 音講論, 1-7-15(1994.3).
- [2] 北井, 西: "ワードスポッティングを用いた電話音声対話システムにおける対話手順の評価", 信学論 A Vol.J 77-A No.2 pp.251-258, 1994年2月.
- [3] 西, 北井: "尤度と正解率の統計的關係に基づく認識結果確認対話戦略", 信学論 A Vol.J 77-A No.2 pp.259-266, 1994年2月.
- [4] 西, 小島, 五味: "人間機械間の音声対話におけるタイミング", 第2回ヒューマンインタフェースシンポジウム, 2212(1986.10)
- [5] 西, 小島: "電話会話における「相づち」の発言促進効果", 第3回ヒューマンインタフェースシンポジウム, 3132(1987.10)
- [6] 嵯峨山: "なぜ、音声認識は使われないか・どうすれば使われるか?", 音声言語情報処理研究会資料, 94-SLP-1-4, (1994.5).

[付録]

留守番電話・伝言サービスの評価アンケート用紙

(20・30・40・50歳代)，(男性・女性) ←年代と性別に○印をつけて下さい

本アンケートは留守番電話や伝言サービスなどの、音声を録音して相手に伝えるサービスにおいて、ガイダンスの内容が、電話をかけた人の印象にどのように影響するかを評価するためのものです。

以下のアンケートに答えて下さい。該当する回答の□にレ印を記入して下さい。予め次の事前アンケートに記入して下さい。

[事前アンケート]

Q1. あなたは、電話をかけた相手が不在で、代わりに留守番電話が応答した時、どのようにしますか？

- a. たいてい、何らかのメッセージを残している。
- b. 3回中2回くらいは、メッセージを残している。
- c. 2回に1回くらいの割合でメッセージを残している。
- d. 3回に1回くらいはメッセージを残している。
- e. たいてい、メッセージを残さずに切る。

Q2. 上の答えでd. またはe. を選んだ方のみお答え下さい。

メッセージを残さない理由を、上げて上げるならば次のどれですか？

- a. 人が出ると思って電話するのに、いきなり留守番電話が出ると驚いてしまうから。
- b. 用件をうまく整理することができず、まごついているうちに思わず切ってしまうから。
- c. 留守番電話では用事を済ませることができないから。
- d. 機械に向かって話すのは不愉快であるから。
- e. 用事があって電話するのではないから。
- f. その他 ( )

[アンケートA]

2番目の留守番電話は最初の留守番電話と違ってメッセージを少しずつつけて録音する点が異なります。このような留守番電話ならばあなたの留守番電話に対する対応はどうなるでしょうか。次の中から選んで下さい。

- a. たいてい、何らかのメッセージを残そうと思う。
- b. 3回中2回くらいは、メッセージを残そうと思う。
- c. 2回に1回くらいの割合でメッセージを残そうと思う。
- d. 3回に1回くらいはメッセージを残そうと思う。
- e. たいてい、メッセージを残さずに切ると思う。

[アンケートB]

3番目の留守番電話は、録音した人に役に立つ情報を流します。このような留守番電話ならばあなたの留守番電話に対する対応はどうなるでしょうか。次の中から選んで下さい。

- a. たいてい、何らかのメッセージを残そうと思う。
- b. 3回中2回くらいは、メッセージを残そうと思う。
- c. 2回に1回くらいの割合でメッセージを残そうと思う。
- d. 3回に1回くらいはメッセージを残そうと思う。
- e. たいてい、メッセージを残さずに切ると思う。