

# コールセンタ対話における話者の知識量推定

宮崎 千明 東中 竜一郎 牧野 俊朗 松尾 義博  
 日本電信電話株式会社 NTT サイバースペース研究所

{miyazaki.chiaki, higashinaka.ryuichiro, makino.toshiro,  
 matsuo.yoshihiro}@lab.ntt.co.jp

## 1 はじめに

近年、多くの企業において、一般消費者からの問合せや注文などを受け付ける総合窓口としてコールセンタが導入されている。コールセンタに蓄積された「顧客の声」は「貴重な経営資源」[8]と考えられており、その分析は企業の意味決定において重要な意味を持つ。しかし、蓄積された通話数はあまりに膨大で、人手ですべての通話を聴取し、分析することは不可能である。よって、分析対象を絞り込む手段として、通話を何らかの基準によって分類する技術が求められている。

通話分類の軸として、我々は、問い合わせしてくる顧客の知識量が有用であると考えている。なぜなら、コールセンタの通話内容は、顧客の知識量によって大きく異なると考えるからである。例えば、どのようなトラブルに陥りやすいか、また、どのような言い回しで説明すれば理解されやすいか、というのは顧客の知識量によって異なる。そこで、本研究では、コールセンタの通話などの情報・知識の伝達や獲得を行う対話から、問い合わせる側の話者の知識量を推定する手法を提案する。

本研究の目的は、話者の知識量を反映する特徴量を探り、高精度な知識量推定を実現することである。アプローチとしては、教師あり機械学習を用いて知識量の自動推定を行う。

## 2 知識量主観判定

### 2.1 知識量判定基準の策定

話者の知識量とは、そもそも客観的に判定できるものであるのかを検証するため、事前調査として、複数名の作業者に一定の基準に基づいて知識量を主観評価させ、その一致度を確認した。

話者の知識量を表 1 に示される 4 段階に設定し、主観判定を実施した。対象とする対話コーパスとしては、問合せ、契約等をタスクとするコールセンタの通話音声を用意した。

表 1 の定義付けにおける「必要最低限の知識」とは、

表 1: 知識量主観評価の判定基準 (4 段階)

知識量	定義
1	通話前は何も知らない
2	必要最低限の知識しかない
3	ある程度の知識はあるが、十分ではない
4	オペレータが説明する内容のほとんどは通話前から知っている

表 2: 知識量主観判定の一致度

	一致率	値
4 段階判定	0.52	0.25
2 段階判定	0.70	0.41

オペレータが会話の前提知識であるとみなし、特段の説明なしに対話に導入するような知識のことである。例えばインターネットサービスプロバイダのコールセンタの場合、オペレータは「インターネットを利用するにはインターネット回線が必要である」という知識について、ユーザも知っているという前提で会話を開始する。この場合、「インターネットを利用するにはインターネット回線が必要である」という知識は必要最低限の知識に該当する。「十分な知識」の量については、オペレータと同等の知識量として定義した。

### 2.2 知識量判定の一致度

表 1 の判定基準に基づいて、261 対話を対象に話者の知識量判定を実施した。具体的には、作業員 2 名 (筆者らを含まない) に個別に対話を聴取させ、各対話のユーザの知識量が、4 段階のうち、どの知識量に該当するかを判定させた。表 3 に知識量主観判定の一致率、および 値 (Cohen's Kappa) を示す。ここで、一致率は全 261 通話中、2 名の判定が一致した割合を示している。

表 2 の通り、知識量を 4 段階で定義し、主観判定をした場合、作業員 2 名による判定の一致率が 0.52、値は 0.25 となり、ともに値が低く、複数名の間で一致した判定は難しいことが分かった。そこで、知識量判定基準を表 3 に示される 2 段階で定義したところ、作

表 3: 知識量主観評価の判定基準 (2 段階)

知識量	定義
小	必要最低限以下の知識しかない
大	ある程度以上の知識がある

業者 2 名の主観判定の値は 0.41 となり、判定は中程度一致することが確認できた。

以上により、話者の知識量を必要最低限以下の知識量であるか、それよりも多いかという 2 段階で判定することは可能であると分かった。以降、2 段階の判定基準に基づいた知識量主観判定を行い、作業者 2 名の判定が一致した 180 対話を対象に、知識量推定技術の評価を行うこととする。なお、180 対話中のユーザの知識量の内訳は、知識量小が 94 対話、知識量大が 86 対話である。

### 3 提案手法

本研究では、教師あり機械学習によってユーザの知識量を推定する。推定のための特徴量としては、ユーザの使用語彙に関する特徴量 (語彙的特徴量)、及び、オペレータとユーザの対話のやりとりに関する特徴 (対話的特徴量) を用いる。

#### 3.1 語彙的特徴量

##### 3.1.1 語彙の専門性

関連研究に倣い、語彙の専門性を知識量推定の特徴量として利用する。検索システムのユーザを対象とした関連研究 [3] では、ある分野において認知度の低い事柄を検索している人は知識の豊富な人であると捉え、検索クエリとして入力された単語の珍しさを表す指標を用いて、ユーザの知識量を推定している。これに倣い、本研究では、話者の使用語彙の専門性を特徴量化する際、その話者が使用する単語の珍しさの平均値を用いる。単語の珍しさとしては、Inverse Document Frequency (IDF) を用いる。総対話数が  $N$ 、単語  $t$  を含む対話数が  $df_t$  のとき、単語  $t$  の IDF 値は以下の方法で算出する。

$$idf_t = \log \frac{N}{df_t}$$

本研究では、ユーザの使用語彙の平均 IDF は、音声認識結果から抽出した発話列を入力として形態素解析を行い、その出力中の名詞と未知語のみを利用して算出する。なお、形態素解析には形態素解析システム「茶筌」<sup>1</sup>を使用した。

<sup>1</sup><http://chasen-legacy.sourceforge.jp/>

表 4: 発話意図判定のためのパターン例

意図	パターン	
質問	ですか。かしら。等	
説明	記述文・同定文	です。でして。等
	理由の説明	ので。から。
	聞き手の意思決定に資する情報・条件の提示	けど。けれど。等
相槌	はい。ええ。うん。	

#### 3.1.2 異なり語数

異なり語数とは、通話においてユーザが何種類の単語を使用したかを指す。同様の用件の通話を終えるのに、あるユーザが他のユーザより多くの種類の単語を使う必要があったということは、そのユーザの意思疎通が非効率であったことを示していると考え、本研究では、ユーザの使用語彙の異なり語数を特徴量として用いる。異なり語数の算出は、音声認識結果から抽出した発話列を入力として形態素解析を行い、その出力中の名詞と未知語の異なり数を数えることによって行う。

#### 3.1.3 対話的特徴量

コールセンタ対話は、主に情報・知識の伝達や獲得を目的としたコミュニケーション [6] で構成されていると考えられる。そこで、そうした対話において重要な発話意図である「質問」、「説明」、「相槌」に該当する発話を抽出し、知識量推定の特徴量として導入する。本研究では「質問」、「説明」、「相槌」に該当する発話を、発話末の言語表現に対する正規表現のパターンマッチによって、表層的に判別する。表 4 は「質問」、「説明」、「相槌」の判定に利用した正規表現パターンの例である。

#### 3.1.4 質問に関する特徴量

質問とは、相手から情報を得るための言語手段であると考えられている [5]。よって、発話される質問の回数には、話者が持ち合わせる情報の量、つまり知識量が反映されていると考えられる。表 4 の例に示されるような発話末の言語表現を手掛かりに質問発話を抽出し、オペレータとユーザそれぞれの質問回数を特徴量として用いる。

#### 3.1.5 疑問詞疑問文に関する特徴量

質問の形式をとる発話の中には、話者の持つ仮説の正否を確認する機能を果たすものが含まれている。しかし、話者の知識量を推定するうえでは、未知の情報を要求する質問を抽出することが重要であると考えられるため、オペレータとユーザそれぞれが発話した「疑問詞疑問文」の回数も抽出し、特徴量として用いる。

疑問詞疑問文とは、疑問詞（「何」、「どう」、「どこ」等）が含まれる疑問文のことであり、本研究では、質問発話の中に疑問詞に該当する言語表現があるかどうかを正規表現で判別することによって疑問詞疑問文を抽出し、オペレータとユーザそれぞれの疑問詞疑問文の発話回数を特徴量として使用する。

### 3.1.6 説明に関する特徴量

本研究では、相手に情報を提供する行為を「説明」と呼ぶ。具体的には、表4に示される3種類の発話を説明とみなす。「記述文」とは、ある事物の属性を記述する文のことで、「同定文」とは、ある役割に値を割り当てる文のことであり[4]。例えば、「日本の首都は東京です。」は「日本の首都」という役割に「東京」という値を割り当てる同定文である。記述文・同定文の抽出は、発話末の「です」等のパターンを用いて判定する。

「理由の説明」とは、自身の先行する発話に対し、発話内容（命題）が成立する理由を付け加えるための発話とする。本研究では、発話末が「ので」または「から」で終わっているものを「理由の説明」として抽出する。

「聞き手の意思決定に資する情報・条件の提示」とは、聞き手が何かをするための参考となるよう、提示される情報のことを指す。これは、言い終わりの「けど」の持つ終助詞的機能[9]であり、本研究では、発話末が「けど」等のパターンで終わっているものを抽出し、説明の一部として加えることとする。

### 3.1.7 相槌に関する特徴量

対話において聞き手が打つ相槌の果たす役割は、相手の発話の継続を促すと同時に、相手の発話内容を理解していることを示すことなどである[1]。よって、ある話者が相槌を打った回数は、その人が聞き手に回った回数や、どのくらい相手の発話内容を理解できたかを表すと考えられ、話者の知識量との関連が期待される。そこで我々は、発話末に「はい」、「うん」、「ええ」が現れている発話を相槌発話とし、その回数を特徴量として使用することとする。

## 4 実験

コールセンタの通話を対象として、知識量の少ないユーザの対話を抽出する実験を行い、本研究で提案した特徴量が、話者知識量推定に有効であるか評価した。知識量の少ないユーザを対象とした理由は、知識の少ない人ほどトラブルや疑問を抱えやすく、コールセンタにおける分析対象として重要なためである。

表5: 各種特徴量を用いた実験の適合率・再現率

特徴量	適合率	再現率
提案手法 (F1-F10)	0.80	0.43
ユーザの bag of words (B1)	0.72	0.28
発話数 (B2)	0.61	0.45

### 4.1 実験手順

対象としたデータは、2節で述べたコールセンタの通話180通話の音声認識結果である。2.2節の手順により、そのうち94対話に「知識量小」、86対話に「知識量大」のラベルが付与されている。なお、音声認識にはVoiceRex[7]を用いた。使用する特徴量は、(F1)ユーザの語彙の平均IDF値、(F2)ユーザの異なり語数、(F3)オペレータの質問回数、(F4)ユーザの質問回数、(F5)オペレータの疑問詞疑問文の回数、(F6)ユーザの疑問詞疑問文の回数、(F7)オペレータの説明回数、(F8)ユーザの説明回数、(F9)オペレータの相槌回数、(F10)ユーザの相槌回数の10種類とした。

各通話の音声認識結果から抽出した特徴量とその通話にあらかじめ付与された知識量ラベルの組を学習データとし、教師あり機械学習の手法であるSupport Vector Machine (SVM)を用いてモデルを学習し、知識量を推定した。カーネル関数には2次の多項式カーネルを用い、学習・推定は10分割交差検定によって実施した。さらに、単語の出現回数や発話数などの単純な特徴量と精度を比較するために、(B1)ユーザの発話における各単語の出現回数 (bag of words)、(B2)オペレータおよびユーザの発話数を特徴量として用いた比較実験も行った。

### 4.2 評価尺度

情報抽出で一般的な、適合率 (precision)・再現率 (recall)を用いて「知識量小」の推定精度を評価した。正しく抽出できた「知識量小」の数をR、「知識量小」として抽出した総数をN、全データ中の「知識量小」の数をCとすると、それぞれ以下の方法で算出できる。

$$\text{precision} = \frac{R}{N}, \quad \text{recall} = \frac{R}{C}$$

本技術は、コールセンタの膨大な通話の中から、分析対象を絞り込むために利用されることを考慮し、本稿では、上記2種類の評価尺度のうち、適合率の向上をより重視する。

### 4.3 結果

表5が実験結果である。提案した特徴量 (F1-F10)によって、適合率0.8、再現率0.43の精度で知識量小を抽出することができた。

表 6: 再現率をそろえた場合の適合率の比較

特徴量	適合率	再現率
提案手法 (F1-F10)	0.80	0.43
ユーザの bag of words (B1)	0.67	0.43
発話数 (B2)	0.65	0.43

表 7: 適合率をそろえた場合の再現率の比較

特徴量	適合率	再現率
提案手法 (F1-F10)	0.80	0.43
ユーザの bag of words (B1)	0.78	0.07
発話数 (B2)	0.80	0.09

表 6, 表 7 は, 正例と負例を分類する閾値を変化させることによって, 再現率をそろえた場合の適合率, 及び, 適合率を概ねそろえた場合の再現率を示している. 表 6 より, 再現率が同じとき, 適合率が最も高いのは提案手法であることが分かる. また, 表 7 より, 比較対象 (B1, B2) の手法で提案手法と同程度の適合率を達成しようとするとき, 再現率が大幅に低下することが分かる. 以上より, 単語の出現回数や発話数などの単純な特徴量と比較して, 提案手法が抽出の正確性, 網羅性ともに優れていることが確認でき, 提案手法の有効性が示された.

#### 4.4 各特徴量の効果

提案した特徴量 (F1-F10) それぞれの効果を検証するために, 一つずつ特徴量を除いて 4 節と同様の実験を行い, 結果を表 8 に示した. 表 8 より, (F1) ユーザの使用語彙の平均 IDF 値は除外しても精度には変化がなく, 本研究で対象とした対話の話者の知識量推定においては寄与が認められなかった. 一方, 以下の 4 つは, 除外すると著しく精度が低下することから, 知識量推定に大きく寄与する特徴量であると分かった.

- ユーザの異なり語数 (F2)
- オペレータの説明回数 (F7)
- オペレータの相槌回数 (F9)
- ユーザの相槌回数 (F10)

上記 F2 については, 3.1.2 節で述べた仮説の通り, ユーザの使用語彙の異なり語数によって, ユーザの意思疎通の非効率性を捉えることができた結果だと考えられる. F7 については, オペレータがユーザに情報を与えた回数によって, ユーザに欠けていた知識の量を捉えることができたものと考えられる. F9 は, オペレータが聞き手に回った回数を表し, ユーザの要件や問合せ内容を理解するのにかけた手間を捉えることができたものと思われる. F10 は, ユーザが聞き手に回った回数を示しており, F7 と同様に, オペレータがユーザに与えた情報量を捉えていると考えられる.

表 8: 各特徴量を除いた場合の推定精度

特徴量	適合率	再現率	適合率の増減	再現率の増減
提案手法 (F1-F10)	0.80	0.43	N/A	N/A
w/o F1	0.80	0.43	0.00	0.00
w/o F2	0.72	0.35	<b>-0.08</b>	-0.08
w/o F3	0.76	0.44	-0.04	+0.01
w/o F4	0.78	0.42	-0.02	-0.01
w/o F5	0.78	0.43	-0.02	0.00
w/o F6	0.78	0.43	-0.02	0.00
w/o F7	0.71	0.37	<b>-0.09</b>	-0.06
w/o F8	0.77	0.44	-0.03	+0.01
w/o F9	0.70	0.40	<b>-0.10</b>	-0.03
w/o F10	0.49	0.43	<b>-0.31</b>	0.00

## 5 まとめと今後の展開

本研究では, コールセンタの通話から, 問い合わせるユーザの知識量を推定する手法を提案した. 使用語彙の異なり語数と発話意図に着目した対話的特徴量を用いることを提案し, 適合率 0.8 の精度で知識量の少ないユーザの通話を抽出できることを確認した.

今後取り組みたいことは, 話者の知識量と, その人の使用語彙や対話に表れる特徴についての定量的・定性的な分析である. 関連研究として, 近年, 話者の verbal intelligence (目的達成のための言語使用能力) と, その人の会話に表れる特徴を分析した研究 [2] がある. こうした研究で報告された知見を取り入れつつ, 特徴量の分析を進めたい.

## 参考文献

- [1] S. K. Maynard. *Japanese Communication: language and thought in context*. University of Hawai'i Press, 1997.
- [2] K. Zablotskay, U. Rahim, S. Zablotskiy, S. Walter, and W. Minker. Conversation peculiarities of people with different verbal intelligence. In *Proc. IWSDS 2011*, pp. 153-159, 2011.
- [3] 佐藤大祐, 安田宣仁, 望月崇由, 鈴木智也, 松浦由美子, 片岡良治. 検索システムユーザの分野別の知識推定. In *Proc. DEIM Forum 2010*, 2010.
- [4] 坂原茂. 役割, ガ・ハ, ウナギ文. pp. 25-30, 1990.
- [5] 植木迪子. 知っていること, 知りたいこと. *独語独文学科研究年報*, Vol. 20, pp. 267-282, 1993.
- [6] 深田博己. *インターパーソナル・コミュニケーション*. 北大路書房, 1998.
- [7] 政瀧浩和, 柴田大輔, 中澤裕一, 小橋川哲, 小川厚徳, 大附克年. 顧客との自然な会話を聞き取る自由発話音声認識技術「VoiceRex」. *NTT 技術ジャーナル*, Vol. 18, No. 11, pp. 15-18, 2006.
- [8] 那須川哲哉. テキストマイニングを使う技術 / 作る技術. 東京電機大学出版局, 2006.
- [9] 白川博之. 「ケド」で言い終わる文. *広島大学日本語学紀要*, Vol. 6, pp. 9-17, 1996.