

## $TF \cdot AoI$ を用いた類似性による発話予測

木村 泰知<sup>†</sup> 荒木 健治<sup>†</sup>

<sup>†</sup> 北海道大学 大学院 情報科学研究科, 札幌市

あらまし 我々は、類似度に基づいて発話単位の予測を行う手法を提案する。音声対話では、背景知識、文脈、対象範囲の限定により、発話を予測することが考えられる。n-gramのような線形予測では、文字単位あるいは単語単位による予測となるが、発話単位の予測を行う手法が望まれる場面も多い。そこで、我々は低頻度に基づく類似度を利用して学習データから類似度の高い文を見つけ、その次に発話された文を予測文とする。類似度は $TF \cdot AoI$  (Term Frequency  $\times$  Amount of Information)の重み付けに基づいて計算される。一致率を類似度と考え比較実験を行った結果、提案手法が候補選択において5ポイントの向上が確認された。

キーワード 発話予測, 類似度,  $TF \cdot AoI$

## Prediction Using Similarity based on $TF \cdot AoI$

Yasutomo KIMURA<sup>†</sup> and Kenji ARAKI<sup>†</sup>

<sup>†</sup> Graduate School of Information Science and Technology, Hokkaido University, Sapporo-shi

**Abstract** We propose the prediction method using similarity based on  $TF \cdot AoI$ . In spoken dialogue, we generally use background, context and limited domain for prediction. Although n-gram is words dictation, we hope the prediction for utterance. Therefore we decide the prediction which is next utterance of the highest similarity utterance in a learning corpus. Our similarity is calculated by using weight based on  $TF \cdot AoI$  (Term Frequency  $\times$  Amount of Information). We confirmed 5 points improvement compared to the matching rate measure.

**Key words** Prediction, Similarity,  $TF \cdot AoI$

### 1. はじめに

発話の予測は、対話処理及び音声認識において有用である。一般的に、予測は単語のような短い単位で行われ、n-gramも単語単位の予測となる。n-gramは統計を用いた音声認識の言語モデルとして利用されている[1][2][3]。音声認識で言語モデルを作成する場合、パープレキシティが低いほど良い。パープレキシティは平均分岐数であるため、音声認識候補が絞られていれば、単語候補が減少する。評価する言語データから言語モデルを作成した場合、認識単語の候補は少なくなる。つまり、次回発話される内容を予測可能となれば、パープレキシティは小さくなる。本稿では、単語単位ではなく発話単位の予測を扱う。話し言葉において、単語単位で予測するには、書き言葉よりもフィラー、言い誤りが含まれるため、時系列の予測では困難な場面が多い。さらに、音声対話では相手の発話が理解されれば協調的な応答にも利用可能となる。協調的な応答は、相手の考えている内容を考慮した応答を行うことである[4][5]。つまり、ユーザが次に発話する内容を予測することで、ユーザの求めている内容を先回りして応答する。

このように発話単位あるいは文単位の予測が可能となれば、

解決できる問題も少なくない。しかしながら、話し言葉を予測する場合には、書き言葉の文法では対応できない場面も数多く存在する。そこで、話し言葉であるフィラー及び倒置にも対応するために、本手法では低頻度語に焦点を当て、学習データから計算した頻度により各単語に重み付けを行い、類似度を計算する[6][7]。本手法では、ファイラーのような雑音に影響されずに、類似性を測定することが可能である。最も類似性の高い文の次の文が、予測文となる。本稿では、本手法に基づいた手法による実験システムの構築を行い、実験を行った。その実験結果の詳細と有効性について述べる。

### 2. 基本的な考え方

人間が音声認識する場合、認識対象の音声だけではなく、背景知識、状況、文脈を利用している。これらの知識を利用することで、発話内容を制限している。発話内容の制限は、発話を予測しているとも考えられる。この予測は以前に経験した会話から考えられる内容を予測することになる。しかしながら、n-gramのような線形予測では、学習データ中に含まれる雑音に影響される。対話の予測で求めていることは、次回の発話される内容である。つまり、単語単位ではなく文あるいは発話単

位の予測が望まれる。しかしながら、今までの予測方法は単語単位である。特に、話し言葉の場合、全てを正確に予測するよりも発話内容の一致が望まれる。

そこで、本稿では類似性に基づいた発話予測を行う。本稿で予測とは、入力された発話と同じ内容の発話が過去に存在した場合、それ以降の会話も同じような展開になると仮定している。したがって、学習データは時系列に並んでいる必要がある。しかし、会話の内容が全く同じであることはほとんどないため、同じ会話の判定を類似度に基づいて評価する。単語単位の予測と異なる点は、単語単位の時系列ではなく、発話単位の時系列により予測を行う点である。図1に予測の例を示す。例えば「おはようございます…」のような発話が存在した場合、学習データ中に存在する最も類似した文の次の文「先日予約したんですけど…」を予測文とする。

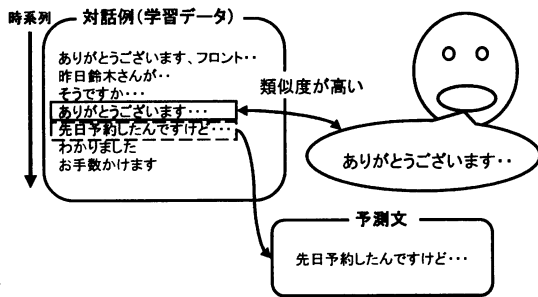


図1 予測方法

### 2.1 TF・AoIによる類似度

本手法では、ベクトル空間モデルの表現形式を利用してユークリッド距離を計算する。我々はユークリッド距離を測定する要素として、「文における単語頻度 × 情報量」を使う。これを、 $TF \cdot AoI$  (Term Frequency × Amount of Information) と呼ぶ。情報量は式(1)で定義される。

$$AoI = -\log_2 P(x) = -\log_2 \frac{f(x)}{N} \dots (1)$$

ここで、 $N$ は総単語数、 $f(x)$ は $x$ がデータ中に出現する頻度である。

まず、 $TF \cdot AoI$ 及び、ユークリッド距離の算出方法と対話処理への適用について説明する。最初に、基準となる文(以降、基準文とする)に対して単語の分割を行う。茶筌[9]により単語分割を行い、単語分割された基準文の「異なり単語数」を $n$ として、 $n$ 次元のデータとする。分割された単語がその文中に出現する頻度(Term Frequency以降 $TF$ )を求める。基準文中に出現する「異なり単語」を $(T_1, T_2, \dots, T_n)$ とした場合、 $TF$ は $(TF_1, TF_2, \dots, TF_n)$ になる。基準文中の出現単語数だけを要素とするため、基準文の $TF$ は全て出現頻度が1回以上となる。重み付けした値、すなわちは $TF \cdot AoI$ は、 $(TF_1 \cdot AoI(T_1), TF_2 \cdot AoI(T_2), \dots, TF_n \cdot AoI(T_n))$ になる。ここで、類似度を測定する対象となる文(以降、

対象文とする)を $Sent_A$ とする。対象文 $Sent_A$ の $TF_a$ は $(TF_{a1}, TF_{a2}, \dots, TF_{an})$ である。基準文と対象文の重み付けされた $TF \cdot AoI$ の値からユークリッド距離を計算する。ユークリッド距離を下記に示す。

$$D_{input \rightarrow data} = \sqrt{\sum_{i=1}^n (TF_i \cdot AoI(T_i) - TF_{ai} \cdot AoI(T_{ai}))^2} \dots (2)$$

$$D = D_{input \rightarrow data} + \alpha \times D_{data \rightarrow input} \dots (3)$$

ここで、 $D_{input \rightarrow data}$ と $D_{data \rightarrow input}$ の違いは、入力文を基準文とした場合と学習データ中に含まれるデータを基準文とした場合の違いである。つまり、基準文の異なり単語数の違いである。 $N$ は学習データ中に存在する総単語数である。類似度を式(4)に示す。

$$Similarity = \frac{1}{D+1} \dots (4)$$

2文が同一文である場合、ユークリッド距離の値は0になるため、最も高い類似度は1となる。

## 3. 比較実験

### 3.1 実験の目的

話し言葉における対話の中で、類似性を用いて予測を行う。本実験では、 $TF \cdot AoI$ による予測手法、 $TF \cdot IDF$ による予測手法、入力文と対象文の一致率による予測手法の比較を行う。

### 3.2 実験方法

予測文は、入力文と学習データの類似度の測定を行い、最も類似している文の次に出現する文を予測文とする。比較する手法は、 $TF \cdot IDF$ による重み付けを行った手法と一致度による手法である。 $TF \cdot IDF$ による重み付けを行ったユークリッド距離を式(5)、その類似度を式(6)に示す。

$$D_{TFIDF} = \sqrt{\sum_{i=1}^n (TF_i \cdot IDF(T_i) - TF_{ai} \cdot IDF(T_{ai}))^2} \dots (5)$$

$$Similarity = \frac{1}{D_{TFIDF} + 1} \dots (6)$$

一致度による評価を式(7)に示す。

$$MATCH = \frac{\text{入力文と比較する文の一致した単語数}}{\text{入力文の単語数}} \dots (7)$$

実験データは話し言葉によるコーパスであるATRのSLDB[8]を用いた。学習データを表1に示す。評価データは学習データ以外のデータを用いるためオープンデータとなる。

	学習データ
発話数	10,512
総単語数	17,4519
異なり単語数	4,239

### 3.3 実験データの収集

本手法は音声対話処理の利用を目的としているため、音声対話を書き起こした ATR の SLDB [8] を利用する。データは受付とお客の会話から構成されている。このデータには、パターン化されている箇所もあり、対話形式になっているため、音声対話及び予測の観点から適切なデータと考えられる。しかし、予測困難な対話も含まれており、1 回しか出現しないような単語も存在する。異なり単語からみた場合低頻度語が数多く含まれており [10]、評価データにこのような単語が含まれている場合、学習データ中に正解が含まれないことになる。そこで、本実験では予測結果の正解文と同一の内容が学習データ中に含まれていることとした。さらに、「はい」、「そうですね」のように少ない単語数で構成されている文は、単語予測と変わらなくなり、発話単位の予測として意味がなくなるため、対象外とした。他にも、次の発話の長さが短いのか長いかわからないまま予測するのは困難であるため、正解の単語数の範囲を与えている。これは音声認識の場合、発話時間や音素数から正解単語数を知ることが可能という理由からである。例えば、正解単語数が 10 単語の場合には 5 単語から 15 単語までの予測結果を選択することとした。対象外とした条件を下記に示す。

- 正解が学習データに存在する
- 正解が 5 単語以上で構成されている
- 予測結果と正解の単語数の差が 5 単語以上の文は候補としない

### 3.4 F 値による評価

本節では、予測内容が一致していることを評価する以外に客観性による評価として、F 値を用いる。再現率、適合率、F 値の式を以下に示す。

$$\text{再現率} = \frac{\text{正解予測単語数}}{\text{正解単語数}}$$

$$\text{適合率} = \frac{\text{正解予測単語数}}{\text{予測単語数}}$$

$$F \text{ 値} = \frac{2 \times \text{再現率} \times \text{適合率}}{\text{再現率} + \text{適合率}}$$

音声認識及び協調的な応答を考えた場合、予測文が 1 文である必要は無く、関連性のある文を選択できれば良い。そこで、予測候補文を 10 文までとした場合に最も良い結果を評価する。

### 3.5 実験結果

実験では、最も類似している一文の次の発話を予測文として評価した。評価データは学習データと異なる会話である。その会話例の中から、学習データ中に正解が存在する 200 文を抽出した。その結果を表 2 に示す。TF・AoI は提案手法、TF・IDF は重み付けが TF・IDF による手法、MATCH は一致率による手法である。TF・AoI による手法が最も良い結果であった。

表 2 比較実験の結果

手法	$\alpha$	候補 10 位以内	再現率	適合率	F 値
MATCH	-	50	0.7555	0.7199	0.7373
TF・AoI	0.0	54	0.7476	0.7299	0.7387
TF・AoI	0.0	53	0.7312	0.7238	0.7275

最も高い値であった TF・AoI について、さらに向上させる方法を検討する。本実験では、1 発話単位で比較を行ったが、比較の単位を変更することで向上する可能性がある。そこで、次節では 1 発話単位ではなく 2 発話単位での比較及び、発話直前の利用する単語数の比較を行う。

## 4. 性能向上に関する実験

### 4.1 属性数の比較実験

#### 4.1.1 実験の目的

類似度は比較する単位及び係数により結果が異なるため、本実験では予測に適切な設定を求める。

#### 4.1.2 実験方法

本実験では、2 発話単位で類似度を測定する場合と、予め予測に利用する単語数を決めておき、属性数を決定する場合の実験を行う。前節の実験では、1 発話を基準文としており、基準文の異なり単語数を属性数としていたが、本実験では適切な単語数を求める。そこで、扱う単語数を変化させ、類似度の測定を行う。まず、2 発話の単位の類似性を測定する場合の実験を行い、次に予測に用いる単語数を予め決めておいた場合の実験を行う。単語数は、直前に発話された文末の単語からかぞえ、2 単語間隔で、4 単語から 30 単語まで変化させる。評価データは 200 文である。

#### 4.1.3 実験結果

2 発話単位は文脈を利用していることになるため、予測精度が向上すると考えられたが 1 発話よりも低い結果となった。ここで発話とは受付とお客の 1 回の発言である。つまり、1 発話単位で比較した場合は、受付の発話パターン及びお客の発話パターンの情報も考慮しているため、1 発話単位の予測の場合、お客の発話があれば次回は受付の発話であるということが情報として含まれているため、向上したと考えられる。このような要因から、一文単位の予測が良い結果となった。

表 3 に属性数と係数を変化させた比較実験の結果を示す。

表 3 属性数と係数

手法	係数 $\alpha$	再現率	適合率	F 値
TF・AoI	0.0	0.7476	0.7299	<u>0.7387</u>
TF・AoI 2 発話	0.0	0.7124	0.6840	0.6979
TF・AoI 2 発話	0.1	0.7116	0.6858	0.6985
TF・AoI 2 発話	0.2	0.7077	0.6851	0.6962
TF・IDF 2 発話	0.0	0.7265	0.6831	0.7041
TF・IDF 2 発話	0.1	0.7045	0.6734	0.6886
TF・IDF 2 発話	0.2	0.7030	0.6937	0.6983

### 4.2 係数 $\alpha$ の比較実験

#### 4.2.1 実験の目的

TF・AoI の重み付けによる類似度では、基準文の設定や係数の付与により計算結果が異なる。本実験では、基準文の設定と係数の値を変更することで精度向上を試みる。まず、基準文の設定によるユークリッド距離の違いは下記のように  $D_1, D_2$  の 2 つが考えられる。

$D_1$  入力文を基準文とした場合

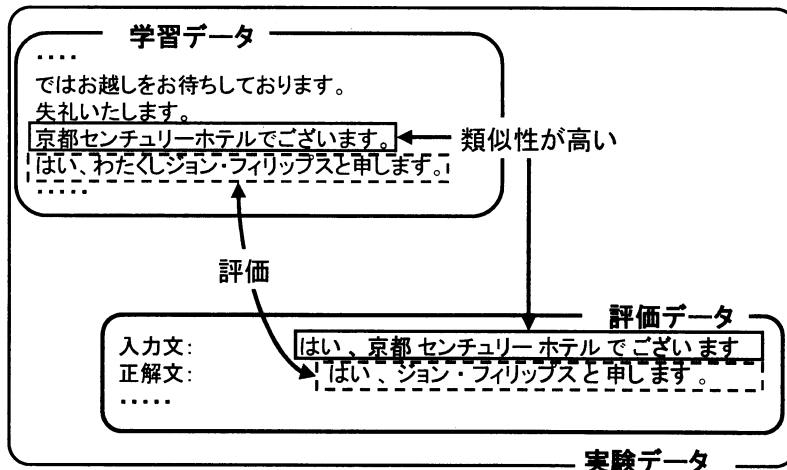


図 2 実験方法

$D_2$  コーパス中の各質問文を基準文とした場合  
類似度の計算結果も異なることから、 $D_1$  と  $D_2$  の計算結果の合計に係数  $\alpha$  を付与した式を  $D_1 + \alpha \times D_2$  として、この係数  $\alpha$  の最適値を求める。

#### 4.2.2 実験結果

係数  $\alpha$  の比較実験の結果を表 4 に示す。  $\alpha = 0.1$  が最適値となり、図 3 で示すように予測候補 10 位以内に完全一致の正解が含まれた数は 200 セット中 60 個存在した。F 値は 0.7488 となり、最も高い結果であった。MATCH 手法と F 値の比較において、ほとんど向上はみられなかった。しかし、完全一致の正解が 25% から 30% に向上しており、5 ポイントの向上が確認された。この結果から比較対象の単語を 1 割程度考慮した方が良いことがわかる。

表 4 係数  $\alpha$  の比較実験

手法	$\alpha$	候補 10 位以内	正解率	再現率	適合率	F 値
MATCH	-	50	25.0%	0.7555	0.7199	0.7373
$TF \cdot AoI$	0.0	54	27.0%	0.7476	0.7299	0.7387
$TF \cdot AoI$	0.1	<b>60</b>	30.0%	0.7641	0.7342	<b>0.7488</b>
$TF \cdot AoI$	0.2	55	27.5%	0.7375	0.7172	0.7272
$TF \cdot AoI$	0.3	49	24.5%	0.7179	0.7046	0.7112
$TF \cdot AoI$	0.4	43	21.5%	0.7085	0.6869	0.6975
$TF \cdot AoI$	0.5	37	18.5%	0.6912	0.6543	0.6723
$TF \cdot AoI$	0.6	36	18.0%	0.6732	0.6391	0.6557
$TF \cdot AoI$	0.7	33	16.5%	0.6638	0.6192	0.6407
$TF \cdot AoI$	0.8	30	15.0%	0.6536	0.6096	0.6309
$TF \cdot AoI$	0.9	27	13.5%	0.6513	0.5983	0.6236
$TF \cdot AoI$	1.0	24	12.0%	0.6395	0.5837	0.6103

次に比較単語数の比較実験の結果を表 5 に示す。比較する単位を予め単語数で設定した場合である。係数は前回の実験結果から  $\alpha = 0.1$  の場合とした。最も良い結果は、発話直前から 14 単語前までを入力文として比較した場合であった。しかしながら、1 発話単位で比較した場合を下回った。

候補10位以内に完全一致の正解が含まれた数

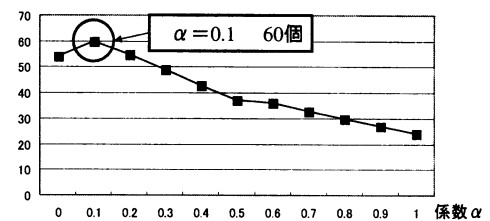


図 3 評価データ 200 文における係数  $\alpha$  の比較実験

表 5 属性数の比較実験

手法	単語数	$\alpha$	再現率	適合率	F 値
$TF \cdot AoI$	1 発話	0.1	0.7641	0.7342	<b>0.7488</b>
TFAOI	4	0.1	0.6426	0.5916	0.6161
TFAOI	6	0.1	0.6693	0.6436	0.6562
TFAOI	8	0.1	0.6912	0.6774	0.6843
TFAOI	10	0.1	0.7006	0.6687	0.6843
TFAOI	12	0.1	0.6920	0.6866	0.6893
TFAOI	14	0.1	0.7092	0.6903	<b>0.6997</b>
TFAOI	16	0.1	0.6881	0.6875	0.6878
TFAOI	18	0.1	0.6904	0.6787	0.6845
TFAOI	20	0.1	0.6865	0.6562	0.6710
TFAOI	22	0.1	0.6936	0.6624	0.6776
TFAOI	24	0.1	0.6975	0.6662	0.6815
TFAOI	26	0.1	0.6983	0.6489	0.6727
TFAOI	28	0.1	0.6936	0.6629	0.6779
TFAOI	30	0.1	0.6904	0.6689	0.6795

#### 4.2.3 考察

$TF \cdot AoI$  は 1 発話単位で係数  $\alpha = 0.1$  の結果が最も良かったため、この実験結果と MATCH の実験結果から比較を行う。  $TF \cdot AoI$  と MATCH が共に正解であった予測の割合は 18%(36/200), MATCH のみが正解であった予測は 7%(14/200),  $TF \cdot AoI$  のみが正解であった予測は

表 6 正 解 例

入力文 正解文	それではお客様のお名前とご連絡先をお願いしますか。 鈴木直子です。
予測結果 [1]	はい、メアリー・フィリップスです。
予測結果 [2]	はい。
予測結果 [3]	トーマス・ネルソンです。
予測結果 [4]	はい、ケーアンドケー運輸 ニューヨーク 支店 です。
予測結果 [5]	三〇八号室の田中です。
予測結果 [6]	はい。
予測結果 [7]	はい。鈴木直子です。
予測結果 [8]	はい、鈴木直子といます。
予測結果 [9]	ジョン・フィリップスといます。
予測結果 [10]	鈴木直子です。
入力文 正解文	かしこまりました。お客様は何名様でいらっしゃいますか。 二人です。
予測結果 [1]	わたしと妻の二人です。
予測結果 [2]	三人です。
予測結果 [3]	わたし一人です。
予測結果 [4]	わたしと妻の二人です。
予測結果 [5]	二人でお願いします。
予測結果 [6]	二人です。
入力文 正解文	はい、ありがとうございます。ニュー 京都 ホテルで ございます。 こんにちは。ジョン・フィリップスと申します。
予測結果 [1]	四条河原町のニュー 京都ホテルで間違いないでしょうか。
予測結果 [2]	四条河原町のニュー 京都ホテルで間違いありませんね。
予測結果 [3]	宿泊の予約をお願いしたいのですが。
予測結果 [4]	ホテルの予約をお願いしたいのですが。
予測結果 [5]	こんにちは。ジョン・フィリップスと申します。

12%(24/200),TF・AoI と MATCH が共に予測を行えなかった結果は 63%(126/200) であった。予測の正解例を表 6 に示す。入力文が「それではお客様のお名前とご連絡先をお願いしますか。」の場合の予測の正解は「鈴木直子です。」であった。この入力文に対して、候補 10 位まで出力した結果を示す。第 10 位の候補に完全一致の正解が存在した。名前の予測は困難であるが、学習データに存在した場合、予測は可能となる。他にも、入力文が「かしこまりました。お客様は何名様でいらっしゃいますか。」の場合、予測候補第 6 位に正解が含まれていた。このような人数を当てることは難しいと考えられるが、状況に依存した状態であれば、ある程度的人数でも予測できると考えられる。

次に誤りの分析を行う。どちらの手法でも完全一致の正解が含まれなかった場合の 126 文に対して、4 つの主な誤りパターンの特徴に分類する。

- A 多頻出の発話であるため、類似候補が多くなり選択が困難
- B 表現が異なるが、正解に近い
- C 類似文が存在しないため予測が困難
- D パターン化されているが、完全一致の予測は困難

この結果を図 4 に示す。複数パターンを選択を可能として、126 文を評価した。

誤りパターン A は、多頻出語から構成されているため、同じ発話が多いパターンである。類似候補が多くなり、入力文が情報を与えていない状態である。例えば、入力文が「かしこまり

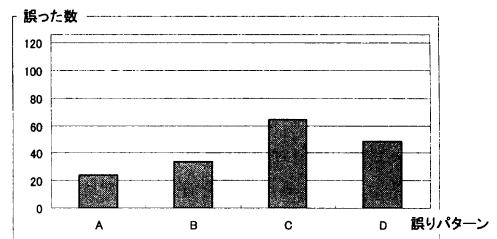


図 4 誤りの分類

ました。」である。「かしこまりました。」は高頻度で出現するため、同じ類似度の候補が複数存在する。このような場合には、前文を考慮して予測する必要がある。

誤りパターン B は正解と似ている内容を出力しているが、表現が一致していない場合である。つまり、これは正解と判断してよい結果であるが、今回は不正解とした。パターン B は 126 セット中 34 個存在した。例えば、入力文が「ありがとうございました。では鈴木様、二十四日にお待ち申し上げます。」正解文が「じゃ、よろしく申し上げます。」という場合である。パターン C は、予測が非常に困難なパターンである。例えば、入力文が「えご予約の際にクレジットカード番号を伺っておりますので、えそちらの口座の方に請求させていただきます。」のような場合である。ここで下線部はフィラーである。入

表7 誤りのパターン

パターン A	
入力文	かしこまりました。
正解文	ええ、それをお願いします。
パターン B	
入力文	ありがとうございました。では鈴木様、二十四日にお待ち申し上げております。
正解文	じゃ、よろしくお願いします。
パターン C	
入力文	えご予約の際にクレジットカード番号を伺っておりますので、えーそちらの口座の方に請求させていただきます。
正解文	分かりました。
パターン D	
入力文	はい、何名様でいらっしゃいますか。
正解文	四人です。
予測候補	七人です。
予測候補	わたしと妻の二人です。
予測候補	わたしと妻の二人です。
予測候補	わたし一人です。
予測候補	そうですねサンドイッチはありますか。
予測候補	三人です。
パターン C+D	
入力文	はい。和室のほうはお値段が一万はっぴゃ、失礼いたしました一万八千円となっておりますが、よろしいでしょうか。
正解文	はい、それで結構です。

力文は長いが予測に必要な情報が不足している場合は困難と考えられる。パターンDは、パターン化されているが、完全一致の予測は困難な場合である。例えば、入力文が「はい、何名様でいらっしゃいますか。」の場合、人数を当てることは背景知識がなければ困難である。ここでは、正解が「四人です。」、予測候補は「七人です。」、「わたしと妻の二人です。」、「わたしと妻の二人です。」などがある。人数を答える点では正解となるが、完全一致した予測候補は存在しなかった。パターンC+Dは、内容は困難であるが、文末表現から判断することが容易な場合である。例えば、入力文が「はい。和室のほうはお値段が一万はっぴゃ、失礼いたしました一万八千円となっておりますが、よろしいでしょうか。」の場合、Yes/No 質問に対して正解の場合には容易と考えられる。この入力に対する正解は「はい、それで結構です。」であった。しかし、質問に対する応答がNoである場合に、訂正内容を予測することは困難である。

これらの結果から、予測可能な場面と予測が困難な場面を切り分ける必要がある。パターンDのように文末表現や質問タイプで次の発話パターンを予測することが大きな役割となる場面もある。さらに、評価方法も問題の難易度により変更する必要がある。今回は学習データ中に正解文が含まれているという問題設定であったため、多頻出語から構成された文が評価データとなることが多かった。しかしながら、簡単な文を予測できるだけでは有効な範囲も限定されてしまうので、今後は未知語を含んだ評価実験も行う必要がある。

## 5. まとめ

予測手法は、音声認識における言語モデル及び協調的応答のように相手の意図を予測して応答するような音声対話処理の場において有効と考えられる。今まで、予測手法は単語単位の予測手法が多く、発話単位の予測を行うことが考えられていなかった。そこで本稿では、 $TF \cdot AoI$ を用いた類似性による発話予測手法を提案した。本手法は学習データに依存することなく重要単語に大きな重みを付与するために、低頻度語を重要単語として重み付けを行った。正解が予測候補に含まれる精度の比較実験を行った結果、本手法は一致率で類似度を測定した手法に対して5ポイントの向上が確認された。

しかしながら、予測を行うためには文末表現や質問タイプで次の発話パターンを判断することも重要である。本手法では低頻度語を重視していたが、文末の表現は高頻度語で構成されているため、低頻度語と高頻度語の使い分けを考える必要がある。さらに、類似している文の次の文が必ずしも予測に適している文であるとは限らない。そこで、今後は文タイプの分類及び入力文と予測文を構成している単語の関連性を考慮する予定である。他にも、発話単位の予測が可能な状況と不可能な状況が存在するため、その条件を明らかにしていく予定である。

## 文 献

- 森 伸介, 宅間 大介, “生コーパスからの単語 N-gram 確率,” 情報処理学会研究会報告 2004-NL-162, pp.183-190, 2004.
- Sehyeong Cho, “Improvement of Language Models Using Dual-Source Backoff,” PRICAI 2004: Trends in Artificial Intelligence.
- Langzhou Chen, Jean-Luc Gauvain, Lori Lamel, and Gilles Adda, “Dynamic Language Modeling for Broadcast News,” ICSLP2004.
- 伊藤克亘, “自然言語処理 -基礎と応用-,” 田中穂積監修, pp.302-322, 電子情報通信学会, 東京, 1999.
- 伊藤 敏彦, 小暮 悟, 中川 聖一, “協調的応答を備えた観光案内音声対話システムとその評価,” 情報処理学会論文誌, Vol.39, No.5, pp.1248-1257, 1998.
- 木村 泰知, 荒木 健治, 柄内 香次, “ $tf \cdot AoI$ の重み付けに基づく類似性を用いた話し言葉における質問文の同定法,” 電子情報通信学会論文誌 D-II, Vol.J87-D-II, No.8, pp.1653-1664, 2004.
- 相澤 彰子, “低頻度語の利用によるテキスト分類性能の改善と評価,” 情処学論, Vol.44, No.7, pp.1720-1730, 2003.
- T. Morimoto, N. Uratani, T. Takezawa, O. Furuse, Y. Sobashima, H. Iida, A. Nakamura, Y. Sagisaka, N. Higuchi and Y. Yamazaki, “A speech and language database for speech translation research,” Proc. ICSLP'94, pp.1791-1794, 1994.
- “日本語形態素解析システム『茶釜』 version 2.2.1 使用説明書”, 松本裕治, 北内啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原 正幸 Dec, 2000.
- 辻 慶太, 芳鐘 冬樹, 影浦 峯, “対訳コーパスにおける低頻度語の性質 訳語自動抽出に向けた基礎研究,” 情処学, NL 研報, NL-138, pp.47-54, 2000.