

オペレータレス自動電話番号検索システムの開発

東田 正信、村上 仁一、奥 雅博
NTT情報通信研究所
{higasida, murakami, oku}@isl.ntt.co.jp

概要

我々はオペレータの介在を必要としないオペレータレス自動電話番号検索システムの研究開発を進めている。本システムでは、利用者からの情報入力端末としてプッシュボタン (PB) 電話機を用いる。利用者はシステムからの音声ガイダンスに従って、電話番号検索に必要な情報 (住所や加入者名義) をPB電話機を用いて入力する。音声ガイダンスは録音編集方式によって生成される。このシステムの最大の特徴は、12個のプッシュボタンだけを用いて日本語を入力する文字情報縮退方式と呼ぶ日本語入力方式を採用したことにある。この方式は従来のプッシュボタンによる日本語入力に比べて約半分の押下数で日本語が入力できるため、利用者の入力に対する負荷を大幅に軽減することができる。

A Fully Automated Directory Assistance System Using Touch-Tone Telephones as Input Devices

Masanobu Higashida, Jin'ichi Murakami and Masahiro Oku
NTT Information and Communication Systems Laboratories

Abstract

We have developed a fully automated telephone directory assistance system that utilizes intelligent interactive information retrieval technologies. Without any operator assistance, customers can access this service directly through conventional touch-tone telephones. The system provides users with stress-free keyword input interface which guides them with recorded speech generated by filling slot-type sentence templates with proper terms. This system is especially characterized by a newly developed input interface that reduces user input workload and time to almost half compared with the conventional coded letter input methods.

1. はじめに

家庭やオフィスに居ながらにして必要な情報を入手することができるオンラインデータベース検索の需要が急速に高まっている。しかし、その利用にはパソコンなどの端末が必要なだけでなく、複雑な操作方法を習得する必要がある。このため、誰でもがどこからでも (特に家庭から) オンラインデータベースを検索して情報を収集することが困難であるのが現状である。

家庭やオフィスから検索して役に立つ情報の1

つに個人や企業、法人の電話番号がある。近年、電話番号もデータベース化され、104番号案内サービスのオペレータが専用端末を用いて検索したり、パソコンや専用端末 (Angel Note) を使って利用者自らが検索できる Angel line と呼ばれるサービスも提供されている。しかし、電話番号のデータベースも上記の理由から家庭やオフィスから容易に利用者が直接アクセスできるものとはなっていない。

そこで我々は、どこからでも容易にアクセス

可能なオペレータレス自動電話番号検索システムの研究開発を進めている。このシステムは、10代から中高年までの広い範囲に受け入れられるように以下のような条件を満たしている：

(1) 端末はPB電話機とする。

PB電話機は日本においてコンピュータ端末の約50倍も普及しており、利用者は家庭やオフィスからこのPB電話機を端末として電話番号検索を行うことができる。

(2) 処理は完全自動とする。

音声認識技術は、街頭やオフィスなどの使用環境を考慮すると、実用的なシステムに適用するには未成熟である。このような状況を踏まえ、PB電話機を入力端末とした使いやすい入力方式（文字情報縮退方式）を考案した。システムと利用者とのやりとりは、知的な対話誘導によって適切に制御される。システムから利用者へ提供される音声ガイダンスには、録音音声と合成音声を適切に組み合わせて用いている。

(3) PB電話機に付いている12個のプッシュボタンのみで入力操作を行う。

このためには、利用者からの入力操作はユーザインタフェースの面から見て改善が必要となる。文字情報縮退方式は、1文字を入力するのにプッシュボタンを1回押下すればよい方式であり、利用者のストレスを軽減した入力方式となっている。

(4) 利用者が所望の結果を得るまでの入力数を最小化する。

入力情報が多かったり、システムとのやりとりが増えると、利用者はシステムとのやりとりに疲れてしまうことが多い。

上記の条件を満足した本システムのプロトタイプは、FDDIによって接続されたクライアント-サーバ(C/S)構成上で動作している。利用者はPB電話機からこのシステムに電話をかけ、電子オペレータとの対話を通して必要な情報を入

力することによって所望の電話番号を検索することができる。

2. オペレータレス自動電話番号検索システムの構成

図1にプロトタイプシステムの構成を示す。入力端末には、家庭やオフィスに広く普及しているプッシュボタン電話機(PB電話機)を用いる。電話網を通してPB電話機からの入力を受け取るシステム本体は、図1に示すようにクライアント/サーバ(C/S)構成をとっており、両者の間は100MbitのFDDIで接続されている。この構成によって、利用者の増加に伴ってシステムの許容量を容易に増やすことができる(スケーラブルな構成)。

(1) サーバ

サーバには、加入者データベースとそのインデックス、録音音声データベース等を載せている。サーバは、クライアントからの要求に従って上記のデータベースを検索し、結果をFDDIを介してクライアントに返す。6,400万加入(企業、法人名義を含む)のデータベース容量は約64GBである。また、検索の高速化のために新たに導入したインデックス(Oku et al. 1997)に対する容量は約20GBである。録音音声データベースは、約1,000文のガイダンス、日本全国の地名をカバーするための部品約20万件、姓名をカバーする部品約20万件、企業名をカバーする部品約8万件から構成される。サーバに搭載されたデータベースの総容量は約100GBである。

(2) クライアント

クライアントは利用者とのやりとり(対話誘導)を受け持つ。利用者からの呼は、クライアント本体に接続された回線制御ボードを通して処理される。このボードは以下の処理を行う：

- (a) 利用者からの呼を受け付ける。
- (b) PB信号を受け取って符号化し、それをクライアント本体に送る。
- (c) クライアント本体からの音声による回答を利用者に送出する。

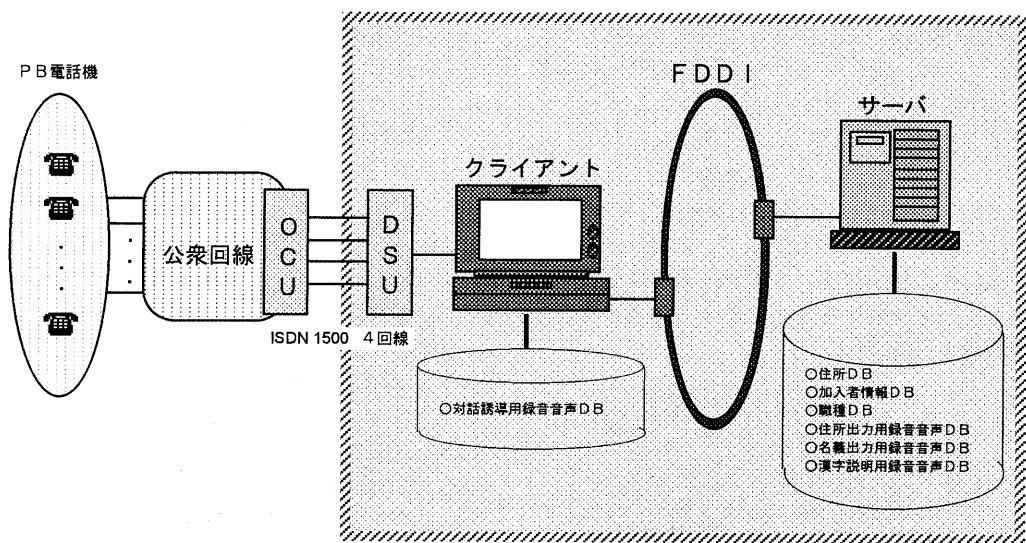


図1：プロトタイプシステムの構成

(d) 通信が終われば呼の完了処理を行う。クライアント本体では、符号化された数字列（利用者による入力）と、現在の対話誘導の場面（住所入力を求めているのか姓名入力を求めているのかなど）とから、サーバ上のどのデータベースをどのような条件で検索すべきかを決定し、その結果をSQL化してサーバに送る。また、サーバからの検索結果に対して、利用者にもどのように回答すべきかを決定し、適切な音声ガイダンスを選択して利用者へ回答する。

(3) 入力端末としてのPB電話機

PB電話機の12個のプッシュボタンによって日本語を入力するために、文字情報縮退方式と呼ぶ従来の入力方式に比較して利用者のストレスの少ない入力方式を考案した。この入力方式の最大の特徴は、1文字を入力するのに1回のプッシュボタン押下で済む点にある。

3. 入力方式

住所や姓名などの情報をPB電話機を用いてどのように入力するのかという問題は、利用者にとっての使いやすさという観点から非常に重要

な問題である。

電話番号検索に必要な情報は、住所と加入者名義（姓名、企業名、法人名）である。アルファベットを使用している国においては、これらの情報を入力するのに27文字のキーがあればよい。これに対して、日本では71文字のひらがなに加えて2,000文字以上の漢字が使われることから、PB電話機を用いた漢字までを含めた実用的な入力方式は確立されていない。かなレベルの入力方式としては、文字循環指定方式、文字コード化方式が用いられており、単語そのものにコードを割り当てる単語コード化方式も用いられている。

3. 1 従来の日本語入力方式

PB電話機、携帯電話、PHSなどの既存の電話機の持つ12個のプッシュボタンを用いて日本語を入力する方式がいくつか提案されている。本節ではこれらの方式について概観する。

図2に12個のプッシュボタンへのデフォルトのひらがな割当てを示す。これにはITU勧告 (ITU 1981) に基づくアルファベットの割当ても含まれ

ている¹。

図2に示すように、アルファベットの割当てと同じように、5つのひらがなが1つのプッシュボタンに割り当てられている。“1”のプッシュボタンには5つの母音（“あ”、“い”、“う”、“え”、“お”）が割り当てられ、他のプッシュボタンには同じ子音を持つ5つのひらがな（“2”にはカ行の5文字、など）が割り当てられている。濁音、半濁音は元の清音に加えて“*”を押下することによって入力することができる（図2参照）²。

1 あいうえお ー: 長音	2 かきくけこ ABC	3 さしせそ DEF
4 たちつてと GHI	5 なにぬねの JKL	6 はひふへほ MNO
7 まみむめも PRS	8 やゆよ TUV	9 らりるれろ WXY
* 濁音 (半)濁音	0 わん QZ	# スペース# 終了##

図2：PBへの文字の割り当て例

(1) 文字循環指定方式

この入力方式は携帯電話において広く用いられている。利用者が個人的な電話帳を作成するときに必要な住所や姓名を入力する際に用いられている。この方式においては、入力モードに切り替えた後にプッシュボタンを押下するとア段のひらがなが表示される。同じプッシュボタンをもう1度押すことによってイ段のひらがなが表示される。同様にして同じプッシュボタンを6度押せば元のア段のひらかなに戻る。例えば、“と”を入力する際には、タ行であるプッシュボタン“4”を5回押下する。

(2) 文字コード化方式

この方式はポケベルやPHSに採用されている

¹ ITU勧告はその後、Qの割当てが“0”から“7”へ、Zの割当てが“0”から“9”へと変更されている。

² この方式は必ずしも標準的に使用されているわけではなく、一般には他のファンクションキーを用いることが多い。

方式であり、若者の間で友人にメッセージを送るのに幅広く利用されている。この方式では、71のひらがな、アルファベット、数字、およびいくつかのコントロール文字をすべて2桁の数字にコード化してある。例えば、“あ”は“11”、“い”は“12”で、“と”は“45”である。この方式を用いれば、ひらがなで簡単にメッセージを作ることができるが、1つのひらがな文字を入力するのに2桁の数字をプッシュボタンで押さなければならない。このため、メッセージを作るのに若干の時間が必要となる。

(3) 単語コード化方式

この方式は電話を介したテレフォンショッピングやチケット予約で利用されている。この方式では、個々の単語や商品、飛行機の便名、列車などに特有のコード（数字列）を割り振っている。従って、この方式を利用するには常にどのコードがどの単語や商品などに対応するのかを記載したコード表を参照しなければならない。

PB電話機を端末として全国の電話番号検索を行うシステムの入力方式としての観点から上記の入力方式を評価すると以下ようになる：

- 単語コード化方式は全国の住所や姓名に対するコードを知るためにコード表を持ち歩く必要があることから最も効率の悪い方式である。
- 文字循環指定方式も、たとえプッシュボタンを正しい回数押してもそれが正しい文字列になるとは限らない点で効率が悪い方式である（例：“111”→“あああ”、“あい”、“いあ”、“う”の4通りの解釈が可能）³。
- 文字コード化方式は、文字コードを覚えることをいとわなければ、他の2つの方式に比べて実用性が高いと考えられる。しかし、文字循環指定方式も文字コード化方式もともに番号案内システムに正しい情報を与えようとする

³ 実際には、これらの区別を行うために、1の間にポーズを置く方法を採用している機種が多い（例：“11”ポーズ“1”=“いあ”）。

と、利用者の入力に対する負担がかなりのものとなることは避けられない。

○上記(1)～(3)のすべての方式に共通する問題として、情報を入力する際に誤る可能性が高いため、利用者の求める情報をうまく見つけられなくなることが挙げられる。

3.2 新しい日本語入力方式(文字情報縮退方式)

3.1節で述べた3方式に比べて利用者のストレスが少ない、より使い勝手のよい入力方式である文字情報縮退方式を考案した。この方式は、利用者の要求に一致する情報をデータベースから見つけたときに特に有効性を発揮するが、この方式はプッシュボタンの押下数が少なく済むため、ヒューマンマシンインタラクションを改善することができる。

本稿で提案する文字情報縮退方式は、1文字を入力するのにプッシュボタンを1回押下する。例えば、“とうきょう”と入力するためには、プッシュボタンを使って“4 1 2 8 1”と数字を5回押下する。プッシュボタン押下数を他の方式と比較すると、文字循環指定方式では“4 4 4 4 4 - 1 1 1 - 2 2 - 8 8 8 - 1 1 1”の16回、文字コード化方式では“4 5 - 1 3 - 2 2 - 8 3 - 1 3”の10回となり、文字情報縮退方式の5回に比べて非常に多いことがわかる。

表1に各方式の平均プッシュボタン押下数とそれに要する時間を示す。表1から明らかのように、文字情報縮退方式は他の方式に比較して、情報を入力するための平均プッシュボタン押下数が少なく、その入力時間も短い。従って、文字情報縮退方式は、使い勝手がよく、利用者を与えるストレスも少ない方式であると言える。

表1：入力方式の比較

入力方式	1文字当たりの平均プッシュボタン入力数	1文字当たりの平均入力時間 [秒]
文字循環方式	2.6	1.9
文字コード化方式	2.0	2.2
文字情報縮退方式(提案方式)	1.1	1.3

文字情報縮退方式はプッシュボタン押下数が少なく入力しやすい方式ではあるが、母音情報の入力を省略した方式であるため、プッシュボタンにより押下したそれぞれの数字がどの段(ア段、イ段など)の文字であるかが曖昧なままである。数字列の長さをNとすると、その曖昧さの数は 5^N となる。この曖昧さをDBを参照しながら解消する。

英単語の入力方式として文字情報縮退方式と同様な方式が提案されている(Kondraske et al. 1986)。電話番号案内で用いる単語は、住所や姓名というように限られた集合に収まるので、(Kondraske et al. 1986)と同じ考え方が適用できる。(Kondraske et al. 1986)で用いられている英単語は非常に狭い範囲に限られているのに対して、我々の場合には住所や姓名というように限られているとはいえ、かなり大きな範囲になる。このため、与えられた数字列に対応して多くの住所や姓名が一致することにより、短時間にそれらのうちのどれが利用者の求めるものであるのかを見いだすことが難しくなる可能性がある。

3.3 エントロピーの計算

文字情報縮退方式による入力が有用であるか否かは、システムや利用者に対する負荷を減らせるか、大規模なデータベースから目的とする市や町の名前や加入者名義を特定することができるかを評価することによって明らかにすることができる。この評価を行うために、我々の扱う環境(住所や姓名の入力)に対するエントロピーを計算する。言語そのものが持つエントロピーに関する研究は数多い(Brown et al. 1992)(Nagao et al. 1994)が、住所や姓名のような限られた範囲から成る環境におけるエントロピーに関する研究はない。

エントロピー計算の結果を表2に示す。表2の結果は、住所に関しては、71文字のひらがなを11個の数記号(10個の数字と1つの記号)にコード化することによって失われる情報は非常に少ないことを示している。また、姓名に関しては住所に比べて失われる情報は若干大きい、やはり失われる情報は少ない。これらのことは、上記のコード化を行っても、住所や姓名に関し

ては単語間の重なり（縮退）があまり大きくないという事実を示している。このように、住所や姓名のように入力される単語の範囲を限れば、文字情報縮退方式による入力の曖昧さの解消は、単純に計算した曖昧さ（ 5^N ）の解消に比べて困難ではない。すなわち、文字情報縮退方式による住所や姓名の入力が有用であることを示している。

表2：住所、姓名に対するエントロピー計算結果

住所レベル	件数	ひらがな表現 [ビット]	数字列表現 [ビット]	減少量 [ビット]
都道府県	47	5.55	5.20	0.35
市区郡	3,722	11.87	11.76	0.11
町大字	371,221	18.10	17.94	0.16
姓、名	157,377	12.76	11.15	1.61

4. 対話誘導

4.1 基本的な考え方

文字情報縮退方式に適した対話誘導では、利用者が求める解を検索するために必要な情報を得ると共に、上記の方式自体に起因する曖昧さの解消も行う必要がある。また、電話番号案内のオペレータは単純にデータベース検索の仲立ちをしているのではなく、利用者からの入力に対して利用者の求める解に効率的に行き着くように知的作業を行っている (Muller et al. 1995)。これらのことを考慮して、利用者に対する対話誘導を以下の4つの戦略を基本に組み立てた：

(戦略1) 利用者の入力が極力最小となるようにする。

(戦略2) 利用者システムに理解に曖昧さがあることを意識させないような応答を行う。

なぜなら、利用者はある文字列を意識して入力を行っており、システムがそれを同定できないとは考えていないからである。

(戦略3) 利用者負担とならない（要求されて当然であると思わせる）情報についての質問を行う。

戦略2を満足するために、曖昧さ解消のためだけの情報を求めることは利用者の負担となってしまうので避けなければならない。

(戦略4) 1度要求した情報、既知の情報から決定できる情報については質問しない。

4.2 対話誘導の例

図3に文字情報縮退方式による入力に対して、検索処理の中でどのように曖昧さを解消していくかの対話誘導例を示す。ここで、yes/no 質問はよく使われるので、yesは“1”に、noは“0”にそれぞれ割り当てられている。

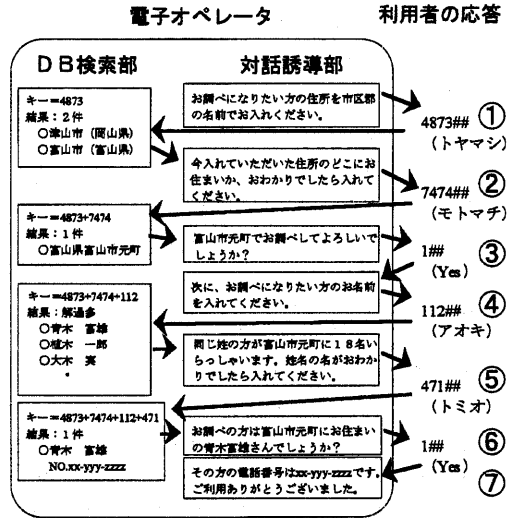


図3：対話誘導の例

電子オペレータは図3に示す対話誘導部において、利用者とのやりとりを音声ガイダンスによって行う。どのようなやりとりを行うか、どのように音声ガイダンスで利用者に指示するかはDB検索部の検索結果によって決定される。この音声ガイダンスは録音編集方式によって生成される。録音編集方式では、基本的にはスロットを持つプレート文に対して、スロットの部分にデータベースを検索した結果（住所や姓名など）を埋め込むことによって最終的な音声ガイダンスを生成する (Higashida et al. 1991) (Ouchi et al. 1992)。すべての住所とほとんどの姓名についてはあらかじめ録音しておき、これらを録音音声データベースに登録している。利用者の入力に対するデータベース検索結果がこの録音音声データベースに登録されていない場合には、合成音声によってスロットを埋める。

以下に図3に示すシステムからの応答が4.1節のどの戦略に基づいているかについて述べる：

- ①最初に市区郡を尋ねる。市区郡が決まれば都道府県も決まるため、都道府県を聞く必要はない(戦略1)(戦略3)(戦略4)。
- ②文字情報縮退方式に起因する曖昧さを解消するために、さらに下位の住所の入力を求める(戦略2)(戦略3)。
- ③調べる住所が1件になると、その住所が正しいか否かを利用者に確認する。
- ④電話番号を調べるのに必要な名前の入力を要求する(戦略3)。
- ⑤検索解が多く、特定できないので、さらなる情報(この例の場合、姓名の名)の入力を求める(戦略2)(戦略3)。
- ⑥検索解が利用者の求めている加入者であるか否かを住所、姓名を用いて確認する。
- ⑦特定できた加入者の電話番号を案内する。

5. 評価

プロトタイプシステムに対して、複数人のモニタにアンケートに答えてもらった。評価は、パフォーマンスとユーザインタフェースの2つの観点から行った。

5.1 パフォーマンス

表3に個人の電話番号を入力情報の条件を変えて検索したときの平均的なパフォーマンスを示す。表3は、利用者からの入力情報を【場合1】～【場合4】の4種類に分けて、それぞれの場合に対する正確さと接続時間とを示している。

表3：パフォーマンス評価結果

場合	入力情報の条件	正確さ [%]	接続時間 [秒]
場合1	(市区郡+町大字) & (姓+名)	99.0	62
場合2	(市区郡) & (姓+名)	85.0	71
場合3	(市区郡+町大字) & (姓)	65.9	99
場合4	(市区郡) & (姓)	16.5	120 ~180

(1) 正確さ

ここでは正確さを、「利用者が適切な情報を入力した後で、利用者が望む電話番号を案内できた呼の全呼に対する割合」と定義する。

この正確さは、利用者がどのような情報を持っているかに大きく依存する。表3より、利用者が市と町レベルの住所と姓名とを正しく入力すれば、ほぼ加入者を一人に特定でき、利用者の所望する電話番号を案内することができる【場合1】。利用者の入力できる情報が少なくなると、その情報の重要さに応じて案内できる割合が減少する。例えば、利用者が市レベルの住所と姓の情報しか持っていない【場合4】では、ほとんどの場合、正しく電話番号を案内することはできない(表3では16%程度の正確さしかなく、加入者を特定することができない)。

(2) 接続時間

利用者とのやりとりを完了するまでに要する時間は利用者が入力した情報に依存する。表3の【場合2】(市レベルの住所と姓名を入力)での時間が【場合3】(市と町レベルの住所と姓を入力)よりも短いのは、姓名の名の情報の方が町レベルの住所の情報よりも重要であるからと考えられる。また、利用者が市レベルの住所と姓しか知らない【場合4】では、その姓が市レベルの住所で特定できるようなまれな場合を除いて、非常に時間がかかることがわかる。

企業を対象とした検索での正確さは利用者が企業名などの情報をすべて知っていても約78%であった。時間に関しては個人とほぼ同じ時間で検索を完了することができる。

5.2 ユーザインタフェース

アンケート結果を見ると、本システムのユーザインタフェースはかなり効率的であると考えられる。アンケート結果からは、予期していた以上に新しい入力方式(文字情報縮退方式)が受け入れられやすいことが読みとれる。このことから、文字情報縮退方式は電話番号案内サービスを行うための簡単で便利な入力方式であると言える。すなわち、電子オペレータとの対話を通して情報を得ていくという処理が、利用者が感じるであろう入力方式に対するストレスを低減させていると考えられる。

6. おわりに

本稿ではオペレータレス自動電話番号検索システムの概要について述べた。このシステムは、PB電話機を入力端末として使用することによって、家庭やオフィスからの電話番号検索を可能とする。システム実現にあたって、利用者の入力に対するストレスを軽減するために、文字情報縮退方式と呼ぶ新たな入力方式を考案し、さらに入力に伴う曖昧さを対話誘導によって解消していく技術を開発した。評価を通して、文字情報縮退方式の有効性および本システムの有効性を確認することができた。

なお、文字情報縮退方式によって入力すべき情報を数字列に縮退させた集合の有するエントロピーが、元の言語の持つエントロピーに対してあまり減少しなければ、PB電話機を用いた他のトランザクション処理（テレホンショッピングやチケット予約など）に本システムで開発した技術を適用することが可能である。

【参考文献】

- (Brown et al. 1992) Peter F. Brown et al., "An estimate of an Upper Bound for the Entropy of English", *Computational Linguistics*, Vol. 18, No. 1, pp.31-40 (1992).
- (Higashida et al. 1991) Masanobu Higashida et al., "A Text Generation Expert System for Real-Time Information Delivery", *Proc. of The World Congress on Expert Systems*, Vol. 4, pp.2326-2333 (1991).
- (ITU 1981) CCITT, ITU, "International telephone service operation recommendations E.100-E.123", Vol.2. - FASCICLE 2.2 YELLOW BOOK, 7th plenary assembly, Geneva (1981).
- (Kondraske et al. 1986) George V. Kondraske and Adanan Shennib, "Character Pattern Recognition for a Telecommunication Aid for the Deaf", *IEEE Tr. on Biomedical Engineering*, Vol. BME-33, No. 3, pp.366-370 (1986).
- (Muller et al. 1995) M. J. Muller et al., "Telephone Operators as Knowledge Workers: Consultants Who Meet Customer Needs", *CHI'95*, pp.130-137 (1995).
- (Nagao et al. 1994) Makoto Nagao et al. "A New Method of N-gram Statistics for Large Number of n and Automatic Extraction of Words from Large Text Data of Japanese", *Proc. of COLING-94*, pp.611-615 (1994).
- (Ouchi et al. 1992) Yukio Ouchi et al., "A New Traffic Information Service Delivered over the Cellular Mobile Phone Network", *Proc. of PTC'92*, pp.141-147 (1992).
- (Oku et al. 1997) 奥 雅博、野田良輔、林 智定、"形態素解析を用いた中間部分一致検索の高速化手法"、*情報処理学会研究会報告 97-NL-121-9*, pp.61-68 (1997).