

アンケート回答に現れる要求意図の認定に関する分析

大塚(乾) 裕子^{†‡}

井佐原 均^{†‡}

[†]神戸大学大学院

[‡]通信総合研究所

1. 背景と目的

本研究では人々の意見として収集された自由回答アンケート回答が適切かつ有効に活用されるための方法として回答の自動分類・情報抽出を行っている。とくに、回答から回答者の意図情報を取り出し、その情報をもとにテキストとしての回答を自動的に分類することを目的としている。この実現のため図1に示すように、本研究では自由回答をデータとし(道路審議会 1996), 1) 回答テキストの自動分類, 2) 回答テキストの言語表現に基づく意図判定基準の作成を, 研究の二つの柱として進めてきた。

1) の分類手法としては, 最大エントロピー法に基づく機械学習によって分類を実現している(乾, 村田, 内元, 井佐原 2002)。これは, a) 従来, 分類の対象データとして中心的であった新聞テキストに比べ回答テキストではアンケート調査の質問に対する限定された分野対象における回答内容が中心的であること, そのため, b) 分類対象としては萌芽的である自由回答の処理に最適な語彙体系や辞書は言語データとして整備されていないこと, という主に二つの理由によるものである。分類実験の結果, 本研究で対象とするデータでは回答文の約 40%が要求, 依頼, 願望, 提案, 命令などの要求表現(国語研 1960), および「べきだ」「なければならない」のように義務・当為などの助動詞相当語句(森田・松木 1996)とといったいわゆるモダリティ(仁田・益岡 1989)を含み, また, これらの多様な表現が要求を示す典型表現である「てほしい」に言い換え可能であることがわかった(乾・井佐原 2000)。このことは, 目的志向の研究において語句の意味機能を定義しようとする場合, 構文論の観点のみによる助動詞相当語句の意味分類に限界があることを示しているといえる。

そこで本研究では, 要求の意図を取り出すために語用論的観点を導入し, 表現の形式だけに依らず回答文から要求文を抽出するため, 2) の手法として回答文が「てほしい」を含む文に言い換え可能かどうかという判定を要求意図判定基準として提案した。これは, 人間の言語解釈過程を学習データとして知識化し利用することを試みたものであり, 意図抽出や意図分類のように現在の言語処理技術で

は難しい課題に対する解決方法の一つと考える。この基準によって判定できる言い換えの可否をタグとして回答文に付与し, これを学習データとして要求か否かの判定を機械学習により行った結果, 約 91%の精度を得た。機械学習の結果としては高精度の値が得られたが, 基準としての信頼性を客観的に評価するため基準を用いて異なる複数の被験者により判定(下記実験 a) した場合には κ 係数による一致率(Cohen 1960)は 0.4~0.6 の中程度あるいは実質的な一致しか見られなかった(乾, 内山, 井佐原 2002)。すなわち人手による再現可能性の検証結果としては高い精度が得られなかった。

この結果を分析したところ, 判定の不一致は「歩道が狭い」のように不満を表明することにより暗黙的には要求意図を示すと解釈できる表現において多く見られることがわかった。また, この暗黙的な要求の表明は「歩道を広くしてほしい」のような明示的な要求に対する要求の動機になる。このことは「歩道が狭い」と「歩道を広くしてほしい」を原因・理由の接続助詞「ので」でつないだ際, 意味が通っていると人が判断できることからわかる。さらに「歩道が狭い」という命題を P とすると, この命題から「歩道が狭いので歩道を広くしてほしい」という要求を得ることは, 命題 P をネガティブに捉えているので P の否定($\sim P$)の状態(歩道が狭くない \Rightarrow 歩道が広い)を要望するという, 人が行っている推論プロセスを顕在化したものであると考えられる(乾, 兼重, 矢嶋, 井佐原 2001)。

そこで, 本稿ではこのプロセスを要求動機判定基準として提案し, a) 要求意図判定基準を用いた判定, および c) 基準を用いない判定との比較実験を行うことで明示的要求と暗黙的要求, および要求でないものが精度よく区別できるかを実験検証する。

比較手法 a: 「てほしい」を含んだ文への言い換えによる要求意図判定基準により要求を取り出す

提案手法 b: 手法 a+ 「ので〜てほしい」という文への言い換えによる要求動機判定基準によって暗黙的な要求も要求として取り出す

比較手法 c: 基準を設けず直観によって要求を取り出す

上記三件の実験結果を比較考察することにより, a~c の判定によって取り出せる要求文に現れる表現の違いを明らかにし, 要求意図認定の諸相について整理する。これらの実験および分析を通し, どのような表現が要求意図と解釈されるかについてまとめる。

2. 比較実験

本節では, a) 要求意図判定基準を用いた判定, b) 要求意図および要求動機判定基準を用いた判定, c) 要求意図

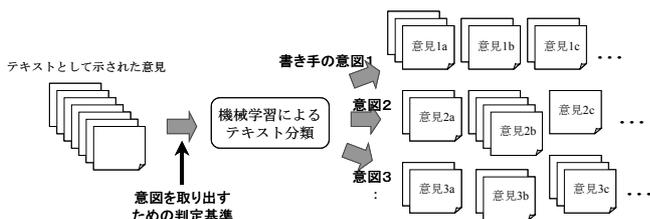


図1 目標とするシステム概要

表 1 要求意図および要求動機判定基準の再現性に関する実験の概要

	実験 a	実験 b	実験 c
実験データ	自由回答(道路審議会 1996)から無作為抽出した共通の 1000 文		
判定対象	要求意図	要求意図および動機	直観で判断できる要求
基準有無	有り	有り	無し
作業指示	<ul style="list-style-type: none"> 回答文の文末を「～てほしい」の形に言い換えて日本語として自然であるかどうか(言い換えの可否) 判断に迷いそうな文も言い換え可能ならばできるだけ言い換える。また、言い換え可能な表現が複数考えられる場合は思いつく限り表記する テキスト全体を言い換えることができないが、何らかの要求表現である場合にも要求を記述する 元文: ガソリン税で道路の整備をしているがこの価格破壊の時に工事費が高すぎると思う 言い換え例: (安くしてほしい)	<ul style="list-style-type: none"> 回答文の文末を「～てほしい」「～ので、～てほしい」の形に言い換えて日本語として自然であるかどうか(言い換えの可否) 元文 1: 大事な事と十分理解しているが、「道」はあくまで人中心に考えるべきであるように思える 言い換え例 1: 大事な事と十分理解しているが、「道」はあくまで人中心に考えてほしい 元文 2: 凸凹している歩道がある 言い換え例 2: 凸凹している歩道があるので、凸凹をなくしてほしい	<ul style="list-style-type: none"> 例示する各回答文には文脈的つながりが無い 回答文が要求であると思う場合は 1, 要求ではないと思う場合は 0 を直観的に判断 無回答やその他の回答は無効になるので、必ずいずれかを選択
被験者数	2 人	3 人	3 人
被験者属性	a1: 男性,工学部学生,20 代 a2: 男性,工学部学生,20 代	b1: 男性,アルバイト,20 代 b2: 女性,アルバイト,20 代 b3: 男性,大学院生(社会人経験有,20 代)	c1: 男性,アルバイト,20 代 c2: 女性,大学生,20 代 c3: 女性,アルバイト(社会人経験有,20 代)

判定基準を用いない判定手法の三つの実験を比較する。これら三実験の概要については表 1 にまとめる。なお、実験 a および c については先行研究(乾, 内山, 井佐原 2002)で行っているため本稿では概要に留め、b)の実験について 2.1 節で詳しく述べる。尚、実験 b で対象とする回答文データは実験 a, c で対象とした 4000 文から無作為に抽出した 1000 文であるため三実験の対象データは共通している。

2.1. 要求意図および要求動機判定基準による実験

1 節に述べたように、この実験の目的は回答文の解釈において判断の一致しにくい「歩道が狭い」のような不満すなわち暗黙的要求を、「歩道を広くしてほしい／広くすべきだ／広くすることが重要だ」のような明示的要求と区別すること、また、要求と要求でないものを区別することにある。a)の「てほしい」を含む文への言い換え可否判定実験との変更点は、暗黙的に要求と判断できる文の文末に「ので」という接続詞をつなげることによって、要求動機を区別することである。そこで、実験にあたっての指示では、1) アンケートの自由回答から要望意見および要望理由を取り出すための作業であることを明記し、2) 自由回答文を、2a)「～てほしい」という文に言い換えられるかどうか、そのうえで、2b)「ので」という語を付けられるかどうかの二つの観点で段階的に言い換え可否を判定し、3) 言い換え可のデータ文については最適と思う言い換えを一文だけ作成することを示した。具体的な事例として、表 1 に示したように「凸凹している歩道がある」という回答文があった場合「凸凹している歩道があるので、凸凹をなくしてほしい」と書き換える例を挙げている。

実験 a では各データに対して 2 人の被験者による作業実験を行ったが、ここでは結果の信頼性の高さをより厳密に測定するため、すなわちデータを作成する基準の厳密さを測定するために被験者を 3 人にした。言い換えの可否判定

に迷うデータについては「？」記号の入力を許可したが、「？」記号を入力してよい条件を修正仕様において流れ図で明記したため実際の「？」入力は少数であった。

2.2. 結果と検証

各実験の結果を表 2 に示す。観察された一致率とは被験者が共に要求である(=言い換え可)・要求でない(=言い換え不可)とした数の割合であり、例えば実験 b の場合、3 人一致の要求判定 A が 715、要求でない(以下、非要求と呼ぶ)という判定 B が 82 であった。ここから計算される観察された一致率とは $0.797 = (A+B)/1000$ である。近年、言語処理の分野でもよく利用されるコーエンの κ 係数は観察された一致率から偶然の一致率を排除したものであるが(Cohen 1960)、クリッペンドルフが指摘するように、基本的に κ 係数は二者一致を求める係数であるため(Krippendorff 1980)、ここではクリッペンドルフの一致係数 α を求める¹。クリッペンドルフの手法では偶然性の排除を不一致の確率から求めている。

$\alpha = 1 - (\text{不一致の観察値} / \text{不一致の期待値}) \dots\dots (1)$
 不一致の観察値を D_0 、不一致の期待値を D_e とすると、

$$\alpha = 1 - D_0 / D_e \dots\dots (2)$$

クリッペンドルフの一致係数では、 m 人のコーダーおよび、二分法ではないカテゴリや数値のシステムを前提としている。 k_{ij} は i 番目の判断者あるいはコーダーによる j 番

¹ (荒木ほか 1999) では、コーエンとクリッペンドルフの一致率の値には差がないと述べているが、実験 a の結果に対して κ 係数を求めたところ 0.755 の値が得られた。一致率 α 0.609 とは約 15% の信頼性の差が生じたことになる。この値は実験の目的や状況に応じては僅差ではない。尚、一致率として信頼できる値は 70~80% 以上といわれているが、検証の目的や状況に応じて妥当な値は変化する可能性があることが指摘されている(Krippendorff 1980)。言語処理の分野においても一致率を評価値として使用する場合、計算方法・得られる値についてさらに検討する必要があると考える。

目の記録を意味する。そこで、信頼性のデータは $m \times r$ の行列となる。このとき、データ内の対の個数を PWC, j 番目データにおける不一致の総和を DWC とする。これらは式 (3) (4) のように定義できる。

$$PWC = rm(m-1) \dots\dots (3)$$

$$DWC = \sum_i \sum_b \sum_c n_{bi} n_{ci} d_{bc} \dots\dots (4)$$

PWC 毎の差の平均として表現される不一致の観察値 D_0 は、次のように定義される。

$$D_0 = DWC / PWC \dots\dots (5)$$

行列全体における対の数は、

$$PT = rm(rm-1) \dots\dots (6)$$

行列全体の不一致は、

$$DT = \sum_b \sum_c n_{b\cdot} n_{c\cdot} d_{bc} \dots\dots (7)$$

これらの値から、不一致の期待値 D_e が定義される。

$$D_e = DT / PT \dots\dots (8)$$

上記の式から、実験 b の結果から得られるデータの信頼性について一致係数 α を 0.54008677 と計算できる。同様に、実験 a, c についても表 2 に示すように一致率を求めた。

表 2 実験結果 (言い換え可否および要求判定の個数)

	被験者	可 (要求)	不可 (非要求)	判定不能	観察された一致率	一致率 α
実験 a	a1	541	459	-		
	a2	760	240	-	0.763	0.609
実験 b	b1	870	121	9		
	b2	797	203	0	0.797	0.540
	b3	783	188	29		
実験 c	c1	602	398	-	0.472	0.270
	c2	157	843	-		
	c3	461	539	-		

最初に行った実験 a の結果から「道路が狭い」のような暗黙的な要求表明については要求の判断が一定しないことが明らかになった。これを「道路を広くしてほしい」のような明示的な要求と区別するため、実験 b では原因・理由の接続助詞「ので」を判定基準に加え「道路が狭い」は「(P:道路が狭い) ので、広く (P の否定の状態 = ~P) してほしい」という書き換え規則を導入したが、表 2 に示すように実験 a の一致率の方が高い、すなわち判断基準としては要求意図判断基準のほうが信頼性が高いこと、明示的要求と暗黙的要求を区別するための基準として提案手法である要求動機判定は統計的検証結果としては不十分であることが示された。そこで次節では、どの点が不十分だったのかを明らかにするため、判定の一致・不一致の内容について詳細に分析する。

3. 一致と不一致の考察

前節の三実験について、被験者ごとの判定の一致・不一致の件数を表 3 にまとめた。表 3 の 1 行目は実験 a, b, c のいずれにおいても被験者が要求 (= 言い換え可) と判定した回答である。これは 1000 文中 137 文であった。このうち 65 件は「してほしい (です/と思う/と思います)」を含むもの、残りの 72 件については「てほしい/ていた

ほしい/お願いします」などの直接要求表現と、「べき」「望む/要望する」「必要である/必要がある」「大切だ/急務である」といった要求意図表現 (乾, 内山, 井佐原) を含むものがほぼ半分の割合ずつを占めていた。

次に、実験 a, b において被験者全員一致で要求と判定されたものは 386 件であった。このうち、直接要求表現を含む回答は下記の例を含む 6 件しかなかった。

- #1 地方自治体の分を含めて、国の借金は 477 兆円になっていることを忘れないでほしい。
- #2 片方の交通手段 (高速道路) を人々の為に解放してほしいのです。
- #3 道路の立体交差もどんどんやって下さい。

全員一致の 137 件には直接要求表現が約 70% 含まれていたことと考え併せると、この結果は直接要求表現を含んでいれば要求の判断が一定することを示している。

一方、要求意図表現およびその他の表現形式については、上述の例に加え、1) 「道路周辺の景観を美しくすることで、落ちついた運転ができ安全な生活ができるのではないだろうか/建設省の若手をどしどし、欧米で勉強させる事が必要ではないでしょうか」のように形式として否定や疑問を含んだ婉曲的な表現、2) 「公共の乗り物を充実させる/料金所の自動化」のような命題の提示、3) 「モラルが向上^(c)すれば、かなり安全になる^(e)はずです/冬は雪対応を優先する^(c)ことが、生活レベルの向上になる^(e)」のような要求内容 (c) を条件節として表現し、後続する節でその効果 (e) を示した表現が見られた。

以上はサールの間接言語行為の成立条件として示されている下記の間接的な依頼・命令の条件内容にも一致することから (Searle 1976; 石崎 2001), 要求意図判定基準は言語の表現に依存せず話者の意図を扱う枠組みである言語行為論 (Searle 1969) の諸条件を言語処理技術として実装できる可能性を持つ規則であると捉えることができる。(石崎 2001 より引用)

- 1) 聞き手の能力に関する準備条件についての質問 (例「塩に手が届きますか?」)
- 2) 命題内容の質問あるいは言及 (例「夕食時にはネクタイを着用すること」)
- 3) 誠実性条件についての言及 (例「それをやらしてもらえれば、感謝します」)
- 4) 行為をする理由についての質問あるいは言及 (例「母親にはもっと礼儀正しくすべきだ」)

特に、表 3(1)の実験 c の結果のように、被験者 c1~c2 が非要求とした回答文の中で 3 人の判断が一致していた 33 件の中には、上述の誠実性条件のように条件と効果を表明することにより暗に要求を示すパターンが 7 件含まれていた。直観によって要求を判定する実験の被験者が 3 人とも要求と判断しなかったことは、このように表明された意図を適切に処理することの難しさを意味する。したがって、このような意図を取り出せることは基準が有効であることを示すと考える。

表 3 判定の一致・不一致の個数を示した実験結果の比較

実験ごとの一致・不一致	言い換え可否 被験者	実験 a		実験 b			実験 c			直感的判定	
		要求意図判定基準	判定基準	要求意図および要求動機判定基準	動機判定基準		判定基準なし				
a,b,c で一致	可	a1	a2	b1	b2	b3	c1	c2	c3	要求	
a,b で一致		137						305	10		258
b で一致	386						81	376	128		
b で不一致	可	8	152	192			114	8	54	要求	
	不可	184	40				78	184	138	非要求	
b で不一致	可および不可	(211)									要求および非要求
a,b,c で一致	不可	74									非要求

表 4 条件と効果の表明による暗黙的な要求のパターン

回答文	パターン	PN
経済活動に影響がでるような交通規制は問題。	CはE	N (1)
都市の機能を地方に分散させることで、様々な面で地方が汚染されてしまう。	CことでE	N (2)
市街地における土地区画整理事業の効果はすでに実証済みであり、商業活動が極めて活発になることから、人口の流出を防げ、都市部の人口一極集中を防ぐことができる。	CはE	N (2)
冬は雪対応を優先することが、生活レベルの向上になる。	CことがE	P (1)
広域的な機能は、利用する住民の利便性を考えて分散し、みんなで使うようにした方が機能の充実も図れる。	CたほうがE	P (1)
自分達が使用する車の事だけ考えずに、緊急時にいち早く緊急車輛が走れるということも「くらしやすい住環境」だと思う。	CということもE	P (1)
モラルが向上すれば、かなり安全になるはずですよ。	CればE	P (1)
地方都市の経済の活性化を図るならば、大都市と地方都市のネットワークが必要でしょう。	EならばC	P (1)
今の技術を使えば、可能と思います。	CればE	P (2)
地域の活性化や、人口増を図るには、地理条件はもとより、交通の便が絶対条件になる。	EにはC	P (2)

実験 b の被験者間では要求判定が一致した 192 例の中で被験者 a1, a2, および c1~c3 が非要求とした表 3(2) のうち、共通判定された回答文は 19 件あった。ここにも、上述のような条件と効果による暗黙的な要求の表明が見られたが、興味深い点としては(1)と(2)では現れる表現パターンに違いが見られることである。この違いを上述の要求内容(c), 効果(e)の観点を用いて C と E の表記で記述した場合、表 4 に示すように (1)では E の要素にポジティブ(P)な効果が、(2)ではポジティブな効果およびネガティブ(N)な効果がともに表現として現れることがわかった。この違いは、本稿での提案手法である要求動機判定基準に拠ると考えられる。この傾向の違いの深い考察については今後の課題とする。

また、表 3(3)の 211 件は、各実験の被験者間で判定に一致が見られない回答文である。これらの回答と非要求と判定された 74 件の回答との大きな違いは、上述のネガティブな要素が回答に現れているか否かである。これは、現状否定的な認識が回答に表現として現れていなければ直接的な要求の動機とはなりにくいことを示している。211 件の内訳についての詳細な分析は本稿では行わないが、傾向としては上述のネガティブな要素・ポジティブな要素に加え、(内山・乾・井佐原 2004) の要求内容の同定との関連が深い。要求内容(c)やその要素である「A について」がないもの、あるいはこれらの内容や要素の表現が抽象的すぎるものについて判定の一致しない傾向が見られた。

4. まとめと課題

本稿では、暗黙的に表明される要求意図を取り出すために要求動機判定基準を提案し、提案手法を用いない場合との比較実験を行った。結果の統計的値としては基準の有効

性を示すことはできなかったが、基準によって判定が一致する場合しない場合を分析することにより、条件と効果の表明による要求の有無など段階的に現れる表現の違いが見られることがわかった。今後はこれらの傾向についてさらに詳細に分析し、どの範囲までが要求として取り出すべき回答かを明らかにしていく。

参考文献：

Cohen, J. (1960) A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, 20, pp.37-46.
 道路審議会基本政策部会 (1996) ボイス・レポート. 21 世紀の道を考える委員会 建設省道路局 建設省都市局.
 石崎雅人 (2001) 言語使用の基礎理論. 石崎雅人・伝康晴『談話と対話: 言語と計算 3』2 章, 東京大学出版会.
 国立国語研究所 (1960) 『話しことばの文型』秀英出版.
 Krippendorff, K. (1980) Content Analysis: An Introduction to Its Methodology, Sage Publication. K・クリッペンドルフ著 三上俊治・椎野信雄・橋元良明訳(1989)『メッセージ分析の技法; 「内容分析」への招待』勁草書房.
 乾裕子・井佐原均 (2000) 表層情報からの意図タグ判定の試みー自由記述アンケートを対象にー. 言語処理学会第 7 回年次大会発表論文集, pp.437-440.
 乾裕子・兼重賢太郎・矢嶋宏光・井佐原均 (2001) アンケートというコミュニケーションーPI 手法を取り入れた意図の抽出方法ー. 人工知能学会研究会資料 SIG-SLUD-A102-9, pp.51-56.
 乾裕子・村田真樹・内元清貴・井佐原均 (2003) 表層表現に着目した自由回答アンケートの意図に基づく自動分類. 自然言語処理, Vol.10 No.2, pp. 19-42.
 乾裕子・内山将夫・井佐原均 (2003) 言い換えによる自由記述アンケート回答の要求意図判定基準の作成および検証. 言語処理学会第 9 回年次大会発表論文集. Pp230-233.
 仁田義雄・益岡隆志 (1989)『日本語のモダリティ』くろしお出版. Searle, J. R. (1969) Speech Acts, an Essay in the Philosophy of Language. Cambridge University Press. 坂本百大・土屋俊訳 (1986) 言語行為. 勁草書房.
 Searle, J. R. (1976) Expression and meaning. Cambridge University Press.
 内山将夫・乾裕子・井佐原均(2004)自由回答アンケートにおける要求内容の分析. 言語処理学会第 10 回年次大会発表論文集.