

雑談対話に含まれる場所を表すフレーズの分析と検出手法の検討

Analysis and Detection Methodology Study for Phrases Indicating Location in Chat-Based Spoken Dialogue

成松宏美^{1*} 杉山弘晃¹ 水上雅博¹
Hiromi Narimatsu¹ Hiroaki Sugiyama¹ Masahiro Mizukami¹

¹ NTT コミュニケーション科学基礎研究所

¹ NTT Communication Science Laboratories

Abstract: The purpose of this study is to detect words or phrases which indicate location in spoken language for chat-based spoken dialogue systems. Humans use various linguistic representations to indicate location. The representations include words which do not exist in corpus or phrases represented as places where humans' actions or events happen. These phrases cause the system to misunderstand the meaning of users' utterances. Not to cause such misunderstanding, we examine the characteristics of these phrases and propose a new method to detect them.

1 はじめに

本稿では、人と人の雑談対話中に現れる“場所”を表すフレーズをシステムが自動で理解できることを目的として、当該フレーズの分析と検出手法を検討する。近年、注目が高まっている雑談対話システムでは、ドメインが限られていない状況下で、ユーザの発話を正しく理解し、その内容にふさわしい応答を返すという、難しい課題がある。このとき雑談システムは、人の発話を理解するために、その単語やフレーズが何を表すものかを正しくグラウンディングする必要がある。例えば、人の発話に含まれる 5W1H 構造を利用して発話を生成するシステムであれば、発話中のどの単語やフレーズが、いつ（時間）、どこで（場所）、だれが（主語）、何を（目的語）、どうやって（手段）のいずれに該当するかを正しく当てはめる必要がある。表 1 に、システムがユーザ発話から単語やフレーズの理解を正しくできた場合と誤った場合とで、どのような発話が生成されるかの例を示す。一つ目の例では、誰とにあたるフレーズが、職場のメンバであることを一つのフレーズとして正しく理解できなかったために、職場を行った場所と勘違いしている。二つ目の例では、仮定節にあたるフレーズが滋賀に行ったらであり、（滋賀の）バウムクーヘンのお店が場所に該当するフレーズである。バウムクーヘンのお店を正しく場所として理解できなかったために、滋賀に行きたいと判断し、その理由を問う質問を進めている。このように、システ

表 1: システムが誤理解した場合の例.

昨日、職場のメンバで焼肉を食べに行った。
○ 職場のメンバですか。仲良しで、いいですねえ。
× おお、職場で焼肉食べたんですか。いいですねえ。

滋賀に行ったら、バウムクーヘンの有名なお店に行きたいです。
○ あっ、バウムクーヘンのお店有名って聞いたことがあります。
× 滋賀に行きたいところがあるんですか？

ムがユーザの発話を正しく理解するためには、発話中のどのフレーズが時間、場所などの項目に該当するかを正しく理解することが求められる。

本稿では、上記例のうち“場所”に着目して、ユーザ発話の分析および検出を行う。場所の理解には、コーパス中で“場所”として登録されていない単語やフレーズであっても場所として正しく理解することが求められる。従って、当該フレーズの区間も正しく検出することも必要である。以下の文は「店 A」および「昨日ラーメンを食べた店」はいずれも場所を表すが、いずれもコーパスに場所として登録されているものではない。

- 店 A は、近々移転するらしい。
- 昨日ラーメンを食べた店は、近々移転するらしい。

店名などをコーパスに登録することも可能であるが、店名等は多数存在し、他の一般名詞等とも重複することから安易に全て登録することは、他の一般名詞を誤っ

*連絡先：日本電信電話株式会社
NTT コミュニケーション科学基礎研究所
〒 619-0237 京都府精華町相楽郡光台 2-4
E-mail: narimatsu.hiromi@lab.ntt.co.jp

て店名と認識する等の誤理解を生じる可能性があるため望ましくない。また、昨日ラーメンを食べた店のように、人の行動やイベントと結びつけて特定の場所を表現する場合もあり、場所を表すフレーズの区間を理解することが必要である。

そこで本稿では、場所を表現する語やフレーズを場所として検出すること、さらにフレーズである場合にその区間を正確に検出することを目的として、人と人の雑談対話中にどのような表現が現れるかをアノテーションを行うことにより分析し、その検出手法を検討する。

2 関連研究

発話文中の特定の単語やフレーズを検知する手法は、場所に限定せず従来より行われており、大きく2つに分類できる。1つ目は、タスク対話においてタスクの対象を把握するための、特定フレーズの理解である [5][6]。旅行の推薦や予約システム等では、システムは、ユーザとの対話を通して、期間や場所、人数等の必要な情報を聞き出し、その情報を用いて該当するツアーの予約や、推薦を行っている。例えば、予約システムであれば、{YY年MM月DD日からDS日間}等の、YY, MM, DD, DSに当てはまる部分をキーフレーズマッチングにより穴埋めする。ユーザの発言が「2017年8月(初旬)から8日間のツアー」であり、システムがかっこ内の日にちを理解できなかった場合には、「出発の日にちを教えてください」と質問することで、不足の情報を聞き出す。理解すべき項目が確定しているため、予めグラウンディングの対象となるキーフレーズを登録しておくことが可能である。一方、本稿で対象とする雑談対話中に出現するフレーズはグラウンディングの対象が限られていないことから、同様のアプローチを取ることが難しい。

2つ目は、文書や発話文に出現する語やフレーズが、何のカテゴリを表現するものかを推定する手法である。1995年に米国のMUC-6[1]で定義された固有表現抽出(Named Entity Task)では、Conditional Random Fields(CRF)をベースとした手法が主流となっている。日本でも、IREX[3]にタスクとして取り上げられ、組織名、人名、地名、固有物名、日付表現、時間表現、金融表現、割合表現の8つの固有表現クラスが定義されている。近年では質問応答システムや情報抽出テキストマイニングの応用を目的として拡張固有表現も提案されている[7]。固有表現抽出タスクとして様々な手法が提案されてきたが、CRFをベースとした識別モデルが成功を取めている[4]。CRFをベースとする固有表現抽出手法では、例えば地名の始点には「B-地点」、地名の続きには「I-地点」、人名の始点には「B-人名」、何にも

該当しないものには「O」のラベルを付与し、文中の地点に対応するフレーズの区間を学習し、推定する。本タスクでは、場所に対して、市や州、日本で言えば県、国、などがカテゴライズされており、90パーセント程度の検出精度を実現している。本手法は、単語だけでなくフレーズも扱えるという点で有用であると考えられる。しかし、扱うデータは、新聞記事等に現れる国名や県名などの固有名詞を前提としており、我々の扱う雑談対話中に現れるような、地名としての固有名詞だけではない場所を表す単語や、人の行動やイベントと組み合わせて表現される場所を表すフレーズに対しての有用性は検証されていない。

人の発話から正しく場所を表す語やフレーズを推定するためには、それに合った学習データの構築が必要である。

3 場所フレーズアノテーション

本章では、人の発話に含まれる“場所”を表すフレーズの特徴を分析するために、人と人の雑談対話データを用いて各発話に含まれる場所を表す語やフレーズにアノテーションを行った。

3.1 アノテーション手順

アノテーション対象データおよびアノテーション手順を説明する。アノテーションの対象とするデータは、人2名がテキストにて雑談を行ったデータである。1対話約40発話からなる、計600対話24888発話に対してアノテーションを行った。

対話の各発話を読んで、発話中に含まれる語やフレーズが場所を表している場合に、アノテーションを行う。一つの発話に複数の場所を表すフレーズが存在する場合には、複数に対してそれぞれアノテーションを行う。場所を表す語やフレーズの判断条件は以下の通り設定した。

1. 各助詞の前(修飾語)を含んだできる限りひとまとまりのフレーズを場所として扱う
例) 街の駄菓子屋さんに行った
(望ましくないアノテーション) 街、駄菓子屋さん、
(望ましいアノテーション) 街の駄菓子屋さん
2. エリアを限定するような表現「～らへん」や「～するところ」等もアノテーションを行う
例) 皇居らへん、昨年ピクニックで行ったところ、等
3. 少しでも場所を表現すると考えられる場合には、できる限りアノテーションを行う。

4. アノテーションの迷い度合いを判断するために、表 2 に示す場所曖昧性ラベルを付与する。

表 2: 場所曖昧性ラベル.

曖昧性ラベル	ラベルの付与基準
a	迷わなかった
b	一応付与したが、付与すべき対象か迷った
c	もっともそうな単語もしくは句に対してアノテーションを行ったが、どこで区切るべきか迷った
d	一応付与したが、本来付与すべき対象か迷った、かつどこで区切るべきかも迷った

場所のアノテーションは人により解釈が異なる可能性を考慮し、条件 4 に示すアノテーションの曖昧性ラベルを付与することとした。

600 対話のうち最初の 30 対話については、アノテータ 1、アノテータ 2 の 2 名でアノテーションを行った。アノテータの一致率を、表 3 に示す。一方を正解アノテータ、もう一方を検出アノテータとして、以下の計算式により一致率を算出した。

$$v = \frac{\text{両者が同一の区間を検出したフレーズ数}}{\text{正解アノテータによって検出されたフレーズ数}} \quad (1)$$

all が曖昧性ラベルに依らず全アノテーション結果に対して算出した一致率、label a が曖昧性ラベルに対して a(迷わなかった) が付与されたアノテーション結果のみに対して算出した一致率を表す。全アノテーションデータに対する一致率は 8 割以上、曖昧性ラベル a が付与されているデータのみ的一致率は約 9 割であった。このことからアノテーションの一致率は十分に高いと判断し、以降の 270 対話に対しては 1 名のアノテータでアノテーションを行った。

表 3: アノテーション一致率.

アノテータ		v	
正解	検出	全データ	ラベル a
アノテータ 1	アノテータ 2	0.87	0.89
アノテータ 2	アノテータ 1	0.83	1.0

3.2 アノテーション結果の分析

前節に示す手順によりアノテーションを行った結果を分析する。アノテーションされた場所フレーズ数と

曖昧性ラベルの内訳を、表 4 に示す。約 7 割の場所フレーズは、迷いなく付与されていることがわかる。また、それ以外の 3 割の内訳を見ると、アノテーションすべき対象か迷ったラベル b, c, d の総和に対して。アノテーションする対象区間を迷うラベル c は比較的少ないことから、人は、フレーズの区間をある程度迷いなく判断していると考えられる。

表 4: アノテーション結果中の曖昧性ラベルの割合.

曖昧性ラベル	a	b	c	d	合計
数	2914	1025	216	47	4202
(割合)	(0.69)	(0.24)	(0.05)	(0.01)	

アノテーションされた場所フレーズが何単語で構成されているかの内訳を表 5 に示す。上段を全アノテーションデータを用いた場合の内訳、下段をラベル 1 が付与されたデータの内訳を示している。結果より、1 単語で構成される場所を表す語は全体の約 66 パーセント、2 単語、3 単語以上の組み合わせで構成されるフレーズが、全体の約 35 パーセントを占めていることがわかる。また、アノテーションすべき対象か迷ったラベル b が付与されたフレーズは 75 パーセント以上が 1 単語で構成されており、アノテーション区間に迷ったラベル c が付与されたフレーズは 60 パーセント以上が 3 単語以上で構成されていた。以降で、各ラベル毎のアノテーション結果を詳細に見ていく。

表 5: 場所を表すフレーズの構成単語数と割合.

曖昧性ラベル	1 単語	2 単語	3 単語以上	合計
a	1928	426	560	2914
(割合)	(0.66)	(0.15)	(0.19)	
b	785	140	100	1025
(割合)	(0.76)	(0.14)	(0.10)	
c	40	39	137	216
(割合)	(0.19)	(0.18)	(0.63)	
b	14	8	25	47
(割合)	(0.30)	(0.17)	(0.53)	
全データ	2767	613	822	4202
(割合)	(0.66)	(0.15)	(0.20)	

表 6 に、ラベル b が付与された発話および対象のフレーズを表す。{} で囲った箇所がアノテーションされたフレーズである。例 1 から 4 のように、一般的で特定の場所すなわち地理的な位置を表すものではない場合、また例 5 から 10 のように、単語単体では場所を表すが、フレーズ全体を見ると場所とは異なる対象を表

表 6: ラベル b が付与されたフレーズの例.

1. 金魚なら昔よく { お祭り } でとってきたのを飼っていました。
2. { お化け屋敷 } は大丈夫ですか。
3. { アウトレット } は時期的に安くなっていました。
4. { ファミレス } より { ファーストフード } に行くことが多いです。
5. 私も、{ 実家 } 暮らしです。
6. { 高所 } 恐怖症なんです。
7. { 中学 } で水泳部でした。
8. よく作るのは { イタリア } 料理とお菓子です。
9. { スペイン } 語も話せるんですか。
10. { ヨーロッパ } 建築で素晴らしいです。

す場合に、アノテーション対象とすべきか迷う傾向にあることがわかる。

次に、ラベル c が付与されたフレーズの例を表 7 に示す。例 1 から 5 のように、「ほど」や「のほう」、「あたり」など、なくても場所を表現するものである場合、また、例 6 から 7 のようにフレーズ全体に固有名詞としての地名が含まれている場合、さらに例 8 から 10 のようにフレーズの修飾関係が曖昧であるような場合に、対象とする区間に迷う傾向があることがわかる。

表 7: ラベル c が付与されたフレーズの例.

1. 今年の夏は時間があつたので、{(国内を三か所) ほど} 旅行しました。
2. {(シカを見に行った公園) かなんか} でぬいぐるみを買いました。
3. {({ スペイン } の方) は人が陽気そうなイメージがあります。
4. {(異人館) のほう} は長らく行ってないですね。
5. {(足立区) の(鹿浜あたり)} に住んでいました。
6. この間({ 下北沢 } の { 店名 A というお店 }) に行ってきました。
7. {(関東) でおいしいといわれているところ} を以前ネットで見ました。
8. 私は(母の { 実家 }) の { 田舎 } 暮らしです。
9. (京都や金沢など { 日本の歴史を感じられる所 }) がおすすめです。
10. { タクシーの運転手さんも (お勧めの場所) } だったんです。

一方で、表 5 より、ラベル a が付与されたフレーズも半数が 2 単語以上で構成されていることがわかる。2 単語以上かつ、固有表現とは異なるフレーズにどのようなものがあるのか表 8 に例を示す。{} 内は実際に場所としてアノテーションされた対象を表し、カッコ () 内は想定される別のアノテーション区間を表す。例 1 から 3 のように、「~の中」や「~の上」等の位置関係を表すものや、例 4 から 5 のように、「~やさん」とい

う口語的にお店を表す表現が付与されているフレーズが含まれる。また、例 6 から 10 のように、その場所の特徴を表す表現や自身の行動やイベントと合わせて表現するフレーズも含まれていることがわかる。

表 8: ラベル a が付与されたフレーズの例.

1. 前に { 電車の中 } で隣の人とおしゃべりしました。
2. { 飛行機の中 } では映画を見て時間を過ごされますか。
3. { ガラスのショーケースの上 } においてありました。
4. { レンタルビデオやさん } にありますか。
5. 暇なときは { 電気屋さん } に行くことが多いです。
6. { 水が美味しいところ } に行きたいです。
7. { 山と海が両方近距離にあるところ } は水が美味しそうです。
8. { 餃子の専門店が集まっているフロア } があります。
9. 確かに { 私の近所の図書館 } も親子連れがたくさんいます。
10. { 今住んでる地方都市 } は福祉が充実してます。

4 場所フレーズ検出

本稿では、固有表現抽出手法として代表的な CRF[4] を用いる。単語区切りに分割した単語列と合わせて、品詞 (Part-of-Speech: POS) タグと場所フレーズのラベルを付与した系列を作成し、学習させる。その手順を次に示す。発話文を入力とし、形態素解析により、単語区切りおよび品詞を付与する。同時に、入力データ中の場所フレーズに対して、場所ラベルを付与する。場所ラベルは、固有表現抽出手法でも使われている、B-LOC と I-LOC を用いる。{} 内の最初の単語には B-LOC を、2 番目以降の単語には I-LOC を付与する。表 9 と図 10 に、入力する発話文を学習データに変換した例を示す。WORD と POS1, POS2 は特徴量、場所ラベルは今回の検出対象のラベルを示す。場所ラベルには、B-LOC と I-LOC の他に、入力文の先頭と終わりを表す START と END ラベル、それ以外であることを表す NL を付与する。

場所ラベルの推定には CRF を用いる。上記の WORD および POS1, POS2 の系列を入力とし、場所ラベルの系列を表す確率が高くなるように学習する。推定においても同様に、入力する発話文から形態素解析を用いて、WORD および POS1, POS2 の系列を生成し、それらを入力として、場所ラベルを推定する。

5 評価

本稿で作成したアノテーションデータを用いて、4 章で示した手法の検出精度を評価する。評価方法は、10 分

表 9: 入力データへの学習データへの変換 (1) .

入力文: 先日スペインに行きました。

WORD	POS1	POS2	場所ラベル
<S>	bos		START
先日	名詞	日時	NL
スペイン	名詞	固有	B-LOC
に	格助詞	*	NL
行き	動詞語幹	*	NL
ました	動詞接尾辞	終止	NL
。	句点	*	NL
<S>	eos		END

表 10: 入力データへの学習データへの変換 (2) .

入力文: 京都の友人宅に行きました。

WORD	POS1	POS2	場所ラベル
<S>	bos		START
京都	名詞	固有	B-LOC
の	格助詞	連体	I-LOC
友人宅	名詞	*	I-LOC
宅	名詞接尾辞	名詞	I-LOC
に	格助詞	*	NL
行き	動詞語幹	*	NL
ました	動詞接尾辞	終止	NL
。	句点	*	NL
<S>	eos		END

割交差検定を用いた。評価尺度は、手法により検出した中で正しくフレーズを検出できた数の割合 (Precision), 全体の正解フレーズの中で手法により正しく検出できた数の割合 (Recall), また precision と recall の調和平均である f 値 (F-measure) を用いる。形態素解析には, JTAG[2] を用いた。比較手法は以下の 2 つである。

CRL IREX で配布された毎日新聞 1995 年にタグ付けされたデータの LOCATION タグが付与されたデータを用いて CRF を学習した場合

PROP 今回作成した学習データを用いて CRF を学習した場合

各手法の検出精度を表 11 に示す。各手法に対して、曖昧性ラベルが a のデータのみを用いた場合の結果と、全データを用いた場合の結果を算出した。

結果より、従来固有表現抽出の対象としてきた学習データでは今回抽出したい場所フレーズは 6 割程度しか検出できないのに対して、提案手法を用いて今回作成したデータの一部を学習することで、検出精度は約 25 ポイント向上していることがわかる。これは、複数の単語から構成されるフレーズや、地名を表す固有名詞として辞書に登録されている単語以外の場所を表すような単語が、学習データ中に含まれていることによる効果であると言える。

表 11: 場所を表すフレーズの検出精度。

	曖昧性ラベル	precision	recall	F-measure
PROP	a	0.91	0.70	0.79
	all	0.89	0.67	0.76
CRL	a	0.58	0.22	0.32
	all	0.66	0.03	0.07

表 12: 検出に成功したフレーズの例。

1. {海} も {山} もあるのでいろいろできました。
2. 最近 {映画館} で見ました。
3. 私は {実家} 暮らしです。
4. {私の近所の図書館} にも親子連れがたくさんいます。
5. 特に {サクラダファミリア} は最高でした。

提案手法により、検出が可能となったフレーズを表 12 に示す。山や海、映画館等は、通常一般名詞や施設名として登録されていることが多いが、本手法を用いることで場所としての検出が可能となっている。また、人の行動やイベントと組み合わせた {私の近所の図書館} 等も場所フレーズとして検出できることがわかった。さらに、検出対象として想定していなかったが、サグラダファミリア → サクラダファミリア等の一文字間違い等の表現は、辞書に登録されていないため固有表現抽出では検出を誤る傾向にあるが、本手法では検出できていた。実際、JTAG による形態素解析の結果は、サクラダ/ファミリアと 2 分割されており、学習データ中に同様に誤った表現が含まれていたために、検出できたと考えられる。雑談等の口語表現を扱う場合には、一文字間違い等により検出できなくなる可能性も考慮することが必要と考えられる。

一方、検出に失敗したフレーズを表 13 に示す。奥多摩や山梨など地名が検出できていないことがわかる。これらの地名等は、CRL において検出の対象とされてきたものであるため、CRL での検出が有益であると考えられる。CRL 学習データとの統合もしくは、PoS ラベルとして CRL の検出結果を用いることで、さらに検出の精度を向上させることが可能であると考えられる。

表 13: 検出に失敗したフレーズの例。

1. 釣りじゃありませんが、以前 {奥多摩} に行きました。
2. あとは、{フィレンツェ} のステーキはオススメです。
3. {山梨} のほうとうもおいしいですよ。
4. {パルマ} や {ポローニャ} は本当においしいものが
5. {店頭} で操作性を比べると良いです。

6 むすび

雑談対話におけるユーザ発話の理解を目的として、場所を表すフレーズに着目して人と人の対話中に現れる語やフレーズの分析および検出を行った。分析フェーズでは、人と人が実際に雑談対話をしたデータに対して、場所を表すフレーズの区間をアノテーションすることにより、検出対象のフレーズの特徴を抽出した。分析の結果より、人は場所を表現する際に、地名等を表す固有表現だけでなく、複数の単語からなるフレーズでの表現や、自身の行動、イベントと組み合わせた表現を用いることがあると分かった。また、今回のアノテーションでは、アノテータによる検出の差が小さかったことから、人が発話を聞いたときにどのフレーズ区間が場所を表すかどうかは一意に定まる傾向にあることがわかった。人は、何らかの同一の基準で、場所を表す単語およびフレーズであることを判断していると考えられる。今回その判断基準の特定には至らなかったが、検出の対象を場所以外の項目に広げる際に、合わせて基準を解明したい。

アノテーションした場所を表す語およびフレーズを検出するにあたり、CRFを用いた検出器を構築した。従来の固有表現抽出に用いられていたデータを用いて学習した場合と、本稿で作成したデータを用いて学習した場合との精度を比較した結果、作成したアノテーションデータを学習に用いることにより、フレーズとしての検出精度が大幅に向上することがわかった。

課題として、今回、2名のアノテータによるアノテーション一致率は、アノテーションしたフレーズが完全一致した場合を対象としており、部分一致した場合については詳細に分析していない。部分一致したフレーズを見ることで、アノテータ間のアノテーションのゆらぎのパターンを分析することができる。加えて、抽出したパターンを用いることで、検出に有効な特徴量を設計できる可能性がある。また、検出した場所フレーズを対話システムに組み込み、実際に人との対話を通して検証および検出の精度を確認する。音声認識誤り等が起こる環境においては、単語単位での特徴量よりも文字単位での特徴量を用いる方がよりロバスト性が高いと考えられる。実環境での実験を通して、手法の改良を行いたい。更に、序論で述べたように、人の発話を正しく理解するためには、場所だけでなく、時間やその他の必要な語やフレーズを理解する必要がある。それらのフレーズについても、本手法の有用性を検証する。

参考文献

- [1] DARPA. *Proc. of the Sixth Message Understanding Conference*. Morgan Kaufmann Publishers, Columbia, MD, USA, 1995.
- [2] T. Fuchi and S. Takagi. Japanese morphological analyzer using word co-occurrence. *Proc. of COLING-ACL*, pp. 409–413, 1998.
- [3] IREX 編集委員会. Irex ワークショップ予稿集. 1999.
- [4] J. Lafferty, A. McCallum, and F. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *Proc. of International Conference on Machine Learning*, 2001.
- [5] A Lee, S. Jung, K. Kim, D. Lee, and G. G. Lee. Recent approaches to dialog management for spoken dialog systems. *Journal of Computing Science and Engineering*, Vol. 4, No. 1, pp. 1–22, 2010.
- [6] J. E. Weston. Dialog-based language learning. *Proc. of Advances in Neural Information Processing Systems 29 (NIPS 2016)*, 2016.
- [7] 橋本泰一, 乾孝司, 村上浩司. 拡張固有表現タグ付きコーパスの構築. 情報処理学会研究報告自然言語処理 (NL), Vol. NL-188, pp. 113–120, 2008.