

自由記述型アンケート回答を対象にした要求特定手続きの提案

Instruction to Submit a Paper to Proceedings of IBS

乾 裕子* 丸元 聡子** 井佐原 均*

By Hiroko INUI, Satoko Marumoto and Hitoshi ISAHARA

1. はじめに

筆者らは、柔軟かつ確実な方法で、より多種多様な意見を収集蓄積し、これらの意見から発言者の意図を適切に取り出すことに関心がある。本研究では、意見収集の一方法としてのアンケート調査に着目し、図-1に示したように自由記述アンケートの回答テキストから回答者の意図を自動的に取り出し、従来の自由回答分析における分類支援を行う研究を進めている¹⁾²⁾³⁾。

アンケート調査の回答形式には、回答を選択肢として予め用意しておく選択型回答と、回答者に自由記述をさせる自由回答とがある。選択型アンケートの調査計画・質問紙の作成方法・実施法・コーディング(分類)法については様々な知見が蓄積されている。一方、自由回答のコーディングには多大なコストがかかるうえに、人の判断による作業は主観的

な分類結果を招くという懸念もある⁴⁾。また、回答を回収しても解析されないまま終わることが多いことも指摘されている⁵⁾。したがって、回答収集後の解析のコストを避けるために、選択型のアンケートを行うことが多い。自由回答は、従来、選択型アンケートを行うための予備調査として小規模に実施する、あるいは選択型アンケートの中で調査者が想定できなかった選択項目、例えば選択肢以外の「その他」に相当する回答と位置付けられていた。

しかし、近年、インターネットの普及やパブリック・インボルブメントに対する関心の高まりから、想定できる意見を選んでもらうのではなく、回答者の自由な個々の意見を積極的に聞くことが重視されている⁶⁾。その結果、意見の集約方法の一つとして、自由回答が選択型アンケートと同様に大規模に実施されるようになってきている⁷⁾。また、狭義にはアンケート調査によって得られる自由回答とは異なる

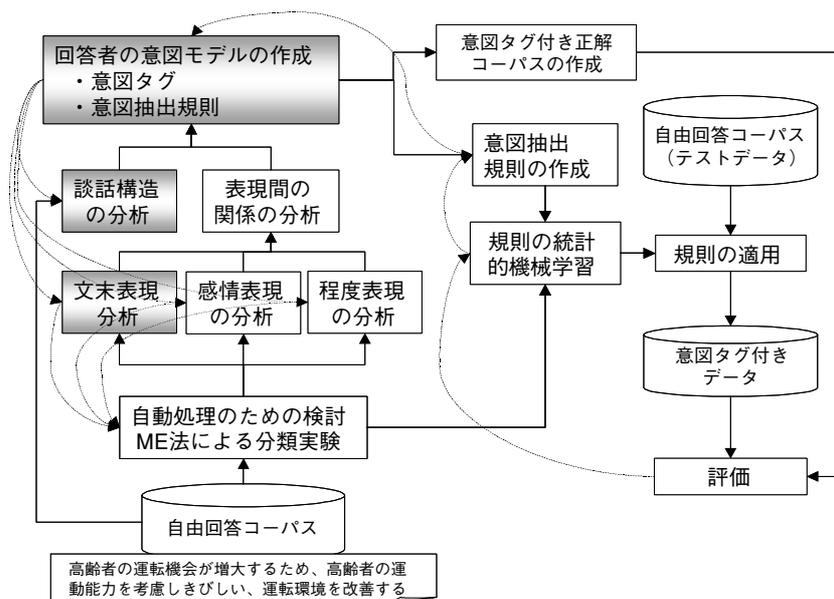


図-1 自由回答アンケートからの意図抽出処理アプローチ

* 通信総合研究所・神戸大学大学院 ** 言語情報研究室

が、企業のホームページの掲示板やコールセンターなどに寄せられる消費者のメールや意見、地方自治体や政府のホームページに集まる住民からのメールは、自由回答同様に意見集約の対象とみなすことができるが、これらの記述されたテキストや発言から、ユーザの志向や消費動向などを発見する研究⁸⁾、質問に対する回答を自動生成するための研究⁹⁾、テレビ番組の視聴者の印象¹⁰⁾や、道路利用者の不満の感情¹¹⁾を取り出す研究も行われている。

しかし、一方で、意見集約の方法として、アンケート調査から回答者の関心や懸念を掘り下げることは難しいとの指摘もある⁶⁾。

以上を踏まえて、本研究のねらいを次に述べる。第一に自由回答のコーディング過程に自然言語処理の要素技術を取り入れることにより、人手作業のコストを軽減し、意見集約の対象データとして、自由回答に記述された意見を活用することである。テキストからの情報抽出や、要約・自動分類などの要素技術が蓄積されてきている言語処理技術を用いれば上記の問題を解決できる可能性がある。これに基づいて、テキスト処理における解釈のための知識データを作成する。

第二に、機械的な処理が可能であるか不可能であるかを問わず、回答者の要求を認識・理解するために、読み手（調査者）が、1) どのような知識構造を用意すべきか、2) 知識構造を構成する要素は何か、3) 各要素は回答からどのように発見できるかを明らかにする（図-1）。さらに、矢嶋の指摘⁶⁾にもあるように関心や懸念を掘り下げるため「明示されていない回答者の意図」にも着目する。

本稿では、回答には回答者の要求意図を示す手がかりが必ず含まれるという仮説を前提とすることにより、回答の要求内容および動機・根拠と、焦点である要求意図との間の有機的構造と推論関係を明らかにする。このため、回答の要求意図、内容、動機について、およびこれらの要素間の関連について回答者の意図がどのように現れているか、また表現として現れていないかという観点から分析を行う。分析で得られた知見をもとに、この構造と関係を記述する方法として要求特定のための知識構造を提案する。

2. コミュニケーションとしてのアンケート調査

(1) アンケート調査と自由記述回答の特徴

アンケート調査は、従来から心理学・社会学の分野においてデータの収集方法として発達し、社会的にも市場調査や意識調査などで多用されてきた。一般に「質問を行い」「回答としてデータを収集し」「解析することで役立つ情報を引き出す」といった表現で定義されるように³⁾、アンケート調査はしばしば、調査者の一方的な行為とみなされている。しかし、本来「質問を行う」という行為の反応として「回答する」という回答者の行為があり、その行為に対して調査者は「データを収集し解析する」という反応を返している。さらに、アンケートの目的や種類によっては、解析した結果を調査側が回答者に対して公開し、それにより回答者は自分の意見や考えがどのように位置付けられたかを知るといった段階を想定することもできる。すなわち、アンケート調査とは調査者と回答者との間のコミュニケーションの一種と捉えることができる。

アンケート調査の方法論は、質問紙調査法としてさまざまな種別・手法が研究考案されてきた。自由回答は一般に自由記述形式とも呼ばれる回答形式のひとつである。自由回答には、回答しにくいために無回答が増える、的外れの回答がある、コーディングに時間がかかるといった欠点⁴⁾⁵⁾や、結果の解釈に関して調査者の主観が入ること⁵⁾から、心理学や社会学におけるアンケート調査では対象外にされる傾向にあった。

(2) 他者想定型テキストとしてのアンケート回答

本研究では、従来の自由記述型アンケート回答の分析・分類における、1) 情報抽出技術などを単純に適用することで、特定の要素のみに着目してしまうため自由記述の長所を活用できない、逆に、2) 回答からどのような情報を取り出すかの焦点化が十分でない、という問題点を解消するため、次のような研究方針を掲げる。それは、先に述べたようにアンケートを調査者と回答者のインタラクションとみなすことにより、アンケート回答を他者想定型テキストと捉えることである。

他者想定型テキストとは、新聞などの一方向型テキストと異なり、テキスト中に「読み手への働きかけ」を含む。このように考えることにより、質問票

における「21世紀の道をどのように考えますか」という設問と、それに対する「環境に配慮することが重要である」といった回答との対を、対話における話し手と聞き手のやりとりとみなすことができ、個々の回答を聞き手（調査者）の信念や意図を変化させる行為と考えることができる。一方、調査者側から見た場合、上記の回答例は「環境に配慮する」という要求内容と「環境に配慮してほしい」という要求態度から成ると解釈できる。これらの考え方は、言語行為論²⁾の命題内容 (prepositional content) と発語内の力 (illocutionary force) として意味論、語用論で議論されてきた。

とくに、新製品の開発、あるいは地域開発・将来の事業計画等の意見集約におけるアンケートの場合には、この発話行為の多くが要求であると推測できる。本研究で扱う将来の道路計画に関するアンケートの自由記述においても同様の傾向を観察することができる。このため、本稿においては回答者の意図の中で「要求」に着目している。同時に「読み手への働きかけ」である以上、読み手である調査者は、回答者の意図を適切に汲み取り、要求は何か、なぜそのような要求がなされたかを判定できなければならない。そのため、要求特定手続きのための知識構造を明らかにする必要がある。

(3) 分析対象データ

本研究で分析対象としたデータは、「21世紀の道を考える委員会」が平成8年5月から7月末に実施したもので、将来的な道路計画に市民の声を活かす目的で行われている⁷⁾。回答者数は35,674人、回答件数は延べ113,316件である。意見は、ハガキ、封書、FAX、電子メールによる回答の他、ホームページへの書き込みによって集められている。回答方法は、あらかじめ設定された道づくりに関する12個のテーマ（渋滞の解消、交通安全の確保、情報通信技術と交通など）の中から、回答者が各自関心の高いテーマを選択する。各テーマに対し、4個程度の参考意見およびグラフや図などの参考資料が提示されている。参考資料をもとに、120字程度の文字を記入できる回答欄に自由に記述し、書ききれない場合には別紙に記入する。

3. 要求の分析

ここでは、まず本研究における分析の方針と概要について述べ、次に事例を挙げて具体的に分析過程を説明する。

(1) 分析方針

アンケート調査が調査者と回答者のインタラクティブな行為であることに着目し、プランニング研究の枠組みである信念欲求モデルを、要求を理解するための知識構造として整形する。この知識構造を用いて、明示されていない要求内容や動機を要求意図からアブダクティブに推論する方法を提案する。要求特定の知識構造は、意図が信念と欲求から生成される一般的な信念欲求モデルを利用して、下記のように記述することができる。

回答者の要求特定の知識構造

= 要求動機 (信念 1, 信念 2, …信念 n)
 + 要求内容 (内容 1, 内容 2, …内容 n)
 + 要求意図

構成要素である要求意図、要求内容、要求動機を特定するにあたって、理論上、各要素がすべて表現されている場合から、いずれか二要素がある場合、一要素だけある場合、すべて表現されていない場合まで8通りの組み合わせが考えられる（表-1）。また、表の要求意図の有無に着目すると、a~dを明示的な要求意図、e~hを暗黙的な要求意図と区別できると考えられる。

上記を仮説として、実データのアンケート回答を対象にこれらのパタンの違いが回答中にどのように現れるか、また、現れていない要素が他の要素から特定できるかどうか分析する。それぞれの違いについては後で詳しく述べる。

表-1 作業仮説：回答に現れうる要求要素のパタン

要求要素のパタン	要求意図	要求内容	要求動機	要素の上位範疇
a	○	○	○	明示的要求意図
b	○	○	×	
c	○	×	×	
d	○	×	○	
e	×	○	○	暗黙的要求意図 あるいは 要求でない場合
f	×	×	○	
g	×	○	×	
h	×	×	×	

(2) 分析作業手順

ここでは、分析の手順を決め、これに基づいて<要求分析データベース>を作成する。以下の手続きにしたがって、分析を進めながら項目列を増やすという発見的方法によって作成している。最終的に構築されるデータベースの項目は、要求特定の知識構造における記述項目でもある。

要求分析およびデータベース構築の作業手順

- 1) 要求意図を取り出す
 - 1 A) 回答文に直接要求表現（てほしい、ください等）が含まれるか否か調べる
 - 1 B) 回答文の文末を「てほしい」という直接要求表現、もしくは「～ことを要望する」という表現に言い換えることが可能か否か調べる
- 2) 回答の1以外の部分を要求内容とみなす
- 3) 要求内容の中に、内容の動機に相当する要素があれば区別する（原因・理由、動機を示す接続助詞などに着目。それ以外は目視により意味から判断する）
- 4) 要求内容を異なる内容ごと（節単位）に分割し、内容間、および内容と動機との順序関係（因果関係含む）を記述する
- 5) 順序関係を記述する際、内容や動機が回答に明確に記述されていなくても、前後関係からこれらの各要素が十分推論可能な場合には、分析者が適宜補完する
- 6) 内容や動機の中に、回答者の感情や信念に対する判断・評価などが現れていれば記述する
- 7) 動機が複数の信念から成っていれば分け、それらの間の関係を記述する
- 8) 7のうち、要求の動機に直接結びつかない信念は区別する

分析にあたってまず、アンケート回答の各事例を対象に要求意図の表層的な有無を区別する（手順1）。次に手順2、3では要求の内容と動機を特定する。ここまでの作業を行った段階で、要求意図分析データベースの基本項目列が整う。

この基本の骨組みをもとに、例えば、要求の動機となる信念が形成されるプロセス（手順5）や、要求の動機が複数の信念から構成されている場合には各信念間の関係（手順7）、また、要求には直接結びつかない信念（手順8）などを、必要に応じて基

本項目列と区別しながら項目列を追加している。

4. 分析結果および考察

ここでは、分析結果と分析過程で得られた知見について考察する。

作業手順1 AとBによる要求意図判定を行った結果、任意の約3000事例に対して表-2の結果が得られた。

表-2 明示的要求と暗黙的要求

要求要素のパタン	要素の上位範疇	直接要求表現	言い換え	3000文中
① a~d	明示的	含む	可	547
② —	—	含む	不可	3
③ a~d	明示的	含まない	可	1190
④ e~h	暗黙的	含まない	不可	1252

注：①~④に含まれない8文については、前後文脈を見ないと判断に迷う例。②の要求表現は引用文中に現れているため対象外とする

手順の1 AおよびBはいずれも表層的な情報を手がかりとはしているものの、本研究が目指すのは表層表現のみからの意図抽出ではない。直接的な要求表現を含むか否か（手順1 A）、要求に相当する表現を含むか否か（手順1 B）、この組み合わせから、少なくとも表-2に示した四つの型があり得る。

直接的な要求表現を含むにも関わらず、要求意図として言い換えができないというのは矛盾しているため実質的には②はありえない。したがって、データとして得られるのは、①・③・④の三つの型である。しかし、④については要求でない意見も含まれている可能性がある。われわれは、この型についても分析を深め、どのような意見と解釈すればよいのかを検討する。表層情報だけを手がかりにするならば、④の型を対象とする必要はないが、回答文において約40%を占めるこのタイプの意見をわれわれは重要であると考えている。

表-1の8パタンのうち、要求要素のパタンCのみ以外は実際にデータとして現れた。これら残りの7パタンでは、1) 要求意図の明確な事例の場合には、1 a) 要求動機から内容を、1 b) 要求内容から動機を、2) 要求意図の不明な事例の場合には、2 a) 動機から内容と要求意図を、2 b) 内容から動機と要求意図を、それぞれ特定できることが判明した。

本研究で使用した回答テキストのアンケート調査は、先に述べたようにあらかじめ設定された道づくりに関する12個のテーマがあり、回答者はそれにしたがって回答することができる。同時にこれは分析者（あるいは調査者）にとって、要求内容が明示されていなくても回答者がテーマを選択した段階で何について「ちゃんとしてほしい」と述べられたのかを推論できるということである。したがって、このタイプの場合、要求内容(C)などを補完するためには、質問紙の設計を考慮する必要がある。

d) 要求意図(I), 動機(R) → 内容(C)

#6 週末になると、小倉競馬場周辺の車の渋滞が
はげしいので(R)、なんとかしてほしい(I)

#7 道路が凸凹になって水たまりができて
いる(R)、ちゃんとしてほしい(I)

上記に見られるように、「何に対して」「何をするのか」といった要求内容が明示的でない例がある。このような場合、aの#2の例と同様に、要求動機に現れている既出の対象（「小倉競馬場周辺の車の渋滞」「道路の凸凹」）を省略補完する。さらに「渋滞をなんとかする／道路の凸凹をちゃんとする」から「渋滞を解消する／道路の凸凹をなくす」を推論するという二つの段階を踏むことにより、要求内容を特定できる。

以上の分析を踏まえ、明示的要求意図における特定のための知識と規則を下記のようにまとめる。

記号の定義：

Demand：調査者が理解すべき回答者の要求，I：要求意図，C：要求内容，R：要求動機（r：信念），φ：空きスロット

要求および要素特定のパターン：

a) $R(r_1, r_2 \dots r_n) + C(c_1, c_2 \dots c_n) + I(i) = Demand$

但し、C中の行為対象はRから省略補完する場合がある

b) $R(\phi) + C(c_1, c_2 \dots c_n) + I(i) = Demand$

但し、NOT(C)に対するネガティブな評価からRをアブダクティブに特定する

c) $R(\phi) + C(\phi) + I(i) \neq Demand$

但し、質問紙の設計に依存してCを特定できる場合がある

d) $R(r_1, r_2 \dots r_n) + C(\phi) + I(i) = Demand$

但し、C中の行為対象をRから省略補完しC'を形成する場合がある。さらにC'からCをアブダ

クティブに特定する

(2) 暗黙的要求

ここでは、直接要求表現を含まず、かつ言い換え不可の文を対象としている。

e) 内容(C), 動機(R) → 要求意図(I)

#8 高速道での事故対面交通箇所通過時に「ぶつかる」という不安を感じます(R)。「八代～人吉間、同じ料金を支払うのであれば早期解消を」(C)

これは、二文からなる一回答中、前文が要求の動機を、後文が内容を示している例である。この事例で「早期解消を（してほしい）」という要求意図が特定できるのは、格助詞ヲで終わる表現的特徴に依存するところが大きい。現在のところ、このパターンにおいて動機と内容からアブダクティブに要求意図が特定できる例は見つかっていない。

f) 動機(R) → 要求意図(I), 内容(C)

#9 歩道がせまい(R)。

これは、#4の解釈プロセスを逆方向に辿る、すなわち「歩道が狭い→（歩道が狭いのは危険である）→（危険は嫌だ）→（広くしてほしい）」という推論プロセスにより要求意図と内容が特定できる事例である。このように要求意図特定手続きにおける推論では双方向の推論が可能である。

g) 内容(C) → 要求意図(I), 動機(R)

#10 高速道路の料金はJHが決めているが、もっと経済原則が反映するような決め方、市場原理が働くような値決め方式があるのではないか。

この事例は、要求意図と動機だけでなく正確に言えば要求内容自体も下記に示す推論プロセスを経ることによって明らかになるものである。これについて、詳しく説明する。

1. もっと経済原則が反映するような決め方、市場原理が働くような値決め方式があるのではないか
2. (いや、あるはずだ(x))
3. (しかし、実際には存在しない(r1→R))
4. (r1がないのは嫌だ)
5. (よって、r1の状態を要求する)
6. (すなわち、もっと経済原則が反映するような決め方、市場原理が働くような値決め方式を作っ(C)てほしい(I))

上記の推論プロセスでは、1～2の間でいわゆる反語表現の解釈を行っている。この補完によって、

#9における信念 x を決めることができ、また、この信念 x により x の否定である信念 $r1$ を決めることができる。これによってパターン b の推論プロセスを経ることができるため、上記3や6に見られるような動機や内容の役割を担うことが明らかになる。推論により、 x の否定が動機であり、 x が結果的に内容となる。最終的に現れる内容が表現として明示されているという点から、この事例を g のパターンと位置付けている。

#11 国道 200 号線の渋滞

この例では、要求内容の対象のみが表現されている。ここで補完が必要なのは、要求内容の要素である「解決する」と要求意図「てほしい」である。これらを表現として補完することにより b のパターンになる。したがって、動機も特定することができる。#10ではネガティブな評価が「のではないか」という反語的表現からの推論によって行われるのに対し、#11の場合には「渋滞」という事象を示す表現から推論される。

h) 要求要素がないもの

#12 別に今のままでいいと思う。

すべての要求要素がなければ回答者の要求は無しとする。

以上の分析から、暗黙的要求意図における要求要素の特定を下記のようにまとめる。

記号の定義：(前述のとおり)

要求および要素特定のパターン：

$$e) R(r1, r2 \dots rn) + C(c1, c2 \dots cn) + I(\phi) = Demand$$

但し、 I は表層的に特定できる場合がある

$$f) R(r1, r2 \dots rn) + C(\phi) + I(\phi) = Demand$$

但し、 R に対するネガティブな評価により NOT(R) から C をアブダクティブに特定する

$$g) R(\phi) + C(c1, c2 \dots cn) + I(\phi) = Demand$$

但し、NOT(C) に対するネガティブな評価から R をアブダクティブに特定する。また、 C から I を特定するとき、機能的側面に拠る場合と内容的側面に拠る場合がある。

$$h) R(\phi) + C(\phi) + I(\phi) \neq Demand$$

以上の分析により、回答者の要求意図を特定するために、表層的な情報を用いて機械的に処理できる要素と、推論知識として規則化が必要な要素があることがわかった。次節では、これらの異なる情報を

得るために回答のどのような表現に着目すべきか、その知識情報について述べる。

(3) 要求特定手続きの提案

以上の分析から明らかになってきたように、明示されていない要求要素を特定する際に必要になるいくつかの情報がある。これらの情報が要求特定の知識構造の項目であり、要求要素をアブダクティブに特定するために利用できる情報であると共に制約や条件となる。

要求分析データベースの作成過程の説明で述べた項目、および分析の説明で述べた項目を下記項目に挙げる。

要求特定手続きに必要な知識情報

- ・要求動機になる信念
 - 信念 $p1 \sim$ 信念 n
 - 信念間の関係 $1 \sim n$
- ・要求内容から動機を特定する際に仮定する信念 x
- ・要求動機に結びつかない信念 (信念 0, 評価, 態度, 態度表明表現)
- ・信念に対する回答者の評価
 - ポジティブな評価
 - ネガティブな評価
- ・要求解釈プロセス
- ・文間の関係 (因果的 / 並列的)
- ・節間の関係 (因果的 / 並列的)
- ・要求内容
- ・要求意図 (要求程度, 要求意図表現)
 - 直接要求表現を含むか否か
 - 直接要求表現への言い換えの可否

なお、上記以外の項目として；

- ・回答者がどのような視点で要求しているか (個人的欲求 / 社会的欲求)
 - cf. 個人的欲求：渋滞は嫌なので、バイパスを作ってほしい
 - 社会的欲求：交通の安全を確保するためには一人一人が交通規則を守るべきです。
 - ・誰に対する要求か (調査者 / それ以外)
- といった情報についても試行的に記述を進めている。

5. おわりに

本研究では、アンケートの回答における回答者の意図を取り出すために、要求に着目し、これを解釈するための要求特定手続きを提案した。この手続きでは、回答には必ず要求が含まれるという仮説にもとづき要求意図を確定することで、他の要求内容・動機といった要求要素を推論し発見できる。これにより、要求を構成する意図、内容、動機が明示されていない場合、必要な要素を発見的に特定できることを実データの分析をもとに報告した。

また、手続きのための要求分析により、機械的な情報抽出処理が可能な表層情報は何か、あるいは規則作成することにより取り出すことが可能な情報は何かを明らかにすることができた。今後は、さらに要求特定知識の記述を進め、作成した知識構造による記述の精度と簡潔さを高めると共に、知識構造内の生起項から暗黙的な項目を自発的に発見できる規則体系の構築を目指している。

謝辞：研究データとして道路審議会基本政策部会「21世紀の道を考える委員会」が実施されたボイス・レポートについて研究利用を快諾して下さった(財)国土技術研究センター調査第二部の前田様、川原様のご厚意に深謝いたします。

参考文献

- 1) 乾裕子・井佐原均：表層情報からの意図タグ判定の試み—自由記述アンケートを対象に—, 言語処理学会第7回年次大会発表論文集, pp 437-440, 2001.
- 2) 乾裕子・兼重賢太郎・矢嶋宏光・井佐原均：アンケートというコミュニケーション—PI手法を取り

入れた意図の抽出方法—, 人工知能学会研究会資料 SIG-SLUD-A 102-9, pp 51-pp 56, 2001.

- 3) Inui,H., Murata,M., Uchimoto,K. and Isahara,H.: Classification of Open-Ended Questionnaires based on Surface Information in Sentence Structure, NLPRS 2001, pp-pp 2001.
- 4) 続有恒・村上英治：心理学研究法9 質問紙調査, 東京大学出版会, 1975
- 5) 有馬昌宏・辻新六：アンケート調査の方法, 朝倉書店, 1987
- 6) 矢嶋宏光：参加型意思決定プロセスとその技術, 土木学会誌 特集「合意形成論」, vol. 87, 6, 2002.
- 7) 道路審議会基本政策部会「21世紀の道を考える委員会」：ボイス・レポート, 建設省道路局建設省都市局, 1996
- 8) 那須川哲哉：コールセンターにおけるテキストマイニング, 人工知能学会誌, Vol. 16, No. 2, pp 219-225, 2001.
- 9) 柳瀬隆史・丸元聡子・難波功・落谷亮：文末の述語表現を利用した質問テキストの解析, 言語処理学会第8回年次大会発表論文集, pp 647-650, 2002.
- 10) 月出奈都子・石崎俊, TV番組に対する自由回答文の印象抽出システム—インターネットアンケート調査による自由回答文の解析—, 言語処理学会第6回年次大会発表論文集, pp 249-251, 2000.
- 11) 高田伸二・屋井鉄雄, アンケート自由記述による道路ニーズ・不満の把握手法の研究, 第35回日本都市計画学会学術研究論文集, pp 571-576, 2000.
- 12) J. R. Searle: Speech Acts. an Essay in the Philosophy of Language, Cambridge University Press, 1969.
坂本百大・土屋俊訳：言語行為, 勁草書房, 1986.
- 13) 森田良行・松木正恵：日本語表現文型, アルク, 1989.