

TR-IT-0262

発話状況の情報を音声認識に用いるための  
予備実験

Preliminary Experiments for the Application of  
Contextual Information for Speech Recognition

荒川直哉

ARAKAWA Naoya

1998年5月

**Abstract**

This report presents preliminary experiments for the application of contextual information to speech recognition. Contextual information here includes speaker roles, speech act types and particular expressions in utterances. The author reports on the statistical identification of 1) speech act types, 2) speech act type cue expressions, and 3) content words (non-functional words). The ATR corpus of travel conversation (SLDB, the Japanese data) was used for the experiments.

© (株)ATR 音声翻訳通信研究所 1998

© 1998 by ATR Interpreting Telecommunications Research Laboratories

# 目次

第1章 序論 .....	1
1.1 話者の役割情報の利用 .....	1
1.2 発話タイプおよび相関する表層表現の利用 .....	1
1.3 話題・内容語の利用 .....	1
第2章 発話タイプ推定 .....	2
2.1 発話タイプコーパス .....	2
2.2 発話タイプ推定 .....	2
2.3 エントロピーとパープレキシティ .....	2
2.4 発話タイプパターンを使用しない条件での推定実験 .....	3
2.5 発話タイプパターンを使用する条件での推定実験 .....	6
第3章 発話タイプパタンの予測 .....	9
3.1 発話タイプパターン .....	9
3.2 発話タイプパターン推定 .....	9
3.3 実験 .....	10
3.4 まとめ .....	12
第4章 内容語の予測 .....	13
4.1 内容語の推定 .....	13
4.2 実験 .....	14
4.3 まとめ .....	15
文献 .....	15
付録 .....	16
付録1 発話タイプ（31種類）一覧 .....	16
付録2 ファイルリスト（計199ファイル） .....	18
付録3 前後する発話タイプの相関 .....	19
付録4 先行発声の発話タイプと発話タイプパターンとの相関 .....	20
付録5 先行発声の発話タイプパターンと発話タイプパターンとの相関 .....	21
付録6 前後する発話の内容語の相関 .....	22

## 第1章 序論<sup>1</sup>

特定の言語表現が現われる確率はそれらが用いられる文脈に依存すると考えられる。例えば、金額が話題になっている場合には「XX円」などの金額の値に関する表現が現われやすいであろう。音声認識では認識結果候補である言語表現の出現確率を候補のスコア付けに用いるので、こうした出現確率の文脈依存性の利用が可能である [1,2,5,6]。このレポートでは言語表現の出現確率の文脈依存性を調査し、発話状況の情報がどの程度音声認識の性能向上に寄与するかを判断するための資料を提示する。ここで発話状況の情報として考えるのは、話者の役割（サービス提供者か享受者か）、先行する発話列あるいはその発話タイプ（<question> <inform> など）、および話題である。また、調査のためのコーパスとしてはATR音声翻訳通信研究所が作成している旅行会話の「音声言語データベース」（以下SLDB）を使用する。

### 1.1 話者の役割情報の利用

以下で使用するSLDBコーパスにおいて、話者の役割とは旅行サービスの申込者かサービス提供側の担当者かという区別のことである。担当者と申込者は特定の言語表現の使用頻度が異なる（例えば担当者の方が丁寧な表現を用いる）と考えられる。

### 1.2 発話タイプおよび相関する表層表現の利用

ここで発話タイプとは、発話の機能あるいは発話によって話者が行おうとする行為の分類を指す。このレポートでは発話タイプとして <yn-question> <inform> など31種類を用いている（付録1）。日本語の場合、発話タイプは文末の表層表現と関連がある（例えば <yn-question> に分類される発話は「か」で終わることが多い）ため、音声認識の対象となる発話がどのような発話タイプを持つかの確率を計算することができれば、その発話の文末表現が現れる確率を計算することが可能になる。発話タイプの出現確率は先行する発話の文末表現から統計的に求められる。推定された発話タイプに基づいて計算される文末表現の出現確率は音声認識候補の出現確率を計算するために用いられ、結果的にその再順序づけに利用される。

日本語では文末表現と発話タイプが関連するのであるから、発話タイプの推定を行わず、前発話の文末表現から現発話の文末表現を推定するという考えられる。このレポートではこの方法に関する実験も行い、発話タイプ推定による文末表現の推定と比較する。

### 1.3 話題・内容語の利用

特定の話題について話している場合、特定の語彙の出現頻度が高くなると考えられる。例えば病気が話題になっている場合には「体温計」といった単語が現われやすいであろう。話題に関連して出現頻度が変わるのは主に自立語などの内容語であろう。前発話の内容語から話題を推定し、そこから現発話の内容語の出現頻度を計算するという方法をとることもできるし [4]、前発話に現れる内容語から直接現発話の内容語の出現確率を計算するという方法を取ることもできる。このレポートでは後者の方法について実験を行った。

---

<sup>1</sup> 謝辞：本レポート中のいくつかのアイデアは妹尾正身氏（現 NTT ソフトウェア）によるものです（Ref. 「統計による次発話予測プログラム」ドキュメント（/dept4/work5/senoo/9708/doc/\*.tex）（1997））。また、田中英輝、岩本秀明両氏、横尾昭男室長をはじめコメントをいただいたATR音声翻訳通信研究所第4研究室の諸姉兄に感謝します。

## 第2章 発話タイプ推定

現発話の発話タイプを先行する発話の情報や現発話の表層表現（発話タイプパターン [下記参照]）などを用いて推定する実験を行った。発話タイプの予測は次章（3章）の発話タイプパターンの予測において用いられる。発話タイプの推定については例えば文献 [7]（日本語に関する調査）あるいは文献 [3]（英語に関する調査）を先行研究としてあげることができる。

### 2.1 発話タイプコーパス

以下の実験で使用する発話タイプコーパスはSLDBの199ファイル（付録2）の各発話（場合によってさらに細かい単位）に人手で発話タイプを付与したものである<sup>2</sup>。発話タイプの一覧を付録1に付す。発話タイプコーパスでは、発声（句点「。」で区切られる単位）を、主に「分割用表現パターン」を用いて「発話タイプ単位」に分割し、分割した「発話タイプ単位」に対してその「発話タイプパターン」（表層表現、主に文末表現）および文脈を参照して発話タイプを付与している。「発話タイプパターン」は、発話タイプを推定する手掛かりとなる表層表現である。日本語の場合、発話タイプパターンは主に文末表現であるが、<wh-question> を判定する場合には疑問詞を含み、<greet> などの場合、発話全体を発話タイプパターンとすることもある。

発話タイプコーパスにおいて、発話タイプ単位は以下の形式で記述されている。実験の学習データでは ENTRY を発話タイプパターンとし、LABEL を発話タイプとした<sup>3</sup>。

```
N_FILE : TBS23003
N_HATUWA : 047000
N_TANI : 047001
SPEAKER : 通訳者
UNIT-UTTR : それから|、|あらかじめ|何|か|予約|を|証明|でき|る|もの|を|頂け|ます|か|。
COMM : /
AMBIGUITY : ("*|ま|す|か" "*|何|か|*|ま|す|か" "*|何|か|*|頂け|ます|か")
ENTRY : *|何|か|*|頂け|ます|か
LABEL : (<action-request>)
```

### 2.2 発話タイプ推定

現発話の発話タイプを、さまざまな発話状況を条件とする条件付き確率を用いて推定する。現発話の発話タイプを  $ift$ 、発話状況の条件を  $D$ 、条件付き確率を  $p(ift|D)$  とすると、推定される  $ift$  は与えられた条件  $D$  に対し  $p(ift|D)$  を最大にするような  $ift$  である。条件付き確率はコーパスから計算する。

### 2.3 エントロピーとパープレキシティ

事象空間  $S$  の事象  $s$  が起こる確率を  $p$  とした場合のエントロピー  $H(S) = - \sum_{s \in S} p(s) \log_2 p(s)$  は、事象空間  $S$  の平均情報量であり、パープレキシティ  $2^{H(S)}$  は事象の複雑さを表すと考えられる。従って、一般的に  $S$  のパープレキシティが少ないほど  $s \in S$  の予測は容易になると考え

<sup>2</sup> /dept4/work5/IR\_IFT/doc/collect.doc

<sup>3</sup> LABEL に多義性がある場合、例えば (<acknowledge> <yes>) の場合、0.5 回ずつ生起するとみなした。

られる。一般的に  $s \in S$  を予測するために  $s$  と関連する条件  $t$  (事象空間  $T$  の事象) による条件付き確率  $p(s|t)$  を用いるとパープレキシティは低下する。条件付き確率  $p(s|t)$  のエントロピーは  $H(S|T) = - \sum_{t \in T} \sum_{s \in S} p(st) \log_2 p(s|t) = H(ST) - H(T)$  により与えられる (ここで  $H(ST)$  は  $S$  と  $T$  の共起のエントロピーであり、 $H(T)$  は  $T$  のエントロピーである)。

## 2.4 発話タイプパターンを使用しない条件での推定実験

### 2.4.1 推定条件

発話タイプの条件付き確率による予測を以下の各項目について調査した。

#### (1) 発話タイプを観測/予測する単位

- ・発話タイプ単位
- ・発声 (句点「。」で区切られる単位)  
 なお、発声に属する最後の発話タイプ単位の発話タイプを発声の発話タイプとする。
- ・Temporizer などを除いた発話タイプ単位  
 (次に示す発話タイプを持つ発話タイプ単位は発話対話構造に大きな影響を与えないと予想されたためである。ここで、「Temporizer など」とは次の発話タイプを指す。  
 <acknowledge> <alert> <apology> <expressive> <temporizer> <vocative>)
- ・Temporizer などを除いた発声

#### (2) 予測用条件

次のような条件の内いくつかを現単位の発話タイプ (ift と略記) の予測のために用いる。

role: 現単位の話者役割  
 p\_ift: 前単位の発話タイプ  
 p\_role: 前単位の話者役割  
 pp\_ift: 前々単位の発話タイプ  
 pp\_role: 前々単位の話者役割

なお、以下条件 B により A を予測する場合を (A|B) と略記する。

### 2.4.2 実験データ

発話タイプコーパス (199 対話) から無作為に抽出した 180 対話を学習データとし、残りの 19 対話をテストデータとした (付録のファイルリスト参照)。なお、対話開始は <start> という発話タイプを持つとした。

実験データ諸元

	学習データ	実験データ
発話タイプ単位	9978単位	1047単位
発声	6588発声	661発声
<temporizer> などを除いた発話タイプ単位	7331単位	745単位
<temporizer> などを除いた発声	5863単位	581単位

### 2.4.3 エントロピーとパープレキシティ

発話タイプ付与単位

Condition	entropy	perplexity
(ift)	3.78974	13.8301
(iftlrole)	3.69398	12.9419
(iftlp_ift)	3.11014	8.63468
(iftlrole p_ift p_role)	2.68633	6.43676
(iftlp_ift pp_ift)	3.10553	8.60713
(iftlrole p_role p_ift pp_role pp_ift)	2.6813	6.41436

発声 (「。」ごと)

Condition	entropy	perplexity
(ift)	3.69568	12.9572
(iftlrole)	3.583	11.9837
(iftlp_ift)	3.04949	8.27921
(iftlrole p_ift p_role)	2.58975	6.01996
(iftlp_ift pp_ift)	3.04836	8.2727
(iftlrole p_ift p_role pp_role pp_ift)	2.58823	6.01359

Temporizer などを除いた発話タイプ付与単位

Condition	entropy	perplexity
(ift)	3.38018	10.412
(iftlrole)	3.35104	9.71976
(iftlp_ift)	2.6785	6.40189
(iftlrole p_ift p_role)	2.305	4.94167
(iftlp_ift pp_ift)	2.67814	6.40032
(iftlrole p_ift p_role pp_role pp_ift)	2.30457	4.94021

Temporizer などを除いた発声

Condition	entropy	perplexity
(ift)	3.46302	8.6346
(iftlrole)	3.35104	10.2038
(iftlp_ift)	2.89027	7.41409
(iftlrole p_ift p_role)	2.4927	5.62829
(iftlp_ift pp_ift)	2.89003	7.41287
(iftlrole p_ift p_role pp_role pp_ift)	2.49241	5.62716

### 2.4.4 的中率

下表 Condition で与えられた (条件付き) 確率により推定を行った。(ift) と (iftlrole) では数の多い <inform> が常に推定された。(iftlrole p\_ift p\_role) および (iftlp\_ift) においてはバイグラムが学習データ中がない場合、デフォルトで <inform> を与えた。トライグラム (iftlp\_ift pp\_ift) による推定で、トライグラムが学習データ中がない場合、バイグラム (iftlp\_ift) による推定を行い、バイグラムもない場合はデフォルトで <inform> を与えた。同様に、トライグラム (iftlrole p\_ift p\_role pp\_role pp\_ift) による推定では、トライグ

ラムが学習データ中がない場合、バイグラム (iftlrole p\_ift p\_role) による推定を行い、バイグラムもない場合はデフォルトで <inform> を与えた。なお、以下カバー率 (cover rate) とは条件付き確率  $p(A|B)$  による予測において (A B) の共起が学習データ中に存在する率を指す。

発話タイプ付与単位

Condition	accuracy	cover rate
(ift)	29.04% (304/1047)	1.0
(iftlrole)	29.04% (304/1047)	1.0
(iftlp_ift)	35.34% (370/1047)	1.0
(iftlrole p_ift p_role)	40.78% (427/1047)	1.0
(iftlp_ift pp_ift)	35.34% (370/1047)	1.0
(iftlrole p_ift p_role pp_role pp_ift)	40.69% (426/1047)	1.0

発声 (「。」ごと)

Condition	accuracy	cover rate
(ift)	31.16% (206/661)	1.0
(iftlrole)	31.16% (206/661)	1.0
(iftlp_ift)	36.91% (244/661)	1.0
(iftlrole p_ift p_role)	40.39% (267/661)	1.0
(iftlp_ift pp_ift)	36.91% (244/661)	1.0
(iftlrole p_ift p_role pp_role pp_ift)	40.39% (267/661)	1.0

Temporizer などを除いた発話タイプ付与単位

Condition	accuracy	cover rate
(ift)	40.81% (304/745)	1.0
(iftlrole)	40.81% (304/745)	1.0
(iftlp_ift)	47.52% (354/745)	1.0
(iftlrole p_ift p_role)	49.80% (371/745)	1.0
(iftlp_ift pp_ift)	47.52% (354/745)	1.0
(iftlrole p_ift p_role pp_role pp_ift)	49.53% (369/745)	1.0

## Temporizer などを除いた発声

Condition	accuracy	cover rate
(ift)	35.46% (206/581)	1.0
(iftlrole)	35.46% (206/581)	1.0
(iftlp_ift)	39.07% (227/581)	1.0
(iftlrole p_ift p_role)	42.69% (248/581)	1.0
(iftlp_ift pp_ift)	39.07% (227/581)	1.0
(iftlrole p_ift p_role pp_role pp_ift)	42.86% (249/581)	1.0

### 2.4.5 まとめ

#### (1) 話者役割情報の効果

一般的に話者役割情報を用いると成績が向上する。ユニグラムの場合は、常に <inform> を予測するため、的中率で (ift) と (iftlrole) の差がない。

#### (2) 先行情報の利用

1つ前の単位の発話タイプを利用した場合（バイグラム）、利用しない場合（ユニグラム）に比べパープレキシティーが低くなる。しかし、2つ前までの単位の発話タイプを利用しても（トライグラム）、バイグラムの場合とそれほどパープレキシティーは変化しない。的中率もバイグラムはユニグラムに比べ高いが、トライグラムでは、バイグラムと同じかむしろ低くなる（低くなるのはデータのスパース度が高くなるためトライグラムの信頼性が低くなるためと考えられる）。

#### (3) 発話タイプを観測/予測する単位による差

概して、Temporizer などを除いた発話タイプ付与単位 > Temporizer などを除いた発声 > 発声 > 発話タイプ付与単位、の順で成績がよい。

## 2.5 発話タイプパターンを使用する条件での推定実験

次章（3章）で発話タイプパターンを予測する際に、前発話の発話タイプパターンによる現発話の発話タイプの予測を利用する。このため、前発話の発話タイプパターンから現発話の発話タイプを予測する実験を行った。発話タイプを観測/予測する単位は「発声」（句点「。」で区切られる単位）とする<sup>4</sup>（発声に属する最後の発話タイプ単位の発話タイプを発声の発話タイプとする）。

### 2.5.1 推定条件

次のような条件の内いくつかを現単位の発話タイプ（ift と略記）の予測のために用いる。

exp: 現単位の発話タイプパターン  
role: 現単位の話者役割

<sup>4</sup> 次の理由により発声を単位として選んだ。

- 1) 発話タイプ付与単位より発声のほうが発話タイプ予測の成績がよい。
- 2) Temporizer など除去するためには前もって発話タイプの推定ができなければならない。
- 3) 発声単位は高い確率で推定することができる [8]。



p\_exp: 前単位の発話タイプパターン

p\_role: 前単位の話者役割

(前々単位の情報の利用は予備実験においてよい成績が出なかったため行わなかった。)

## 2.5.2 実験データ

発話タイプコーパス (199対話) から無作為に抽出した180対話を学習データとし、残りの19対話をテストデータとした (付録のファイルリスト参照)。なお、対話開始は <start> という発話タイプおよび「START」という発話タイプパターンを持つとした。

## 2.5.3 エントロピーとパープレキシティ

Condition	entropy	perplexity
(iftlexp)	0.451151	1.36713
(iftlexp role)	0.404914	1.32401
(iftlexp p_exp)	0.103	1.074
(iftlexp role p_exp p_role)	0.079320	1.05652
(ift  role p_role p_exp)	1.89607	3.72198

(最後の行のみ現発話の発話タイプパターンを用いていない。)

## 2.5.4 的中率

下表 Condition で与えられた (条件付き) 確率により推定を行った。(iftlexp), (iftlexp role) および (ift|role p\_role p\_exp) による推定においては、これらの条件付き確率が学習データ中にない場合、デフォルトで <inform> を与えた。バイグラム (iftlexp p\_exp) による推定で、バイグラムが学習データ中にない場合、(iftlexp) による推定を行い、(iftlexp) もない場合はデフォルトで <inform> を与えた。同様に、バイグラム (iftlexp role p\_exp p\_role) による推定で、バイグラムが学習データ中にない場合、(iftlexp role) による推定を行い、(iftlexp role) もない場合はデフォルトで <inform> を与えた。

下表の cover rate は主たる条件付き確率  $p(A|B)$  による予測において (A B) の共起が学習データ中に存在する率 ( $p(A|B)$  を用いて A の生起確率が計算される率) を指し、cover rate 2 は、バイグラムがない場合にユニグラムによる推定ができた割合を指す。

最後の2行の網掛けは現発話の発話タイプパターンを使用していないことを表わす (現 ift の前発話タイプパターンによる推定は3章の「発話タイプパターンの推定」で利用される)。最後の行は、(ift|role p\_role p\_exp) を  $\sum_{p\_ift} p(\text{ift|role } p\_ift \text{ } p\_role) \times p(p\_ift|p\_role \text{ } p\_exp)$

として近似した場合の結果である ( $p\_ift$  は推定される前発話の発話タイプ)。上式が値を持たない場合 (バイグラム ( $p\_ift|p\_role \text{ } p\_exp$ ) がない場合)、デフォルトで <inform> を与えた。

Condition	accuracy	cover rate	cover rate 2
(iftlexp)	80.94% (535/661)	94.40% (624/661)	-
(iftlexp role)	78.97% (522/661)	91.83% (607/661)	-
(iftlexp p_exp)	82.00% (542/661)	52.65% (348/661)	41.75% (276/661)
(iftlexp role p_exp p_role)	80.64% (533/661)	45.84% (303/661)	45.99% (304/661)
(iftl role p_role p_exp)	39.03% (258/661)	91.22% (603/661)	-
(p_iftl p_role p_exp) (iftl role p_ift p_role)	37.82% (250/661)	91.98% (608/661)	-

### 2.5.5 まとめ

#### (1) 話者役割情報の効果

話者役割情報を用いるとパープレキシティは低下するが、的中率も下がってしまう。原因はパープレキシティの低下に比べ、データ（バイグラムなど）の信頼度の低下が大きいためと推測される。

#### (2) 先行情報の利用

1つ前の単位の発話タイプを利用した場合（バイグラム）、利用しない場合（ユニグラム）に比べパープレキシティが低くなり、的中率も上昇する。

参考：最も成績の良かった (iftlexp p\_exp) による予測を学習データに対して行うと (Closed 実験)、的中率は 96.30% (6344/6588) であった。

### 第3章 発話タイプパタンの予測

現発話の発話タイプパターン（主に文末の表層表現）を先行する発話の発話タイプパターンなどの情報を用いて推定する実験を行った。結果として、1つ前の発話の情報や話者役割情報を用いると、発話タイプパタンの出現予測の精度が高まる（3.3 参照）ことがわかった。

#### 3.1 発話タイプパターン

第2章でも述べたように、発話タイプを推定するために用いられる表層表現を発話タイプパターンと呼ぶ。日本語の場合、発話タイプを推定する手掛かりは主に文末に現れるが、<wh-question> を判定する場合には疑問詞が手掛かりになり、<greet> などの判定には発話全体が手掛かりになることがある。

発話タイプコーパス（2.1 参照）において、発話タイプ単位は以下の形式で記述されている。実験の学習データでは、ENTRY を発話タイプパターンとし、LABEL を発話タイプとした<sup>5</sup>。

```
N_FILE : TBS23003
N_HATUWA : 047000
N_TANI : 047001
SPEAKER : 通訳者
UNIT-UTTR : それから|、|あらかじめ|何|か|予約|を|証明|でき|る|もの|を|頂け|ます|か|。
COMM : /
AMBIGUITY : ("*|ま|す|か" "*|何|か|*|ま|す|か" "*|何|か|*|頂け|ます|か")
ENTRY : *|何|か|*|頂け|ます|か
LABEL : (<action-request>)
```

#### 3.2 発話タイプパターン推定

##### 3.2.1 推定条件

先行発声の発話タイプパターンや発話タイプから現発声の発話タイプパターン候補を条件付き確率を用いて求める。ここで発声とはコーパスにおいて「。」で区切られる単位である<sup>6</sup>。発声に属する最後の発話タイプ単位の発話タイプを発声の発話タイプとする。

次のような情報を発話タイプパタンの予測に用いる。

```
ift:      現発声の発話タイプ
role:     現発声の話者役割
p_ift:    前発声の発話タイプ
p_exp:    前発声の表層表現
p_role:   前発声の話者役割
```

なお、現発声の表層表現を exp と略記する。

考慮する先行発話の数は0（ユニグラム）および1（バイグラム）のみとした。トライグラム・補間の実験は予備調査で精度の向上を見なかつたので行なわなかつた（cf. 2章の発

<sup>5</sup> LABEL に多義性がある場合、例えば (<acknowledge> <yes>) の場合、0.5 回ずつ生起するとみなした。

<sup>6</sup> 2章では4つの単位について調査を行ったが、ここでは次の理由により発声を単位として選んだ。

- 1) 発話タイプ付与単位より発声のほうが発話タイプ予測の成績がよい。
- 2) Temporizer など除去するためには前もって発話タイプの推定ができなければならない。
- 3) 発声単位は高い確率で推定することができる [8]。

話タイプ予測結果)。

### 3.2.2 推定方法

現発声の発話タイプパターンを、さまざまな発話状況を条件とする条件付き確率を用いて推定する。現発声の発話タイプパターンを  $exp$ 、発話状況の条件を  $D$  とし、条件付き確率を  $p(exp|D)$  と書くことにする時、推定される  $exp$  は与えられた条件  $D$  に対し  $p(exp|D)$  を最大にするような  $exp$  である。条件付き確率はコーパスから得る。

#### i) 直接法

様々な条件についてコーパスから得られた条件付き確率を直接利用して  $exp$  を求める。

#### ii) 間接法

$exp$  を求める際に、前発声の発話タイプ  $p\_ift$  あるいは現発声の発話タイプ  $ift$  あるいはその両方を經由して条件付き確率を求める。具体的には、

##### 1) 前発声の発話タイプ $p\_ift$ を經由する方法

$p(exp|role\ p\_role\ p\_exp)$  を  $\sum_{p\_ift} p(exp|role\ p\_role\ p\_ift) \times p(p\_ift|p\_role\ p\_exp)$  で近似することにより  $exp$  を予測する。

##### 2) 現発声の発話タイプ $ift$ を經由する方法

$p(exp|role\ p\_role\ p\_exp)$  を  $\sum_{ift} p(exp|role\ ift) \times p(p\_ift|role\ p\_role\ p\_exp)$  で近似することにより  $exp$  を予測する。

##### 3) 前発声の発話タイプ $p\_ift$ と現発声の発話タイプ $ift$ を經由する方法

$p(exp|role\ p\_role\ p\_exp)$  を  $\sum_{p\_ift} \sum_{ift} p(exp|role\ ift) \times p(ift|role\ p\_role\ p\_ift) \times p(p\_ift|p\_role\ p\_exp)$  で近似することにより  $exp$  を予測する。

#### iii) 複合法

直接法のうち、実際に音声認識に用いることができ、十分なデータがあれば予測精度が高いと考えられるのはバイグラム ( $exp|role\ p\_role\ p\_exp$ ) による直接法である。しかし、十分なデータがない場合、このバイグラムはスパースになり信頼度が低くなると考えられる。そこで、このバイグラムの信頼度が高い場合は直接法を用い、そうでない場合は間接法を用いることにする。バイグラムの信頼度は条件部 ( $role\ p\_role\ p\_exp$ ) の学習データ中の生起回数により判断する。間接法としては、実験の結果成績が最も良かった 2) を用いた。

### 3.3 実験

発話タイプコーパス (199 対話) から無作為に抽出した 180 対話を学習データ、残りの 19 対話をテストデータとし (付録ファイルリスト参照)、学習データについてエントロピーとパープレキシティを求め (エントロピーとパープレキシティについては 2.3 節を参照)、学習データから得た条件付き確率を用いてテストデータにおける的中率を求めた。なお、対話開始は  $\langle start \rangle$  という発話タイプおよび「START」という発話タイプパターンを持つとした。

### 3.3.1 エントロピーとパープレキシティー

Condition	entropy	perplexity
(exp)	7.29966	157.55
(explrole)	6.78034	109.922
(explp_ift)	5.92667	60.8281
(explrole p_ift p_role)	4.8934	29.7207
(explift)	4.05513	16.6233
(explrole ift)	3.60225	12.1447
(explp_exp)	4.07518	16.8559
(explrole p_role p_exp)	3.22942	9.37889

網掛け部は発話タイプを条件部に持つ場合

### 3.3.2 的中率（直接法）

学習データからの条件付き確率によりで何らかの exp が求まらない場合、デフォルトとして  $p(\text{explrole})$  を最大にするような exp を与えた。カバー率 (cover rate) は学習データからの条件付き確率により何らかの exp が求まった割合である。

Condition	accuracy	cover rate
(exp)	5.60% (37/661)	100%
(explrole)	8.62% (57/661)	100%
(explp_ift)	12.10% (80/661)	100%
(explrole p_ift p_role)	16.49% (109/661)	100%
(explift)	28.29% (187/661)	100%
(explrole ift)	31.62% (209/661)	100%
(explp_exp)	12.41% (82/661)	94.55% (625/661)
(explrole p_role p_exp)	15.13% (100/661)	91.23% (603/661)

網掛け部は発話タイプを条件部に持つ場合

### 3.3.3 的中率（間接法）

間接法で何らかの exp が求まらない場合、デフォルトとして  $p(\text{explrole})$  を最大にするような exp を与えた。カバー率 (cover rate) は間接法で何らかの exp が求まった割合である。

Condition	accuracy	cover rate
(p_ift p_role p_exp)	14.52%	91.98%
(explrole p_ift p_role)	(96/661)	(608/661)
(ift p_role p_exp)	15.43%	91.23%
(explrole ift)	(102/661)	(603/661)
(p_ift p_role p_exp)	15.13%	91.98%
(ift role p_role p_ift)	(100/661)	(608/661)
(explrole ift)		

### 3.3.4 的中率 (複合法)

表の見方:

LIMIT: バイグラム (explrole p\_exp p\_role) の条件部 (role p\_exp p\_role) の生起する数を  $C$  とする時、 $C > \text{LIMIT}$  ならば (explrole p\_exp p\_role) による予測を行い、 $C \leq \text{LIMIT}$  ならば現発声の発話タイプの予測を経由する間接法

((iftlp\_role p\_exp) および(explrole ift) を使用) による予測を行った。

cover rate 1: バイグラム (explrole p\_exp p\_role) を用いた直接法で何らかの exp が求まる割合

cover rate 2: (explrole p\_exp p\_role) を用いた直接法で exp が求まらず、かつ現発声の発話タイプの予測を経由する間接法で何らかの exp が求まる割合

cover rate 3: 直接法でも間接法でも exp が求まらず、デフォルトとして  $p(\text{explrole})$  を最大にするような exp を与えた場合の割合

なお、LIMIT = 0 の場合の結果は直接法で (explrole p\_role p\_exp) を用いた場合の結果と同じである。

LIMIT	accuracy	cover rate 1	cover rate 2	cover rate 3
0	15.13% (100/661)	91.23% (603/661)	0% (0/661)	8.77% (58/661)
10	15.43% (102/661)	65.36% (432/661)	25.87% (171/661)	8.77% (58/661)
20	15.58% (103/661)	52.65% (348/661)	38.58% (255/661)	8.77% (58/661)
30	15.73% (104/661)	46.29% (306/661)	44.93% (297/661)	8.77% (58/661)
40	16.04% (106/661)	38.58% (255/661)	52.65% (348/661)	8.77% (58/661)
50	16.04% (106/661)	33.28% (220/661)	57.94% (383/661)	8.77% (58/661)
60	15.58% (103/661)	29.20% (193/661)	62.03% (410/661)	8.77% (58/661)

### 3.4 まとめ

発話タイプを与えられた情報として用いない推定のうちでは、複合法で LIMIT=40 または 50 とした場合の成績が最も良い。間接法では、現発話タイプの前発声と前発話者役割からの推定を経由して発話タイプパターンを推定する場合の成績が最も良い。発話タイプを与えられた情報として用いない直接法のうちでは (explrole p\_role p\_exp) を用いた場合の成績が最も良い。

## 第4章 内容語の予測

現発話に出現する内容語（自立語および接尾辞）を先行する発話の内容語などの情報を用いて推定する実験を行った。結果として、1つ前の発話中の自立語の情報を用いると、自立語の出現予測の精度が高まることがわかった。

・内容語の品詞リスト

<固有名詞> <サ変名詞> <普通名詞> <数詞> <代名詞> <人名> <住所名> <日時> <本動詞>  
<形容詞> <形容名詞> <副詞> <連体詞> <感動詞> <接尾辞> <名詞句>

### 4.1 内容語の推定

実験用のコーパスは SLDB の形態素解析データ (JMOR) [9] であるが、学習およびテスト用対話としては2章、3章で用いたものと同じものを用いた（付録ファイルリスト参照）。内容語を観測/予測する単位としては、句点「。」ごとの単位=発声を用いることにした。内容語は1つの発声中に複数生起することがあるが、それらのすべてを推定対象とした。

#### 4.2.1 推定条件

内容語を予測するための条件としては前発声中の内容語のほかに、現発声、前発声の発話者の役割を用いた。<sup>7</sup>

role: 現発声の話者役割

p\_cont: 前発声のいずれかの内容表現

p\_role: 前発声の話者役割

なお、現発声のいずれかの内容表現を cont と略記する。

#### 4.2.2 推定方法

Cont を発声  $S_i$  中の内容語、p\_cont を1つ前の発声中の内容語とすると、内容語のバイグラムは  $p(\text{cont}|\text{p\_cont}) = (\text{cont}$  が p\_cont の次の発声に起こる回数) ÷ (p\_cont が起こる回数) として定義できる。なお、対話開始は「START」という内容語からなるとした。

次の2つの方法による推定を行った。いずれの場合も、(cont|p\_cont) による推定の場合は、ユニグラム (cont) で最も頻度の高い内容語（「はい」<感動詞>）、(cont|role p\_cont) あるいは (cont|role p\_role p\_cont) による推定の場合は (cont|role) で最も頻度の高い内容語（やはり「はい」<感動詞>）をデフォルトで与える。

##### 1) 平均による予測

発声  $S_i$  に対し、その直前の発声  $S_{i-1}$  の内容語  $A_1, \dots, A_m$  を条件として  $S_i$  中のある内容語が X である確率  $p(X|A_1, \dots, A_m)$  を

$$p(X|A_1, \dots, A_m) = \frac{1}{m} \sum_{k=1}^m p(X|A_k)$$
 として計算する。

$S_i$  の話者役割情報 role,  $S_{i-1}$  の話者役割情報 p\_role を使用する場合は、それぞれ

$$p(X|A_1, \dots, A_m, \text{role}) = \frac{1}{m} \sum_{k=1}^m p(X|A_k, \text{role})$$

$$p(X|A_1, \dots, A_m, \text{role}, \text{p\_role}) = \frac{1}{m} \sum_{k=1}^m p(X|A_k, \text{role}, \text{p\_role})$$
 として計算する。

##### 2) 最大値による予測

発声  $S_i$  に対し、その直前の発声  $S_{i-1}$  の内容語  $A_1, \dots, A_m$  を条件として  $S_i$  中のある内容

<sup>7</sup> 1つの発声中に内容語が複数生起する場合、それらの内容語の予測条件は等しく、同一の予測がなされる。

語が  $x$  である確率  $P(X|A_1, \dots, A_m)$  を  $p(X|A_1, \dots, A_m) = \max_m p(X|A_m)$  として計算する。 $S_i$  の話者役割情報  $role$ ,  $S_{i-1}$  の話者役割情報  $p\_role$  を使用する場合は、それぞれ  $p(X|A_1, \dots, A_m, role) = \max_m p(X|A_m, role)$   
 $p(X|A_1, \dots, A_m, role, p\_role) = \max_m p(X|A_m, role, p\_role)$  として計算する。

## 4.2 実験

付録のファイルリスト中の学習データ180対話およびテストデータ19対話を用いて、予測実験を行った。使用する条件付き確率のエントロピー、パープレキシティも求めた（エントロピー、パープレキシティについては2.3を参照のこと）。

### 4.2.1 エントロピー、パープレキシティ

Condition	entropy	perplexity
(cont)	9.05738	532.774
(cont role)	8.81616	450.742
(cont p_cont)	6.24525	75.8593
(cont role p_cont)	5.66904	50.8804
(p_cont role p_role p_cont)	5.22978	37.5249

### 4.2.2 的中率 (ユニグラム)

Condition	accuracy	cover rate
(cont)	3.29% (94/2857)	100%
(cont role)	3.29% (94/2857)	100%

### 4.2.3 的中率 (平均による方法)

Condition	accuracy	cover rate
(cont p_cont)	5.01% (143/2857)	99.68% (2848/2857)
(cont role p_cont)	5.32% (152/2857)	99.51% (2843/2857)
(p_cont role p_role p_cont)	5.32% (152/2857)	99.37% (2839/2857)

### 4.2.4 的中率 (最大値による方法)

信頼度の低いデータを、学習データ中の条件部の生起回数がある値 (limit) より低い場合に足切りすることにより排除した。

#### 1) (cont|p\_cont) による予測

limit	accuracy	cover rate
0	4.31% (123/2857)	99.68% (2848/2857)
20	4.59% (131/2857)	95.34% (2724/2857)
40	4.66% (133/2857)	88.06% (2516/2857)
60	4.55% (130/2857)	84.53% (2415/2857)



## 2) (controlrole p\_cont) による予測

limit	accuracy	cover rate
0	4.48% (128/2857)	99.51% (2843/2857)
10	4.80% (137/2857)	96.32% (2572/2857)
20	4.69% (134/2857)	79.07% (2259/2857)
60	4.48% (128/2857)	72.66% (2076/2857)

## 3) (p\_controlrole p\_role p\_cont) による予測

limit	accuracy	cover rate
0	4.48% (128/2857)	99.37% (2839/2857)
20	4.80% (137/2857)	90.02% (2572/2857)
40	4.69% (134/2857)	78.75% (2250/2857)
60	4.48% (128/2857)	72.42% (2069/2857)

## 4.3 まとめ

現発声の話者役割情報や直前の内容語の情報をを用いると的中率が改善する。しかし、前発声の話者役割はあまり役に立たないようである。平均による方法と最大値による方法では、平均による方法の方が成績がよい。最大値による方法では、条件付き確率の条件部の生起回数に応じた足切りをすると成績が向上する。

## 文献

- [1] 荒川直哉, 竹澤寿幸, 加藤直人, 森元暎: “発話状況の情報を音声認識に用いた音声対話システム”, 第54回情報処理学会全国大会論文集 (1997).
- [2] 荒川直哉, 竹澤寿幸, 森元暎: “状況がわかれば聞き取れる”, ATR Journal #29 (1997).
- [3] J. Quantz: “Speech-Act Recognition for Dialogue Processing,” ATR-ITL Technical Report TR-IT-0258 (1998).
- [4] K. Jokinen, T. Morimoto: “Topic Information and Spoken Dialogue Systems,” *Proceedings of NLPRS'97* (1997).
- [5] 巖寺俊哲, 竹澤寿幸, 石崎雅人, 森元暎: “次発話予測による音声認識結果の再順序付け”, 第53回情報処理学会全国大会論文集 (1996).
- [6] 岩本 秀明, 妹尾 正身, 荒川 直哉, 森元 暎: “発話状況に基づく音声認識候補の再順序付け”, 言語処理学会第4回年次大会 (1998).
- [7] M. Nagata, T. Morimoto: “An Information-Theoretic Model of Discourse for Next Utterance Type Prediction,” *Transaction of Information Processing Society of Japan*, 35 no.6, pp.1050-1061 (1994).
- [8] 竹澤寿幸, 森元暎: “発話単位の分割または接合による言語処理単位への変換”, 情報処理学会研究報告, 音声言語情報処理 18-4 (1997).
- [9] 浦谷則好, 竹澤寿幸, 松尾秀彦, 森田千帆: “音声言語データベースの構成” ATR 音声翻訳通信研究所テクニカルレポート TR-IT-0056 (1994).

## 付録

### 付録1 発話タイプ（31種類）一覧

#### 1) 頻度順発話タイプ（発声）

発話タイプ	頻度	出現回数／全体の単位数
<inform>	0.307605	(2026.5/6588)
<acknowledge>	0.0865209	(570/6588)
<confirmation-question>	0.0828021	(545.5/6588)
<action-request>	0.0695203	(458/6588)
<wh-question>	0.0673953	(444/6588)
<accept>	0.0510777	(336.5/6588)
<thank>	0.0467517	(308/6588)
<information-request>	0.0426533	(281/6588)
<yn-question>	0.0416667	(274.5/6588)
<desire>	0.0300546	(198/6588)
<yes>	0.0281573	(185.5/6588)
<farewell>	0.0267152	(176/6588)
<permission-request>	0.016697	(110/6588)
<believe>	0.0162417	(107/6588)
<apology>	0.0159381	(105/6588)
<promise>	0.0142684	(94/6588)
<offer>	0.0113843	(75/6588)
<suggest>	0.00956284	(63/6588)
<greet>	0.00865209	(57/6588)
<instruct>	0.00470553	(31/6588)
<good-wishes>	0.00440194	(29/6588)
<offer-follow-up>	0.00402247	(26.5/6588)
<vocative>	0.00364299	(24/6588)
<thanks-response>	0.00303582	(20/6588)
<temporizer>	0.00227687	(15/6588)
<reject>	0.00212508	(14/6588)
<alert>	0.0016697	(11/6588)
<no>	0.00030358	(2/6588)
<do-you-understand-question>	0.00015179	(1/6588)
<offer-follow-up>	0	(0/6588)
<invite>	0	(0/6588)

注) 回数が小数点を持つのは、多義性を持つ場合分数の回数を与えているためである。

2) 頻度順発話タイプ (発話タイプ付与単位)

発話タイプ	頻度	出現回数/全体の単位数
<inform>	0.284977	(2843.5/9978)
<temporizer>	0.122369	(1221/9978)
<acknowledge>	0.0980156	(978/9978)
<confirmation-question>	0.056374	(562.5/9978)
<accept>	0.0514131	(513/9978)
<action-request>	0.0465023	(464/9978)
<wh-question>	0.0447986	(447/9978)
<yes>	0.0329725	(329/9978)
<thank>	0.031369	(313/9978)
<yn-question>	0.0286129	(285.5/9978)
<information-request>	0.0283624	(283/9978)
<desire>	0.0256564	(256/9978)
<greet>	0.0211465	(211/9978)
<apology>	0.0201443	(201/9978)
<farewell>	0.017739	(177/9978)
<believe>	0.0122269	(122/9978)
<permission-request>	0.0111245	(111/9978)
<vocative>	0.0105232	(105/9978)
<promise>	0.0102225	(102/9978)
<expressive>	0.00992183	(99/9978)
<offer>	0.00761676	(76/9978)
<suggest>	0.00641411	(64/9978)
<alert>	0.0039086	(39/9978)
<no>	0.00370816	(37/9978)
<instruct>	0.0034075	(34/9978)
<good-wishes>	0.00310684	(31/9978)
<offer-follow-up>	0.00275606	(27.5/9978)
<thanks-response>	0.00250551	(25/9978)
<reject>	0.00200441	(20/9978)
<do-you-understand-question>	0.00010022	(1/9978)
<invite>	0	(0/6588)

付録2 ファイルリスト (計199ファイル)

テストデータ (19ファイル)

TAC23031 TAS12010 TAS13024 TAS32008 TAS33014 TBS23001 TCC22011 TCC23023  
TCC23062 TCS12024 TCS13025 TCS33007 TDS32005 THS13002 TKS22001 TOS13003  
TRS12001

学習データ (180ファイル)

TAC22011 TAC22012 TAC22013 TAC22014 TAC23011 TAC23013 TAC23021 TAC23023  
TAC23033 TAC23034 TAS12001 TAS12002 TAS12005 TAS12007 TAS12008 TAS12009  
TAS12012 TAS12014 TAS12016 TAS12018 TAS12020 TAS12022 TAS12024 TAS13001  
TAS13002 TAS13006 TAS13008 TAS13009 TAS13010 TAS13011 TAS13012 TAS13014  
TAS13016 TAS13018 TAS13020 TAS13022 TAS22001 TAS22002 TAS22021 TAS22031  
TAS23001 TAS23002 TAS32001 TAS32002 TAS32003 TAS32005 TAS32007 TAS32009  
TAS32010 TAS32011 TAS32013 TAS32014 TAS33001 TAS33002 TAS33010 TAS33012  
TAS33015 TBS12001 TBS12002 TBS12004 TBS13001 TBS13003 TBS22001 TBS22003  
TBS23003 TBS32001 TBS32003 TBS33001 TBS33003 TCC22013 TCC22031 TCC22032  
TCC22034 TCC22042 TCC22044 TCC22074 TCC22091 TCC22093 TCC22094 TCC22112  
TCC23011 TCC23013 TCC23021 TCC23031 TCC23033 TCC23034 TCC23042 TCC23044  
TCC23052 TCC23054 TCC23064 TCC23072 TCC23074 TCC23082 TCC23084 TCC23092  
TCC23094 TCC23102 TCC23104 TCC23112 TCC23114 TCC23122 TCC23124 TCS12019  
TCS12021 TCS12022 TCS12023 TCS12026 TCS13017 TCS13018 TCS13019 TCS13021  
TCS13023 TCS32017 TCS32019 TCS32021 TCS33005 TDS12003 TDS12004 TDS13005  
TDS13006 TDS13008 TDS32007 TDS33007 TDS33009 TGS12001 TGS12002 TGS13001  
TGS23001 TGS32001 TGS33001 THS12001 THS12002 THS12003 THS13001 THS22001  
THS22003 THS23001 THS32001 THS32003 THS33001 TIS12001 TIS13001 TIS22001  
TIS23001 TIS32001 TIS33001 TKS12001 TKS13001 TKS23001 TKS32001 TKS33001  
TOS12001 TOS12003 TOS13001 TOS13005 TOS22001 TOS22003 TOS23001 TOS23003  
TOS23005 TOS32001 TOS32003 TOS33001 TOS33003 TOS33005 TRS12002 TRS13001  
TRS13002 TRS22001 TRS23001 TRS32001 TRS33001 TSS12001 TSS12002 TSS13001  
TSS22001 TSS23001 TSS32001 TSS33001

付録3 前後する発話タイプの相関

一部の発話タイプのバイグラムには高い相関が存在する。このようなもののうち、重要度の高いものを取りだすことを考えて、次のような指標（バイグラムの出現確率とバイグラムの相関の積）を用いることにする。

$$i(a,b) = p(a,b) \log_2 \frac{p(a,b)}{p(a)p(b)}$$

これは、 $I(A,B) = \sum_{a \in A} \sum_{b \in B} p(a,b) \log_2 \frac{p(a,b)}{p(a)p(b)}$  で定義される相互情報量の計算に際しての貢献の大きさを示す。

発声の発話タイプバイグラムについて  $i(a,b)$  が大きい上位 35 個を以下に示す。

a	b	$i(a,b)$	$p(a b)$	$(ab)/(b)$
<yes>	<confirmation-question>	0.0766263	0.279304	(152.5/546)
<accept>	<action-request>	0.0699392	0.358079	(164/458)
<acknowledge>	<inform>	0.0408134	0.156049	(316/2025)
<greet>	<START>	0.0331471	0.25	(45/180)
<thank>	<farewell>	0.027662	0.530612	(52/98)
<inform>	<wh-question>	0.0275912	0.527027	(234/444)
<inform>	<START>	0.0251135	0.733333	(132/180)
<inform>	<information-request>	0.0248139	0.601423	(169/281)
<confirmation-question>	<confirmation-question>	0.0243625	0.214286	(117/546)
<farewell>	<thank>	0.0236353	0.221739	(51/230)
<acknowledge>	<desire>	0.0210516	0.348485	(69/198)
<inform>	<inform>	0.0208453	0.351605	(712/2025)
<inform>	<yn-question>	0.019048	0.54827	(150.5/274.5)
<apology>	<action-request>	0.0185434	0.100437	(46/458)
<thank>	<thank>	0.0179473	0.226087	(52/230)
<accept>	<information-request>	0.0151619	0.188612	(53/281)
<wh-question>	<acknowledge>	0.0147832	0.149123	(85/570)
<acknowledge>	<believe>	0.0144613	0.401869	(43/107)
<confirmation-question>	<permission-request>	0.0140583	0.381818	(42/110)
<thanks-response>	<thank>	0.0128091	0.078261	(18/230)
<accept>	<permission-request>	0.0115944	0.281818	(31/110)
<action-request>	<acknowledge>	0.0115675	0.136842	(78/570)
<believe>	<wh-question>	0.010442	0.072072	(32/444)
<information-request>	<accept>	0.010131	0.126677	(42.5/335.5)
<inform>	<apology>	0.0096174	0.617647	(63/102)
<yn-question>	<acknowledge>	0.00951445	0.093860	(53.5/570)
<wh-question>	<accept>	0.00948345	0.154993	(52/335.5)
<farewell>	<farewell>	0.00890532	0.204082	(20/98)
<desire>	<wh-question>	0.00870851	0.085586	(38/444)
<yes>	<yn-question>	0.00839745	0.105647	(29/274.5)
<good-wishes>	<thank>	0.00805299	0.060870	(14/230)
<promise>	<accept>	0.00626228	0.059613	(20/335.5)
<inform>	<greet>	0.0056715	0.631579	(36/57)
<desire>	<inform>	0.00563918	0.040988	(83/2025)
<information-request>	<acknowledge>	0.00519813	0.074561	(42.5/570)

付録4 先行発声の発話タイプと発話タイプパターンとの相関

先行発声の発話タイプパターンを b、後続の発話タイプパターンを a とした場合の  $i(a,b)$  (付録3参照) が大きい順に a,b の組み合わせ上位 35 個を以下に示す。

a	b	$i(a,b)$	$p(a b)$	$(ab)/(b)$
* ご ざ い ま す	<START>	0.0694315	0.644444	(116/180)
* か し こ ま り ま し た	<action-request>	0.0401045	0.229258	(105/458)
は い そ う で す	<confirmation-question>	0.0314411	0.111824	(61/545.5)
* で す	<information-request>	0.0234238	0.252669	(71/281)
* あ り が と う ご ざ い ま し た	<farewell>	0.0227481	0.387755	(38/98)
* お 待 た せ い た し ま し た	<action-request>	0.0192622	0.0829694	(38/458)
* ご ざ い ま す	<greet>	0.0189775	0.578947	(33/57)
そ う で す か	<inform>	0.0188889	0.0508516	(103/2025.5)
も し も し	<START>	0.0179181	0.127778	(23/180)
* あ り が と う ご ざ い ま し た	<thank>	0.011384	0.134783	(31/230)
* か し こ ま り ま し た	<desire>	0.0113753	0.176768	(35/198)
そ う で す	<confirmation-question>	0.0108428	0.0384968	(21/545.5)
* そ の と お り で す	<confirmation-question>	0.0106976	0.0366636	(20/545.5)
* で す	<inform>	0.00973679	0.0750432	(152/2025.5)
* で す	<wh-question>	0.00955981	0.123874	(55/444)
* 分 か り ま し た	<action-request>	0.00942403	0.0960699	(44/458)
は い	<information-request>	0.0093595	0.0960854	(27/281)
* と い い い ま す	<information-request>	0.00890115	0.0498221	(14/281)
* お 願 い し ま す	<wh-question>	0.00849584	0.0878378	(39/444)
* ど う い た し ま し て	<thank>	0.00841119	0.0521739	(12/230)
* で す ね	<confirmation-question>	0.00753283	0.0843263	(46/545.5)
デ フ ォ ルト	<confirmation-question>	0.0070697	0.0458295	(25/545.5)
* と 申 し ま す	<information-request>	0.00666386	0.0569395	(16/281)
* 分 か り ま し た	<inform>	0.00655344	0.0488768	(99/2025.5)
* 思 い い ま す	<wh-question>	0.00632621	0.045045	(20/444)
こ ん に ち は	<START>	0.00531389	0.0388889	(7/180)
え え そ う で す	<confirmation-question>	0.00524694	0.0183318	(10/545.5)
* に な り ま す	<inform>	0.00500127	0.0241916	(49/2025.5)
* 失 礼 い た し ま す	<thank>	0.00479973	0.0434783	(10/230)
* い た し ま す	<accept>	0.00477043	0.0417288	(14/335.5)
* お 待 ち く だ さ い い ま せ	<acknowledge>	0.00468313	0.022807	(13/570)
* お 待 ち く だ さ い	<acknowledge>	0.00465354	0.0280702	(16/570)
* 失 礼 い た し ま す	<good-wishes>	0.00455947	0.3125	(5/16)
* と な っ て お り ま す	<apology>	0.00444315	0.0882353	(9/102)
* ご ざ い ま す ね	<desire>	0.00436973	0.0707071	(14/198)
* ま す	<inform>	0.0043603	0.0454209	(92/2025.5)
* に な り ま す	<wh-question>	0.00431244	0.0427928	(19/444)
は い	<action-request>	0.00428608	0.0480349	(22/458)
* ご ざ い ま す ね	<permission-request>	0.00426833	0.1	(11/110)
* ま す	<yn-question>	0.0041742	0.0837887	(23/274.5)

付録5 先行発声の発話タイプパターンと発話タイプパターンとの相関

先行発声の発話タイプパターンを b、後続の発話タイプパターンを a とした場合の i(a,b) (付録3参照) が大きい順に a,b の組み合わせ上位 40 個を以下に示す。

a	b	i(a,b)	p(a b)	(ab)/(b)
* ご ざ い ま す	START	0.0694315	0.644444	(116/180)
* か し こ ま り ま し た	* お 願 い し ま す	0.0205238	0.242991	(52/214)
も し も し	START	0.0179181	0.127778	(23/180)
* ご ざ い ま す	も し も し	0.0130967	0.833333	(20/24)
* お 待 た い せ い た し ま し た	* お 待 ち く だ さ い	0.0129108	0.326531	(16/49)
* で す	* で す	0.0117884	0.149051	(55/369)
* と い 申 し ま す	* ご ざ い ま す	0.0114976	0.0833333	(23/276)
は い そ う で す	* ご ざ い ま す ね	0.0107074	0.160714	(18/112)
* お 待 た い せ い た し ま し た	* お 待 ち く だ さ い ま せ	0.0103055	0.413793	(12/29)
は い そ う で す	* で す ね	0.0099409	0.0836502	(22/263)
そ う で す か	* ま す が	0.0088591	0.275862	(16/58)
* 思 い い ま す	* 何 時 * で し ょ う か	0.0081280	0.434783	(10/23)
* で す	* お 願 い い た し ま す	0.0081073	0.30137	(22/73)
* あ り が と う ご ざ い ま し た	* あ り が と う ご ざ い ま し た	0.00678623	0.155963	(17/109)
* あ り が と う ご ざ い ま し た	* お 待 ち い た し て お り ま す	0.0067738	0.555556	(10/18)
* で す ね	デ フ ォ ル ト	0.0066710	0.216867	(18/83)
* ど う い た し ま し て	* あ り が と う	0.0066508	0.162791	(7/43)
* 失 礼 い た し ま す	* あ り が と う ご ざ い ま し た	0.0064350	0.0917431	(10/109)
* あ り が と う ご ざ い ま し た	* お 願 い し ま す	0.0058808	0.289474	(11/38)
そ う で す か	* と い な っ て お り ま す	0.0057558	0.196721	(12/61)
* 分 か り ま し た	* く だ さ い	0.0055824	0.366667	(11/30)
* い ら っ し ゃ い ま せ	こ ん に ち は	0.0055256	0.5	(4/8)
* そ の と お り で す	* ご ざ い ま す ね	0.0054474	0.0714286	(8/112)
* お 待 ち く だ さ い ま せ	* か し こ ま り ま し た	0.0054143	0.0416667	(11/264)
* で す け れ ど も	* ご ざ い ま す	0.0053518	0.0289855	(8/276)
こ ん に ち は	START	0.0053139	0.0388889	(7/180)
* で 結 構 で す	* よ ろ し い で し ょ う か	0.0050913	0.189189	(7/37)
* で す	* 何 名 * で し ょ う か	0.0050245	0.555556	(10/18)
* い た し ま す	* か し こ ま り ま し た	0.0049010	0.0492424	(13/264)
* で す	* て い ま す	0.0048777	0.225352	(16/71)
* か し こ ま り ま し た	* お 願 い し た い い の で す が	0.0046556	0.571429	(8/14)
* ご ざ い ま す ね	* か し こ ま り ま し た	0.0044539	0.0606061	(16/264)
そ う で す	* ご ざ い ま す ね	0.0043571	0.0625	(7/112)
* か し こ ま り ま し た	* で す	0.0042775	0.0785908	(29/369)
デ フ ォ ル ト	デ フ ォ ル ト	0.0042425	0.108434	(9/83)
* か し こ ま り ま し た	* た い い ん で す が	0.0041752	0.333333	(9/27)
* お 待 ち く だ さ い	* で す ね	0.0041599	0.0418251	(11/263)
* 分 か り ま し た	* に な り ま す	0.0041531	0.14	(14/100)
* に な り ま す の で	* に な り ま す	0.0040708	0.05	(5/100)
* お 待 ち し て お り ま す	* あ り が と う ご ざ い ま し た	0.0039577	0.0642202	(7/109)

付録6 前後する発話の内容語の相関

1) 前発声中の内容語を b、後続発話中の内容語を a とした場合の  $i(a,b)$  (付録3参照) が大きい順に a, b の組み合わせ上位 40 個を以下に示す。(「START」は会話の始まりを示す。)

a	b	$i(a,b)$	$p(a b)$ (ab)/(b)
零+ゼロ+数詞	零+ゼロ+数詞	0.0035352	0.319355 (99/310)
かしこまりました+感動詞	願+ネガ+本動詞	0.00303894	0.220049 (90/409)
フロント+普通名詞	START	0.00299878	0.272222 (49/180)
九+キュウ+数詞	九+キュウ+数詞	0.00293185	0.383721 (66/172)
はい+ハイ+感動詞	願+ネガ+本動詞	0.00291925	0.315403 (129/409)
はい+ハイ+感動詞	START	0.0027757	0.538889 (97/180)
零+ゼロ+数詞	二+ニ+数詞	0.00269238	0.169675 (94/554)
三+サン+数詞	三+サン+数詞	0.00261163	0.205451 (98/477)
はい+ハイ+感動詞	三+サン+数詞	0.00260169	0.264151 (126/477)
二+ニ+数詞	二+ニ+数詞	0.00244257	0.18231 (101/554)
九+キュウ+数詞	零+ゼロ+数詞	0.00239972	0.206452 (64/310)
零+ゼロ+数詞	三+サン+数詞	0.00232075	0.169811 (81/477)
四+ヨン+数詞	零+ゼロ+数詞	0.00229711	0.241935 (75/310)
はい+ハイ+感動詞	二+ニ+数詞	0.00227111	0.220217 (122/554)
様+サマ+接尾辞	かしこまりました+感動詞	0.00216303	0.256506 (69/269)
三+サン+数詞	号+ゴウ+接尾辞	0.00211982	0.56 (56/100)
円+エン+接尾辞	円+エン+接尾辞	0.00209648	0.259259 (56/216)
三+サン+数詞	室+シツ+接尾辞	0.00209249	0.606742 (54/89)
八+ハチ+数詞	八+ハチ+数詞	0.00208581	0.305732 (48/157)
二+ニ+数詞	三+サン+数詞	0.00203395	0.178197 (85/477)
三+サン+数詞	二+ニ+数詞	0.00198787	0.15343 (85/554)
三+サン+数詞	零+ゼロ+数詞	0.0019783	0.229032 (71/310)
零+ゼロ+数詞	四+ヨン+数詞	0.00196254	0.173028 (68/393)
三+サン+数詞	一+イチ+数詞	0.00195847	0.195822 (75/383)
四+ヨン+数詞	二+ニ+数詞	0.00191714	0.140794 (78/554)
四+ヨン+数詞	四+ヨン+数詞	0.00190443	0.178117 (70/393)
四+ヨン+数詞	三+サン+数詞	0.00189903	0.155136 (74/477)
二+ニ+数詞	零+ゼロ+数詞	0.00189806	0.229032 (71/310)
様+サマ+接尾辞	はい+ハイ+感動詞	0.00189733	0.089318 (97/1086)
はい+ハイ+感動詞	零+ゼロ+数詞	0.00185529	0.280645 (87/310)
零+ゼロ+数詞	七+ナナ+数詞	0.00185393	0.26699 (55/206)
室+シツ+接尾辞	フロント+普通名詞	0.00183955	0.421053 (32/76)
九+キュウ+数詞	二+ニ+数詞	0.0018172	0.108303 (60/554)
九+キュウ+数詞	番号+バンゴウ+普通名詞	0.00181307	0.578125 (37/64)
はい+ハイ+感動詞	一+イチ+数詞	0.00180197	0.240209 (92/383)
三+サン+数詞	はい+ハイ+感動詞	0.00179706	0.093002 (101/1086)
号+ゴウ+接尾辞	フロント+普通名詞	0.00178234	0.421053 (32/76)
二+ニ+数詞	はい+ハイ+感動詞	0.00176692	0.095764 (104/1086)
三+サン+数詞	番号+バンゴウ+普通名詞	0.00176643	0.6875 (44/64)
零+ゼロ+数詞	室+シツ+接尾辞	0.00173442	0.483146 (43/89)



2) 前発声中の内容語を b、後続発話中の内容語を a とした場合の  $i(a,b)$  (付録3参照) が大きい順に数詞を含まない a,b の組み合わせ上位 40 個を以下に示す。(「START」は会話の始まりを示す。)

a	b	$i(a,b)$	$p(a b)$ (ab)/(b)
かしこまりました+感動詞	願+ネガ+本動詞	0.00303894	0.220049 (90/409)
フロント+普通名詞	START	0.00299878	0.272222 (49/180)
はい+ハイ+感動詞	願+ネガ+本動詞	0.00291925	0.315403 (129/409)
はい+ハイ+感動詞	START	0.0027757	0.538889 (97/180)
様+サマ+接尾辞	かしこまりました+感動詞	0.00216303	0.256506 (69/269)
円+エン+接尾辞	円+エン+接尾辞	0.00209648	0.259259 (56/216)
様+サマ+接尾辞	はい+ハイ+感動詞	0.00189733	0.089319 (97/1086)
室+シツ+接尾辞	フロント+普通名詞	0.00183955	0.421053 (32/76)
号+ゴウ+接尾辞	フロント+普通名詞	0.00178234	0.421053 (32/76)
な+ナ+本動詞	な+ナ+本動詞	0.00171422	0.155251 (68/438)
願+ネガ+本動詞	はい+ハイ+感動詞	0.00171416	0.077348 (84/1086)
はい+ハイ+感動詞	様+サマ+接尾辞	0.00170807	0.201245 (97/482)
部屋+ヘヤ+普通名詞	部屋+ヘヤ+普通名詞	0.00166373	0.174497 (52/298)
もしもし+感動詞	START	0.00162176	0.138889 (25/180)
そう+ソウ+副詞	様+サマ+接尾辞	0.00154958	0.143154 (69/482)
はい+ハイ+感動詞	教え+オシエ+本動詞	0.00153343	0.504587 (55/109)
待+マ+本動詞	待+マ+本動詞	0.00145976	0.175355 (37/211)
ドル+ドル+接尾辞	ドル+ドル+接尾辞	0.00142969	0.201031 (39/194)
人+リ+接尾辞	何名+ナンメイ+普通名詞	0.00142362	0.642857 (27/42)
予約+ヨyak+サ変名詞	はい+ハイ+感動詞	0.00138642	0.055249 (60/1086)
はい+ハイ+感動詞	予約+ヨyak+サ変名詞	0.00132774	0.258964 (65/251)
はい+ハイ+感動詞	はい+ハイ+感動詞	0.00130916	0.110497 (120/1086)
待+マ+本動詞	はい+ハイ+感動詞	0.00129729	0.047882 (52/1086)
室+シツ+接尾辞	はい+ハイ+感動詞	0.00129387	0.038674 (42/1086)
分か+ワカ+本動詞	な+ナ+本動詞	0.0012754	0.118721 (52/438)
万+マン+接尾辞	円+エン+接尾辞	0.00127454	0.162037 (35/216)
号+ゴウ+接尾辞	はい+ハイ+感動詞	0.00125911	0.039595 (43/1086)
バス+バス+普通名詞	バス+バス+普通名詞	0.00122659	0.240741 (26/108)
待+マ+本動詞	少々+ショウショウ+副詞	0.00121894	0.403226 (25/62)
待+マ+本動詞	かしこまりました+感動詞	0.00121701	0.126394 (34/269)
はい+ハイ+感動詞	部屋+ヘヤ+普通名詞	0.00120307	0.218121 (65/298)
円+エン+接尾辞	千+セン+接尾辞	0.00119923	0.260163 (32/123)
人+リ+接尾辞	様+サマ+接尾辞	0.00119401	0.082988 (40/482)
鈴木+スズキ+人名	はい+ハイ+感動詞	0.00117016	0.044199 (48/1086)
はい+ハイ+感動詞	よろし+ヨロシ+形容詞	0.00114987	0.389831 (46/118)
はい+ハイ+感動詞	電話番号+普通名詞	0.00114987	0.389831 (46/118)
円+エン+接尾辞	万+マン+接尾辞	0.00112119	0.20915 (32/153)
ドル+ドル+接尾辞	料金+リョウキン+普通名詞	0.00111144	0.346667 (26/75)
フィリップス+人名	名前+ナマエ+普通名詞	0.0010895	0.176471 (24/136)
ドル+ドル+接尾辞	いくら+イクラ+代名詞	0.00108297	0.511111 (23/45)