

# 対面会話における 相槌の機能の決定要因について

-Gradient Boosted Decision Tree による分析-

経済学部 3 年

松村優哉

（指導教員：鈴木亮子）

## 目次

1	序論	1
1.1	日本語における相槌	1
1.2	相槌の研究史	1
1.3	本稿の目的と構成	2
2	相槌の機能に関する先行研究	3
2.1	言語学の観点から	3
2.1.1	日本語教育の立場	3
2.1.2	会話分析の立場	5
2.1.3	対照言語学の立場	6
2.2	自然言語処理の観点から	7
2.3	先行研究に関する問題点と本稿の新規性	8
3	分析	9
3.1	データと相槌の抽出	9
3.1.1	データの概要	9
3.1.2	本稿における相槌の定義と抽出	10
3.1.3	数理的処理を行うためのラベリング	13
3.2	相槌の機能の再考	16
3.3	基礎分析	20
3.4	応用分析	20
3.4.1	分析手法について	20
3.4.2	分析結果	22
4	分析結果の考察	24
4.1	新情報と旧情報	24
4.2	相槌の相槌	24
4.3	格助詞	24
4.4	助詞	26
4.5	動詞	28
4.6	考察のまとめ	28
5	結論と展望	29
5.1	研究のまとめと結論	29
5.2	今後の課題と展望	29

付録A	トランスクリプトの記号一覧	30
付録B	Gradient Boosted Decision Tree アルゴリズム	31
B.1	決定木(Decision Tree)	31
B.1.1	木の成長	31
B.1.2	木の剪定	32
B.1.3	最適な木のサイズの決定	33
B.2	Gradient Boosted Decision Tree	33
B.3	Xgboost パッケージで用いたパラメーターの値	35

## 1 序論

本章では、本稿で主題とした相槌について、それを研究する重要性と追及する問いについて述べる。まず1.1で、日本語における相槌の位置づけを説明する。次に1.2で相槌という言語現象がどのように研究されてきたかを述べる。そして1.3で、本稿の目的と構成を明確にする。

### 1.1 日本語における相槌

本稿で扱う「相槌<sup>\*1</sup>」という言語現象は、研究者が関心を寄せる前からその名称が与えられていた。例えば、『日本国語大辞典』の「あいづちを打つ」の項目には以下のように記述されている。

#### 【あいづちを打(う)つ】

他人の話に調子を合わせる。

- \* 古活字本毛詩抄〔17C前〕一二「女興の讒言する時にあいづちを打れたかりしか」
- \* 俳諧・毛吹草追加〔1647〕中「木玉さへあい槌(ツチ)をうつ砧(きぬた)かな〈長治〉」
- \* 滑稽本・和合人〔1823~44〕四・下「自分が間拔をして買った事をぬりかくさんと、頻に茶見蔵に相槌を打たせんと気をもむ故」
- \* 芋粥〔1916〕〈芥川龍之介〉「両方に満足を与へるやうな、相槌(アヒツチ)を打った」

これらの例から分かるように、17世紀には既に「あいづちをうつ」という動詞と目的語の組み合わせが用いられていた。つまり、日本人は日本語の歴史の中で「あいづちを打つ」ことを知っており、それを会話や文章の中で使用してきたことになる。しかし、最初に言語学の研究対象として注目された英語の類似表現には元々そのような名称はなく、以降様々な研究者がこの言語現象を命名をしていくことになる。この詳細は1.2で述べる。また、現代においても相槌が日本語で多く使われていることは我々の直感からも分かる。

加えて、日本語の会話において相槌が多く用いられ他言語よりもその使用される頻度が多いことはすでに多くの研究で示されている(White (1986), Miller (1991), Clancy et al. (1996))。メイナード(1993)の調査によると日本人はアメリカ人の2.9倍もの相槌を打つという。そのため、日本語教育においても相槌の重要性は早くから認識されてきた(岡崎(1987), 畠(1988), 水谷(1988), 松田(1988))。一方、日本語教育の現場で相槌の指導が必ずしも成果を上げておらず、学習者の中には相槌を打つことに心理的な抵抗があるという調査結果(任・李(1995), 楊(2002))もある。この結果を見るに、相槌が日本語の会話において非常に重要な役割を果たしていると考えられ、その先行研究も多くある。次節では、相槌という言語現象がどのように研究されてきたかについて述べる。ただし、相槌の研究の中でも特にその機能について研究されたものは2章で紹介をする。

### 1.2 相槌の研究史

Clancy et al. (1996)によると、後に backchannel(バックチャンネル) と命名される一連の言語現象につ

---

\*1 「あいづち」には学術上「あいづち」「相槌」、「相づち」の表記があり研究者によってその表記は異なる。本稿では「相槌」と表記する。引用などでそのほかの表記が現れることがあるが、それらについても区別をせず同じ言語現象を指すこととする。

いて初めて言及したのはFries (1952)であろうという。Fries は、電話の会話の中に現れた yes, unh huhh, yeah, I see, good, oh, oh dear, fine, yes, I know, that's right などのワードを一つのグループとみなし、“those single free utterances”(自由に発せられるこれらの短い発話)は、対面会話で見られる頷きなどの表現と同様に“continued attention”(続けて注意を払っていく)表現であるとした。

このような言語表現が新しい研究対象として指摘された後、まなざしや頷きなど非言語的な表現との関連性が言及されるようになる。Dittmann and Llewellyn (1968)は mm-hmm, I see などを“listener responses”(聞き手の応答)と命名し、頷きの約70%とこの聞き手の応答が共起することを明らかにした。また、Kendon (1967)は“accompaniment signals”(伴奏的シグナル)と命名し、これらには mhm, yes など「注意を向けていることを示す」とI see, yes などの「賛同を示す」ものがあるとした。

現在、上記のような言語表現については backchannel という語が最もよく使用されるが、backchannel という単語を使用したのはYngve (1970)である。Yngve は主要な情報伝達以外の会話全般のことを“backchannel communication”と呼んだ。ここでいう backchannel は言語現象そのものではなく情報伝達の経路のことをさしており、Yngve は音声を発している者はすべて話者としてとらえていたため、相槌を打つのは全て発話者であるという点で、後の研究者とは異なっている(大浜 (2006))。しかしその後の研究では、backchannel という語を Yngve の言うところの backchannel という経路を通じて伝わる言語表現そのものとして使われるようになる(Duncan (1974), White (1986))。

この頃の backchannel の研究から刺激を受けて、日本でも相槌の研究がされていくようになる。1.1で述べてように日本語では古くから相槌という表現が使われていたため名称についての議論はなく、backchannel の研究を応用する形で1.1で述べたような日本語に相槌が多いことや日本語教育に関連した研究などが進められていったと考えられる。

### 1.3 本稿の目的と構成

本稿における最大の目的は、上記の研究や2章の先行研究を受けて相槌の機能が対話の中でどのように決まるのかを解明することにある。2章で紹介するように相槌の定義および相槌が持つ機能は研究者によって違いがある。また、分析に用いたデータも古い場合が多い。つまり相槌の機能の決定要因を探るためには、まず先行研究で提示されてきた相槌の機能が現在の言葉の使われ方の中でも有効なものなのか確認する必要がある。そこで、本稿の目的を以下の二つに設定する。

1. 対面会話における相槌の機能を再度明確にする。
2. 相槌の機能が会話のどのような要素によって決定づけられるのかを分析する。

本稿の流れとしては、まず2章で相槌の機能に関する先行研究を4つの立場に分けて紹介する。3章では実際に汎用的な会話のデータを用いて相槌を抽出し、その機能の検証をした後、それらの機能が会話の中のどのような要素で決定されるのかを機械学習の手法である Gradient Boosted Decision Tree を用いて分析する。4章では、その結果を受けてその要因と相槌の機能が決まる仕組みについて考察を行う。最後に5章で本研究の結論と、今後の展望として言語学以外の分野への応用の可能性について述べる。

## 2 相槌の機能に関する先行研究

これまで、日本語の話し言葉における相槌の機能に関して大きく分けて二つの分野で研究がなされてきた。

一つは言語学、その中でも特に日本語教育、会話分析、そして対照言語学の立場である。日本語教育で会話に注目する場合、話し手と聞き手の観点から非日本語話者と日本語話者の会話を比べることで日本語の特徴的な部分を抽出することを目的としている(堀口(1997))。会話分析では談話の一種である会話における相槌の頻度や機能について考察し、「言語とは何か」という言語学の究極的な問いへの一助となることを目的としている(メイナード(1993))。対照言語学では日本語と他言語との比較に重点を置いている。このような意味で、日本語教育と会話分析にも対照言語学の要素が含まれるが、分野の学問的目的が異なるため三つに分けて紹介をする。

もう一つの分野は、自然言語処理の立場である。この分野における相槌の研究は人工知能の分野での応用が期待されている。言語学が言語現象を記述し分析することを目的としているのに対し、自然言語処理では統計モデルや機械学習の手法を用いて数理的な分析を行う。

本章では相槌の先行研究のうち、その機能について分析しているものを上述の2つの分野から紹介する。そのうえで、先行研究の問題点を指摘することで先行研究からの連続性と本稿の新規性を明確にする。

### 2.1 言語学の観点から

言語学的に相槌を研究したものはさらに三つの分野に大別される。日本語教育、会話分析、対照言語学の立場である。

#### 2.1.1 日本語教育の立場

日本語教育の立場から相槌を分析した研究として、堀口(1997)がある。堀口は、非日本語学習者のための「日本語会話」という授業の中で得られた会話データを用いて、日本語の話し言葉における相槌の特徴を指摘した。

堀口(1997: 40-42)はまず、相槌とは何であるかということについて次のように述べている。

次のような点があいづちの特徴として抽出される。

- ① あいづちが出現する位置は、話し手の発話権\*2の中である
- ② あいづちを行使するのは、聞き手である
- ③ あいづちの機能は、聞いていることを伝える、わかったということを伝える、話の進行を助けるなどである
- ④ 相槌を表す言語形式は短い

①は、話し手は発話を続けているのだから、あいづちを行使するのは話し手ではないもの、つまり聞き手であるということである。その意味で②は同じことを指しているが、堀口は「会話においてことばを口から

---

\*2 【発話】( 発 utterance の訳語) 言語学で、口を開いてものを言う行為、または、それによって出てきた一区切りのことばをいう(『日本国語大辞典』)。「発話権」とは、会話におけるその権利のことを指す。

発する者が原則的には話し手であるにもかかわらず、あいづちのためのことばを口から発するのは聞き手の立場のままであるということは、相槌を特徴付けるものである。」として堀口は②をあいづちの特徴として明示した。また、③は聞き手の内部的認知行動であるがそのことを外に示すことで結果的に話の進行を助けるものであるとしている。④は、あいづちの必要条件でも十分条件でもなく形式上の特徴である。

そして、以上の記述をまとめて「あいづちは、話し手が発話権を行使している間に、聞き手が話し手から送られた情報を共有したことを伝える表現である」(堀口(1997: 42))としている。

相槌を上記のように特徴づけたうえで、堀口(1997: 52-60)では③のあいづちの機能である「聞いていることを伝える、わかったということを伝える、話の進行を助けるなど」を以下の5つに分解している。

1. 聞いているという信号
2. 理解しているという信号
3. 同意の信号
4. 否定の信号
5. 感情の表出

堀口は、これらの機能について1987年に放送された『徹子の部屋』というテレビ番組での会話を用いて説明をしている。なお、2章における用例表記は原文に従うものとするが、会話内で注目すべき語は太字にすることとする。

(1) 聞いているという信号

- 1 A: タクシーの運転手さんが一番あれですよねえ  
2 B: はあ  
3 A: 車に乗るとね  
4 B: え

(堀口 1997: 53より引用))

(1)では、BはAの1行目の「あれ」の内容がよくわからないので、2行目で「はあ」と曖昧な相槌で「よくわからないけど、聞いてますよ」という態度を示し、3行目の「車に乗るとね」を聞いてもまだ内容を理解できないため4行目で「え」と軽い相槌を打つことで「聞いてますよ」という態度を示している(堀口(1997: 53))。

(2) 理解しているという信号

- 1 A: で、あの方たちは、ほら、ああいう映画のファンだから  
2 B: ええ  
3 A: もうトラさんのファンで  
4 B: ええ  
5 A: もう盆と正月は楽しみにしていますって  
6 B: ええ  
7 A: よく言われます

(『徹子の部屋』(堀口 1997: 54より引用))

(2)では、Aの1行目の「あの方たち」や「ああいう映画」が「トラさんのファン」および「トラさんの映画」であることをBは知っているので、2行目や4行目で「ええ、分かります」という相槌を打っている。そのことからBはAの言う「(トラさんが封切られる)盆と正月を楽しみにしている」ことの意味も理解できるので、その態度を6行目で「ええ」という相槌で示している(堀口(1997: 54))。

(3) 同意の信号

1 A: 必ず夢から トラさんの夢から入っていくでしょう?

2 B: そうそうそう

(『徹子の部屋』(堀口 1997: 56より引用))

(3)2行目の「そうそうそう」は、1行目のAの発言に聞き手Bが「その通りだ」と同意していることを示している(堀口(1997: 56))。

(4) 否定の信号

(AがBに以前彫ってもらったはんこについて)

1 A: じゃ、わたしのはわりと、わりとお始めになったころ。

2 B: 初期だから恥ずかしくて、今ね。もうこれから使わないで。

3 A: いえ

4 B: また作って、今度持って

5 A: いえ

(『徹子の部屋』(堀口 1997: 57より引用))

(4)では、AがBに以前彫ってもらったはんこについて1行目で言及したところ、Bは2行目で「そのはんこは自分の初期のころの物だから使わないでほしい」と言い、それに対しAは2行目で「いえ」という相槌で「賛成ではないが話を聞いている」という態度を示している。4, 5行目でも同様にBの「Aのためにまたはんこを作る」と言ったことに対してAが「いえ」と否定の信号を送っている(堀口(1997: 57))。

(5)ではAの発言に対してBが「まあ」と驚きを示していることが分かる(堀口(1997: 60))。

(5) 感情の表出

1 A: アトムと同級生にウランちゃんという女の子がいるというんで、

2 B: まあ、ウランちゃん。

(『徹子の部屋』(堀口 1997: 60より引用))

### 2.1.2 会話分析の立場

日本語教育とは別の立場として会話分析がある。これは2.1.3で述べる対照言語学と似たものである。この会話分析の立場から相槌を研究したものとしてメイナード(1993)がある。メイナード(1993: 58)は日本語の相槌を次のように定義した。

「あいづち」とは話し手が発話権を行使している間に聞き手が贈る表現(非言語行動を含む)で、短い表現のうち話し手が順番を譲ったとみなされる反応を示したものは、あいづちとしない。

つまり、相槌を聞き手の反応としてとらえていることになる。これをさらに詳しく説明するため、メイナード(1993: 58)は次のようにも述べている。

話し手が発話順番を終えてすぐに聞き手が送る反応は、話し手の行動の勢力範囲内にあると考えられることから、分析の対象とする。一般に「間発話順番状況<sup>\*3</sup>」に見られるあいづちのような短い表現は、会話参加者の一種の「いいよども」の中の「埋め込み表現」(filler)と考え、分析の対象としない。

以上のような相槌の定義を受けて、メイナード(1993: 160)は日本語の相槌の機能として6つを挙げた。

1. 「続けて」というシグナル(continuer)
2. 内容理解を示す表現
3. 話し手の判断を支持する表現
4. 相手の意見、考え方に賛成の意思表示をする表現
5. 感情を強く出す表現
6. 情報の追加、訂正、要求などをする表現

このうち、1～5は2.1.1で述べた堀口(1997: 52-60)の1～3, 5とほぼ同じである。これに加えて、メイナード(1993: 160)は「更に話し手がそのまま発話順番を守っている間に聞き手が簡単な情報追加や訂正などをすることもある」として、6番目の機能を加えている。

また、メイナード(1993: 81-83)では、データ収集について会話は1985年5月に行われたことなどが述べられている。

### 2.1.3 対照言語学の立場

堀口(1997)やメイナード(1993)が触れなかった相槌の機能に言及しているものとして大浜・西村(2005)や大浜(2006)がある。これら二つの研究では、英語と比較した日本語の相槌の機能として特に「ターン交替」が言及されている。大浜・西村(2005)における定義を以下に引用する。大浜・西村(2005: 79)は、まず語が5つの形式のどれかにあてはまるかによって判断し、さらに4つの条件のうちいずれかにあてはまればそれを相槌と認定するとした。

#### I 形式

- (a) 「うん、あー、ええ」などのいわゆる相づち詞
- (b) それに類する「なるほど」、「そうですね」、「本当」等の語
- (c) 相手の発話の全部あるいは一部分の「繰り返し」
- (d) 「言い換え」
- (e) 相手の発話を途中から引き受ける「文完成」

#### II 条件

- (f) 話し手が発話権を行使している間に聞き手が送る表現
- (g) 相づちに対する相づちと言えるもの

<sup>\*3</sup> 話し手が発話順番を終了し、聞き手が話し手として発話順番を開始するまでの間の状況(メイナード(1993: 57))。

- (h) 少し間をおいて取られたターンの開始時に現れる表現
- (i) 質問に応じる形で現れた「はい、ええ、うん」など

本稿では基本的にはこの定義を3章の分析で用いている。

大浜(2006: 41)では相槌について、「英語圏の研究者は「ターンに関する機能」、特に continuer としての機能を見る傾向にあり、日本の研究者は「理解」や「反応」を示すなど「聞き手の行動に関する機能」を指摘する傾向」にあると述べている。そこで大浜は、相槌の機能の重要性が日本語と英語では異なるのではないかと考え、日本語におけるターン交替に着目をした。大浜(2006: 42-52)は、日本語のターン交替には7つのタイプ\*4があることを指摘している。

大浜(2006: 60, 189-205)は1999年5月から11月にかけて収集されたインタビュー形式の会話資料を用いて、日本語におけるターン交替と相槌との関係について分析をした。その結果、会話参加者一人一人の相槌の頻度は各会話参加者が会話中に使用するターン交替形式の種類とその数によって決定されることを明らかにした。

## 2.2 自然言語処理の観点から

自然言語処理は文書などの書き言葉を対象とした研究と、対話など音のある言葉を対象にする研究とがあり前者のほうが研究の量としては圧倒的に多いが、相槌の研究に関しては後者がほとんどである。なぜならば、相槌とはそもそも話し言葉で使われるものであり、書き言葉の中で使うことはないからである。そして、この分野における相槌の研究は主に機械による対話システムなどが応用先として想定されている。

そのような自然言語処理の観点から相槌の機能を分析したものとして榎本・石本(2009)がある。榎本・石本は「うん」という語に注目し、その発話機能である「応答」・「承認」・「相槌」を音響情報と統計モデルにより弁別できることを示した。具体的には、8つの会話データに現れた「うん」239トークンに対し、一般化線形モデルによるあてはめを行った。

ここで注意すべきなのは、「相槌」の機能判別をしたのではなく「うん」という語が持つ機能の判別をしたという点である。その点で本稿で行う分析とは根本的に異なるが、自然言語処理の立場から相槌の「うん」に対して研究が行われていることは注目すべきである。

榎本・石本(2009)は、「うん」の機能を定義する際に「ターン構成単位(turn constructional unit; TCU)」という概念を使用している。TCUとは、ある話者が発話を始めた時に、会話参加者全員が発話の終点の内容を理解できる最小の単位である。また、その終点のことを TRP(transition relevant place) と呼ぶ。榎本・石本は、この TCU を分析の前提として「うん」の発話機能を二通りに分け、それを音響情報で分類することを目指した。以下に「うん」の二つの発話機能を引用する。

1. 応答・承認の「うん」：発話権を有したものが発することができる発話であり、相手の TCU 終了後に発される発話である。特に、「うん」は同意や承認を表明するものであり、TRP において直ちに発せねばならない。特に、応答の開始の遅れや不在は相手への不同意の表明となるため、ここで発話することは聞き手にとって必須である。
2. 相槌の「うん」：発話権を有さないものが発することができる発話であり、相手の発話権を侵害

\*4 取得放棄、取得再放棄、他者選択、自己選択、再保持、割り込み、最終自己選択。詳細は大浜(2006)を参照。

しない発話とされる。従って、他者の TCU の途中に挿入することができる。ただし、相槌の挿入は聞き手の任意であり、同様の箇所です必ずしも生じるとは限らない。

この二つの「うん」の機能を弁別するため、榎本・西本は本稿でも3章で用いる『千葉大学3人話し言葉コーパス』と、『日本話し言葉コーパス』から4会話をそれぞれ5分間ずつ用いて分析を行った。具体的には、一般化線形モデルを用いて予測を行い減少法による stepAIC を行ったところ、「うん」の基本周波数の変化幅、平均、最大値と音声の強さの平均の4つの変数を用いることで、「相槌」の「うん」を88.8%弁別できたという。

## 2.3 先行研究に関する問題点と本稿の新規性

2.1と2.2では相槌の機能に特化した研究について言語学と自然言語処理という二つの観点から概観してきた。

言語学の観点では、前後の文脈やターン交替などの個々の機能に関する考察によって相槌の機能について議論をしている。しかし、データは30年前のものであったり新しくても10年以上前のデータである。現代の言葉の使い方が10年前と同じとは限らないため、この点に関しては最近のデータを用いて再検証する必要がある。そのうえ、データは研究者自身が集めたものであるため、研究間で統一性を持たせることが難しい。

自然言語処理の観点では、この点は解消されている。『千葉大学3人話し言葉コーパス』は2014年にリリースされたものであるため新しいデータであり、3.1で述べるように統一性や再現性も期待できる。また、分析についても一般化線形モデルという統計モデルを用いており、数理的なアプローチを用いているといえる。しかし、分析の前段階で発話機能を決定する際、TCU やTRP といったもので機械的に分類をしており、文脈等への考慮が足りないため、言語の実用的側面から外れている解釈も否めない。

本稿の新規性は、これら二つの観点を補完し、言語学的かつ統一的・数理的方法を用いて相槌の機能を分析している点である。具体的には、相槌にみられる9つの機能<sup>\*5</sup>が決まる要因を Gradient Boosted Decision Tree を用いて分析をする。また、相槌の機能についてこの分析手法が使われた研究は存在しないため、目的・手段ともに新規性があるといえる。

次章では、実際にデータから相槌を抽出してそこにみられる機能について議論をした後、Gradient Boosted Decision Tree を用いてその機能がどのような要因で決まるのかについて実際に分析を行う。

---

\*5 9つの機能については、3.1.2で詳しく議論をする。

### 3 分析

本章では、1.3で述べた二つの目的を達成するため、実際の会話データを用いて分析を行う。まず3.1で、分析に用いるデータと相槌の抽出方法を説明したうえで、本稿の立場を明確にする。本稿における相槌をどのように定義するのかもここで記述する。次に3.2で会話データと先行研究で言及された機能を照らし合わせ、データから考えられる新たな相槌の機能を提示する。3.3では抽出されたデータを参照し、対面会話でどのような機能の相槌が多く使われるのかを分析する。最後に3.4で、相槌の機能がどのような要因で決まるのかを分析する。

#### 3.1 データと相槌の抽出

##### 3.1.1 データの概要

本稿では、データとして『千葉大学3人会話コーパス』(以下、3人会話コーパス)を用いる。このコーパスには、大学生・大学院生・ポストドクターを含む同性3人からなる友人同士12組の雑談が収められている。2014年6月10日公開の Release1 では、12組の会話のうち、1会話各約10分ずつを対象として以下の情報が公開されている。

- 音声データ
- 転記テキスト(トランスクリプト)
- 形態論情報

なお映像データは公開されていないため、本稿では、映像によってはじめて取得可能な非言語情報(頷きなど)は分析の対象外とする。

また、2.3でデータの質について指摘をしたが、本稿で用いる3人会話コーパスではその使用説明書に次のような記載がある。

それ以前の対話コーパスが、人工的な状況設定(与えられた課題を解くなど)の元でのやり取りを防音室などの非日常的な空間の中で収録したものや、逆に、日常の会話場面をボイスレコーダなどで収録した低品質なものが中心であったのに対して、本コーパスでは、なるべく日常場面に近い自然なインタラクションを高品質なデータとして収録することを狙いとした。

つまり、このコーパスは言語研究で重要となる「日常場面に近い自然な」会話を収録しており、本稿の目的である「対面会話での相槌」の分析に用いるのに適切であるといえる。

本稿ではこの12会話×10分のうち、12会話すべてについて最初の2分間(0秒~120秒、計24分間)を対象に分析を行う。なお、このコーパスに収録されている会話は録音の直前に会話のテーマが与えられている。例えば、3人会話コーパスのchiba0132, chiba0232は以下のような会話で始まる。本章以降は、用例表記は付録Aに従う。

- (6) 1 C: でね:恋の話なんですけど  
2 C: わたしの話ではなく:  
3 C: もちろん

4 A: <笑>

5 C: (F\_\_あのね:)(D\_\_ア)塾講の:ことなんだけど:

(『3人会話コーパス』 chiba0132より)

(7) 1 B: (F\_\_えー)トピックは(0.236)腹の立つ話です

2 B: はい

3 B: Aさん

4 A: <笑>

5 A: あー

6 A: 腹の立つ話ですか

(『3人会話コーパス』 chiba0232より)

(6)、(7)ではそれぞれ恋の話、腹の立つ話について話すことが事前に決められていることがわかる。ここでchiba0132~chiba1232の12会話でどのような話題についての会話が行われたのかを表1にまとめる。各

表1 各会話の話題

会話番号	話題
1	恋の話
2	腹の立つ話
3	くさい話
4	腹の立つ話
5	くさい話
6	大事件の話
7	恋の話
8	くさい話
9	くさい話
10	恋の話
11	腹の立つ話
12	腹の立つ話

会話では(6)、(7)、表1のように話題が与えられているが、会話の進行状況によって相槌の頻度などが変化するのには先行研究が存在しないため不明である。そのため、今回の分析では機械的に分析区間のはじめを0秒とした。どの会話も話題提供→会話スタートという形式をとっているため、会話のスタートは基本的には(6)や(7)に似た形であり、分析区間の開始時間による相槌の偏りはないと考えられる。

次項では、データから相槌を抽出するにあたり本稿での相槌を定義し、立場を明確にする。

### 3.1.2 本稿における相槌の定義と抽出

相槌を抽出するにあたり、本稿で分析する相槌について定義を明確にしなければならない。本稿では、相槌の定義について大浜・西村(2005: 79)と基本的に同じ立場をとる。具体的には先行研究を参考に、会話に

において相槌を規定する基準として語の「形式」とそれに対する会話上の「条件」の二つを設けた。形式については(a)~(e)のいずれかであり、条件のうち最低一つを満たすものを「相槌」と定義して、コーパスから相槌を抽出した。定義の文章表現についてはほぼ大浜・西村を踏襲しているが、一部筆者が修正を行った。

## I 形式

- (a) 「うん、あー、ええ」などのいわゆる応答詞
- (b) (a)に類する「なるほど」、「そうですね」、「本当」等の語
- (c) 相手の発話の全部あるいは一部分の「繰り返し」
- (d) 相手の発話の一部の「言い換え」
- (e) 相手の発話を途中から引き受ける「文完成」

## II 条件

- (f) 話し手が発話権を行使している間に聞き手が送る表現
- (g) 相槌に対する相槌といえるもの
- (h) 少し間\*6をおいて取られたターンの開始時に現れる表現
- (i) 質問への応答表現

まず、Iの「形式」については堀口(1988)など複数の研究者が(a)~(e)とほぼ同じ形式\*7を採用していることから、本稿の定義に含めることは妥当と判断した。I(a)で「応答詞」という表現を用いているのは、大浜・西村がもともと使用した「相槌詞」という表現は語義矛盾に陥るためである。そこで、友定(2015)が「質問や命令・依頼など隣接ペアを構成する談話の後件に現れ、談話文脈を断絶する割り込みには当たらない」と定義しているところの「応答詞」という表現が最も近いと考え、書き換えをした。(a)は後述する用例(10)の3行目、(b)は(13)の3行目、(c)は用例(20)がそれぞれ当てはまる。これらは一見してすぐにこの形式だと分かるものなので詳細な説明をしないが、(d)「言い換え」と(e)「文完成」については具体例を一つずつ挙げて説明を行う。

まず、「言い換え」については以下のような相槌が典型例といえる。

- (8) (Aが「腹の立つ話」について話題の提示を求めたことに対しての会話。CはユニクロでバイトをしておりBはそれを知っている)
- 1 B: どうですか
  - 2 B: じゃ:先(0.516)ユニクロ班として
  - 3 C: 接客として

(『3人会話コーパス』chiba1132より)

ここでは、Cがユニクロでバイトで接客をしていることを知っているBが、Cに「ユニクロ班として」と話題を求めたことに対し、Cが「接客班として」と相槌を打っている。ここでは、「ユニクロで働く=接客をする」という関係である。このようなケースでは言い換えの相槌とする。

次に、「文完成」の相槌については以下のような例が挙げられる。

\*6 ここでの間とは、沈黙ではなくタイムラグを示すものを指す。付録Aの記号では(.)や(0.3567)であらわされるものを指す。

\*7 堀口(1988)はここでいう「文完成」を「先取り」と表記するなど、表現の揺れは存在する。

- (9) 1 A: ボッ [ クスがあるの  
2 C: [あるあるあるあるあるあるあるあるあるあるある

(『3人会話コーパス』 chiba0532より)

これはAが「ボックスがある」と言おうとしていることに対してCが先取りをして相槌を打ったといえる。このような相手の意図を引き継いでいるようなとき、これを文完成という。

次に、[II]の相槌であるための「条件」について議論をする。

まず(f)は、メイナード(1993: 58)の定義を大浜が適用したものである。例えば、以下の用例(10)の3行目および4行目はこの条件の例である。このような相槌は日本語の相槌として典型的なものである。

- (10) 1 C: で文学を教える時に:(0.837)こう例えば宇野千代が出てきたらさ:  
2 C: もう宇野千代の恋愛遍歴について多少はこう言及しなきゃいけない [ とか:  
3 B: [ うん  
4 A: うん  
5 C: そういう時に(0.307)ほらこういう時ってこうじゃないって話をしなきゃいけないじゃない

(『3人会話コーパス』 chiba0132より)

つまり、Cが発話権を行使している間に聞き手であるAとBがCに対して相槌を打っている。

(g)は、Iwasaki (1997)がBBC(back-back-channel)と呼ぶもので、以下の用例(11)の3行目の相槌がそれに当たる。ここでは、Bが自身の発話が完了した後にCの相槌に対して相槌を打っている。

- (11) 1 B: やってないやってない  
2 C: へえ:  
3 B: うん

(『3人会話コーパス』 chiba0332より)

このような例を相槌と認めることによって聞き手が打つものだけではなく聞き手以外の会話者が打つ相槌も拾えるようになるという(大浜(2005))。この観点は3人の会話を分析する本研究において重要であるため、本稿でもこれを相槌の条件として採用する。

(h)は、Clancy et al. (1996)が backchannel のサブタイプとして分類している resumptive opener にあたるものである。具体的には、以下の2行目の例が挙げられる。

- (12) 1 A: カフェコンパーナとか [ あの辺がそうなるかな  
2 B: [ うん  
3 B: なんか: そう  
4 B: 名前が違う

(『3人会話コーパス』 chiba0632より)

このように、相槌を起点に発話権の交替が起こっている場合、それをターン交替と呼ぶ。このような形式はメイナード (1993) では「間発話順番状況」として相槌の定義からは除外されている。「間発話順番状況」

の中ではターンをとるよう指示する言動が含まれている場合が多く、発話者と聞き手をその行その行で特定することができないため、(I)の定義上では相槌にはならないとしている。しかし、用例(12)からも分かる通り2者の会話では発話者と聞き手、および発話権の交替を確認することができる。また、大浜(2005)によると、直前のターン終了時に間あるいは沈黙があり、会話に小さな中断ができてでも resumptive opener により中断前の発話に直接関連付けることができるという。以上を踏まえて、本稿ではターン交替に先行するものも相槌と認め、これを条件に加える。

(i)のように質問への応答として現れる「はい、うん、ええ」は、先行研究では相槌とはみなされていない場合が多い(堀口(1997)など)。しかし、大浜・西村(2005)によると、このような応答詞が現れる際の疑問文は、質問より相槌の表現の中の「繰り返し」や「言い換え」に近いものだという。このことから、(h)のような「はい、うん、ええ」を同意や納得の相槌ととらえることができる。

以上の形式や条件に従い相槌の抽出を行った結果、355トークンを相槌と認定できた。

### 3.1.3 数理的処理を行うためのラベリング

抽出したトークンひとつひとつに対して相槌の機能の分布や関連要因の分析を数理的に行うために、先行研究を参考にして以下のような情報を付与するラベリングを行った。「～か否か。」と表されているものは否=0としてダミー変数を付与した。

1. 形式・・・定義(形式)5つのうちどれに当てはまるか。
2. 区切れダミー・・・発話の途中で打たれた相槌か否か。
3. 機能・・・相槌のが持つ機能。これについては3.2で具体例を示しながら議論する。
4. 新情報ダミー・・・相槌が対話の中で初めて登場する話題や言葉に対して打たれたか否か。
5. 相槌の直前の語
6. 品詞・・・相槌の直前の語の品詞。
7. 重なり・・・相槌が発話権を持つ人の言葉と重なっているか否か。
8. 相槌の相槌・・・自分ではない人が打った相槌に対する相槌か否か。
9. 相槌の重なり・・・相槌の相槌であると認められた際、その相槌が重なっているか否か。

以下ではこのラベリング項目に対しての説明を行う。まず、1の「形式」は3.1.2での述べた相槌の定義の中にある形式の5つと対応している。この5つのどれであるかを識別するため、「応答詞」「なるほど」「繰り返し」「言い換え」「文完成」を変数名としてダミー変数を付与した。

2の「区切れダミー」は、相槌が発話の途中で間があるか否かを判別するものである。例えば用例(13)では間があるため1、(14)では間がないため0を付与する。

- (13) 1 B: 全然違う道にいった  
 2 C: <笑>  
 3 A: ふうん

(『3人会話コーパス』chiba0332より)

- (14) 1 B: なんかね:(0.211)サポーターがすごい蒸れるの  
 2 B: ひざ [ ひざにあてるやつ ]

3 A: [ あー ]

(『3人会話コーパス』chiba0332より)

基本的には、トランスクリプト上で(.)または(0.396)のように間あるいは極めて短い沈黙、あるいは(13)のような間があるとされているところと相槌が重なるところを1とした。また、ラベリングの際はすべての会話について筆者自身が音声聞いており、トランスクリプト上は記されていないが明らかに間があると判断される場合は1とした。

ラベリング項目の「機能」については先行研究とデータの照らし合わせにより決定したが、3.2で詳しく検証する。

次に、4の「新情報ダミー」について、これはChafe (1987) の主張に基づくものである。Chafe は、会話の中で人が事物を指すときに、その現れ方は情報の活性段階に応じて activated, semiactivated, nonactivated の三通りの状態があると指摘した。Chafe は、activated information はすでに言及し継続的に言及する事柄を、semiactivated information はすでに言及したが別の事柄を挟んだ後で再び言及するものを、nonactivated information は初めて言及する事柄を、それぞれ意味するとした。本稿ではこれを参考にし、情報を新出のものと既出の二通りに分けることとする。用例(15)の7行目の相槌は、1行目の既出情報を受けて打たれている。

(15) (Bがアルバイトの同僚に怒りを覚えた話)

1 B: (W\_\_ [ ンマジ—まじ](0.366)まじ腹立った

2 A: [ ひでえな

3 A: [ それ

4 B: あれはね [ :(0.401)あれはありえない

5 C: (D\_\_エ)

6 C: <息>

7 A: [ うん

(『3人会話コーパス』chiba1232より)

(15)では、引用部分の前後20秒はBがアルバイトに対し怒りを覚えた話が続いている。ここで、7行目のAの「うん」は5行目のBの「ありえない」という発言に対しての相槌である。Aはそれまでの会話でBが強い怒りを覚えていることを知っているため、ここでのBの「ありえない」はAにとって既出の情報である。よって、このA「うん」に対して既出情報としての0を付与する。

ラベリング項目の5の「相槌の直前の語」、6の「品詞」について説明する。後述する「相槌の重なり」が認められるものについては重なっている語の一つ前の語を直前の語とした。付与した品詞は以下の通りである。

- 名詞
- 動詞
- 形容詞
- 副詞
- 接続詞

- 助動詞
- 助詞
- 格助詞
- 接続助詞
- 副助詞
- 終助詞
- 係助詞
- 連体詞
- 感動詞

例えば(15)の7行目の「うん」に対しては、直前の語は「ね」を、その品詞は「終助詞」を付与した。

助詞はさらに細分化し、基本的には格助詞、接続助詞、副助詞、終助詞、係助詞のいずれかに分類したが例外として以下の3トークンは単に「助詞」と分類した。その3トークンを以下に示す。

- (16) 1 C: 疲れてたら座るん [ だよね  
2 A: [ うんそうそうそう

(『3人会話コーパス』 chiba0532より)

- (17) 1 A: よっぽど日本人はウイナーコーヒーって言うん [ でしょうね  
2 C: [ ふうん

(『3人会話コーパス』 chiba0632より)

- (18) 1 C: でなんか(0.405)あれって思ったん [ だよ  
2 B: [ うん

(『3人会話コーパス』 chiba0832より)

(16)~(18)では、相槌の直前の語はすべて「ん」であるが、これは全て「のだ」(「のでしょう」)の音便変化である。「のだ」の品詞に関しては「のだ」の意味をどのように解釈するかで研究者間でも分かれており、「のだ」を一語とする見方(国立国語研究所(1951))と「の」と「だ」で分けるべきだという見方(林(1964))がある。本稿の目的は「のだ」の品詞を考察することではないので、(16)~(18)で現れる「ん」(←「の」)は助詞としておく。

次に、ラベリング項目7の「重なり」について、具体的には用例(19)のような場合は発話の重なりがないので0、(18)のような場合は「うん」が「だよ」と重なっているため1とした。

- (19) 1 B: (F\_\_あーのね:)(R\_\_後藤)さんに代わってデバの録画をしてる(0.396)してたときがあったの一回  
(.)  
2 A: あーうん (.)  
3 C: うん  
4 B: で世話してんの忘れて歌ってたのね

(『3人会話コーパス』 chiba0932より)

つまり、相槌がその瞬間における主たる話し手の発話と重なっているかどうかで判断をする。

ラベリング項目8の「相槌の相槌」は、3.1.2の(g)の条件にあてはまるものであり、具体例として用例(20)の5, 6行目がある。

(20) (Aが「腹の立つ話」について話題の提示を求めたことに対して)

- 1 B: どうですか
- 2 B: じゃ:先(0.516)ユニクロ班として
- 3 C: 接客として
- 4 C: 接客とし [ て
- 5 B: [ 接客として
- 6 C: 接客として

(『3人会話コーパス』 chiba1132より)

これらのBとCの一連の「接客として」は、2行目のBの「ユニクロ班として」をCが「言い換え」た相槌「接客として」を二回続けて言った後、Bがそれに誘発されるように5行目で「接客として」と「繰り返し」の相槌を打ったと考えられる。この状態が「相槌の相槌」である。また、6行目のCの「接客として」も同様に「相槌の相槌」である。

ラベリング項目9の「相槌の重なり」は「相槌の相槌」が認められる際に相槌と相槌が重なっているか否かを判定した。具体的には、上述のような用例(20)の5行目の「接客として」では1を、6行目では0を付与した。

### 3.2 相槌の機能の再考

前節では本稿の分析で用いるデータとラベリングの方法についての説明を行った。本節では残るラベリング項目の「機能」について2.1の堀口(1997)、メイナード(1993)、および大浜・西村(2005)と大浜(2006)で示された相槌の機能の説明とデータと照らし合わせ、その妥当性を検証する。

会話データに現れた相槌の機能は、概ね先行研究で指摘されたもので説明することができるが、一部その表現に修正が必要だと考えられる部分がある。さらには、先行研究では指摘されていないが付け加えるべきであろう機能も存在する。そこで、最初に先行研究で指摘された7つの機能について説明し、その次に付け加えるべき2つの機能について指摘を行う。

まず、堀口(1997: 52)の「聞いているという信号」とメイナード(1993: 160)の「「続けて」というシグナル」は同一のものとみなせる。つまり、聞き手が発話者に対して「私はあなたの話を聞いているのでどうぞ続けてください」ということを意味する相槌だと判断できる。本稿ではこの相槌の機能を「促し」と呼ぶこととする。具体例として、以下の用例(21)が挙げられる。

- (21) 1 B: でもねウイーンの(0.68[3](F\_\_あの:)(F\_\_その)大通り(0.157)有名な所(0.52[3]には(0.172)全部揃ってるんですよ  
2 C: [うん [うん

(『3人会話コーパス』 chiba0632より)

ここでは、発話権を持つBに対して聞き手のCが2回相槌を打っているが相槌の後もBは発話権を失うことなく話を続けている。このことから、ここでのCの相槌は「聞いている、続けて」という「促し」の相槌だと解釈できる。

堀口(1997: 53)の「理解しているという信号」とメイナード(1993: 160)の「内容理解を示す表現」も同一のものとしてよいだろう。これは、相手の話の内容に理解・納得の意を示すことであるから、納得の相槌と呼ぶこととする。具体例として以下の用例(22)の5行目が挙げられる。

- (22) 1 A: でもウイナーコーヒーってほんとはウイナーコーヒーじゃないんです(D\_\_ス)ってね  
 2 B: うん  
 3 B: なんか(D\_\_ム)(.)別の名前があるらしくて  
 4 A: そうそうそう  
 5 C: ふうん(?)  
 6 B: でメニューにないの

(『3人会話コーパス』 chiba0632より)

(22)での話題はウイナーコーヒーであるが、5行目ではBの「別の名前があるらしい」という情報についてCが「ふうん」という言葉で返している。さらに、この「ふうん」のイントネーションは上昇音調である。つまり、「別の名前がある」という情報について「なるほど、納得」という信号を送っていることになる。この納得の信号をうけて、BはCが自分の情報に納得していることを理解し、「メニューにない」という新たな情報を出していると考えられる。したがって、(22)の5行目のような相槌を「納得」の相槌とする。なお、(22)の2行目のBが打った相槌の「うん」は続く3行目の発言につながるので、先に2.1.3で言及した「ターン交替」の相槌と考えられる。

次に、堀口(1997: 56)の「同意の信号」とメイナードの「話し手の判断を支持する表現」「相手の意見、考え方に賛成の意思をする表現」は同意の相槌として同一視できる。この機能の例は、上記の用例(22)の4行目である。ここでは、Bの「別の名前があるらしい」という情報について「そうそうそう」と相槌を打っている。これは、Aも「ウイナーコーヒーではなく別の名前がある」ことをBと同じく知っており、それをBに同意として示すために打ったものと考えられる。

堀口(1997: 59)の「感情の表出」とメイナードの「感情を強く出す表現」は感情を表す機能として一本化できる。しかし、堀口(1997: 57)の「否定の信号」に関しても不快感という感情としてとらえることも可能かもしれない。そこで、会話データの中から感情を表していると思われるものをさらに抽出した。その結果、感情を表すものは「驚き・気づき」に関するものと「やや否定的な感情」の二つに分けることができた。さらに「やや否定的な感情」は「考える」と「疑問」に細分化できた。具体例を以下に示す。

- (23) 「気づき」(アルバイトの話題)  
 1 C: そう  
 2 C: そいすごい(0.864)(W\_\_キト|人)がいて  
 3 C: なんか(0.489)(F\_\_あの:)ズボンの補正ってあるじゃん  
 4 B: あーあー(W\_\_ア|あー)  
 5 C: (F\_\_あの:)  
 6 A: 裾直しね

- 7 C: 裾直し
- 8 B: うん

(『3人会話コーパス』 chiba0932より)

(24) 「考える」(会話の冒頭付近でテーマは「腹の立つ話」と言われて)

- 1 A: なんかあります
- 2 A: どうぞどうぞ
- 3 B: あー
- 4 C: あー
- 5 A: いいね
- 6 B: どうですか
- 7 B: じゃ:先(0.516)ユニクロ班として
- 8 C: 接客として

(『3人会話コーパス』 chiba1132より)

(25) 「疑問」(新しくサークルに入るかについて)

- 1 C: 余裕だよ:
- 2 A: <笑>
- 3 C: でも三年生からさすがにいないよね:
- 4 A: <笑>
- 5 A: うーん
- 6 A: 二年ならまだね:
- 7 B: (W\_\_ウー|うん)
- 8 C: うん
- 9 C: うんうんうんうん
- 10 B: うん
- 11 A: も今のサークルがちょっと(0.592)こう(0.217)厳しい(0.186)じゃないけど
- 12 A: 縛りが強い(0.241)から
- 13 C: そうなの?
- 14 A: うーん
- 15 A: なかなか休むとと%か:(0.361)まずい

(『3人会話コーパス』 chiba0332より)

用例(23)では、Cが提示した「裾直し」という話題についてBが思い出しように「あーあー」という相槌を打っている。これは、Cの頭の中に「裾直し」という概念はあったもののAの言葉によって気づかされたと考えられることができる。なお、8行目の「うん」は納得の相槌と解釈できる。これによって、Bが「裾直し」という概念に4行目で気づき、8行目でそれを確認したと考えることができる。

用例(24)は用例(8)につながる会話である。ここは冒頭付近で会話のテーマが「腹の立つ話」とであると提示された場面だが、この提示に対して、Aが関連する事例の提示をほかの二人に求めている。この要求に対

しBは、すぐに思いつかないことの表れとして「あー」という相槌を打っている。そこで、このような「考える」の相槌を打ったと考えられる。

(25)では、Aが「サークルの縛りが強い\*8」と言ったことに対し単純に「そうなの?」と疑問を投げかけたと解釈できる。

2.1.3で挙げた大浜(2005, 2006)の主張する「ターン交替」の機能について、これは3.1.2で示した相槌の定義の条件(h)に相当するものである。相槌を抽出する際の定義として「ターン交替」があったが、これは会話の中で大きな役割を持っている。したがって、ターン交替も相槌の機能の一つとして採用する。

ここまでの議論により、先行研究で主張されていた相槌の機能に関して実際の会話データと照らし合わせた結果、先行研究が指摘する機能が今回のデータでもみられることが示された。しかしながら、先行研究で触れられていない相槌の機能が2つ存在した。それは、「返答」と「繰り返し」である。

「返答」と「繰り返し」は、「ターン交替」と同じく3.1.2で述べた定義の中に含まれ、それぞれ(i)と(c)に相当する。しかし、これらも会話の中で重要な機能を果たしていると考えられる。「返答」については、いうまでもなく聞かれた側が「はい」「いいえ」と意思表示をすることであるので会話上の機能を持つといえる。「繰り返し」については、形式としては明白なので3.1.2では言及しなかったが会話上の機能については以下の例が参考になる。

(26) (アルバイト先で不満を言ってくる同僚について)

- 1 A: 面と向って言うの
- 2 B: 面と向って言う
- 3 A: まじで?
- 4 B: うん

(『3人会話コーパス』 chiba1132より)

このように、相手の言った発言をそのまま繰り返すことで相手への確認、自分の発言の強調をしていると考えられる。なお、4行目の「うん」は返答の相槌と解釈した。

以上により、会話データの中にある相槌の機能を網羅的に提示することができた。本稿で考える相槌の「機能」を以下に整理する。なお、分析方法の都合上、すべての相槌トークンについて付与する機能は1つのみとした。

- 促し
- 納得
- 同意
- 気づき
- 考える
- 疑問
- ターン交替
- 返答
- 繰り返し

---

\*8 参加が強制されていること

### 3.3 基礎分析

本稿の目的である相槌の機能を決める要因を分析するにあたり、3.2で述べた相槌の機能が会話の中でどのくらいの頻度で現れるのかを確認する必要がある。そこで、抽出を行ったトークンを機能についてグルーピングを行い数え上げを行った。結果を以下表2に記す。表2より、日本語の対話においては相槌の機能と

表2 相槌の機能の頻度

機能	個数	割合	累積割合
促し	137	38.6%	38.6%
返答	46	13.0%	51.5%
同意	46	13.0%	64.5%
納得	42	11.8%	76.3%
繰り返し	23	6.5%	82.8%
ターン交替	23	6.5%	89.3%
気づき	17	4.8%	94.1%
考える	17	4.8%	98.9%
疑問	4	1.1%	100.0%
合計	355	100.0%	

して「促し」が全体の38.6%を占め最も多いことがわかる。累積割合で見ると、「促し、返答、同意、納得、繰り返し、ターン交替、気づき、考える」までで全体の98.9%を占めており最後の「疑問」の相槌がほとんど使われていないことがわかる。この疑問の相槌を外れ値とみなし応用分析から外すこともできるが、本稿の応用分析では相槌の機能全体を分析することを目的としているためそれをせず、9つの機能がどのような要因で決まるのかを分析する。

### 3.4 応用分析

前節では、日本語の対面会話の相槌には「促し」の機能が多く使われていることを述べた。本節では、促しを含めた相槌の「機能」が会話のどのような要素で決定されるのかについて分析する。

#### 3.4.1 分析手法について

決定要因を分析するにあたり、分析手法が満たすべき条件として以下の三つが挙げられる。

1. 単なる相関関係ではなく、説明変数→機能というように因果関係として説明できること。
2. 抽出した説明変数によってデータがどの程度説明できるかを数値として算出できること。
3. どの説明変数が機能の決定にどの程度貢献しているかの指標が得られること。

本稿では以上の条件を満たす分析手法として Gradient Boosted Decision Tree を用いる。 Gradient

Boosted Decision Tree は、機械学習<sup>\*9</sup>の手法の一つであり弱学習器を決定木とする勾配ブースティングのことである。また、勾配ブースティングは統計的機械学習の手法の中でも集団学習<sup>\*10</sup>の一つである。決定木および Gradient Boosted Decision Tree の詳細なアルゴリズムは付録Bを参照されたい。ここでは、それぞれについて簡単な説明を行った後に Gradient Boosted Decision Tree が上記の三つの条件を満たしていることを説明する。

決定木は分類木と回帰木に分けることができる。目的変数が離散値である場合は分類木を、連続値である場合は回帰木を用いる。本分析では目的変数は相槌の機能であり、さらに各機能に上下関係がないことから名義尺度であるといえる。つまり離散値であるといえるため分類木を用いる。例えば、3つのクラスA, B, Cを2つの説明変数 $x_1, x_2$ で分類する分類木を図1のように示す。付録Bで述べる CART アルゴ

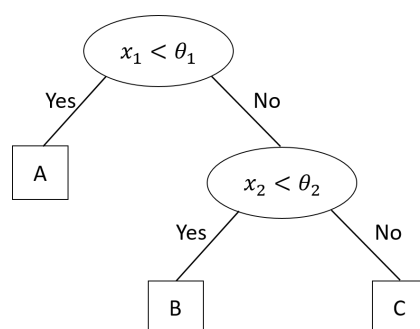


図1 分類木の例

リズムから重要度の高い説明変数<sup>\*11</sup>を算出し、その値の条件にあてはまるかどうかで分類を行う。この場合、 $x_1 < \theta_1$ ならばAに分類され、 $x_1 < \theta_1$ かつ $x_2 < \theta_2$ ならばB、 $x_1 < \theta_1$ かつ $x_2 < \theta_2$ ならばCに分類される、ということになる。

弱学習器を決定木とすることにより、本項の冒頭で述べた三つの条件を満たしたことになる。1については、図1からも分かる通り分類過程で変数をフロー化しているため、相関関係ではなく一方向の因果関係としてとらえることができる。3についても、アルゴリズムの中で変数の重要度を算出しているため、説明変数の重要度について指標を得ることができる。2については、分類のアルゴリズムに従って学習データ(in sample)を再び分類することや学習には用いていないデータ(out sample)に対する分類予測を行い、正解データと照らし合わせることで予測的中率を得ることができ、モデルの精度の指標およびデータをどの程度説明できているかの指標とすることができる。

Gradient Boosted Decision Tree は先にも述べた通り弱学習器を決定木とした集団学習の一つである。この手法では、決定木を逐次的に構築することで予測精度を高めている。つまり、すでに学習し終わった分類木で正しく分類できなかったデータを優先的に学習し次の分類木を構築する。このときに分類木から次の木に重み付けをしていることになる。Gradient Boosted Decision Tree は分類木を構築する際、学習データからランダムにサンプリングを行うことで、汎化性を保っている。以上のことを図示すると図2のように

<sup>\*9</sup> 音声対話システムや人工知能の基礎となる分野である。既存のデータから学習し、未知のデータ(out sample)への予測を行う。機械学習のモデルの性能を判断する指標として、学習データ(in sample)へのあてはまり具合とout sample への予測力がある。詳細は後藤・小林 (2014)を参照。

<sup>\*10</sup> 複数の性能の低い学習モデル(弱学習器)を統合して高い精度の予測を得る手法。

<sup>\*11</sup> 機械学習の分野では統計学でいうところの説明変数を特徴量と呼ぶのが普通であるが、本稿では便宜上説明変数で統一する。

なる。

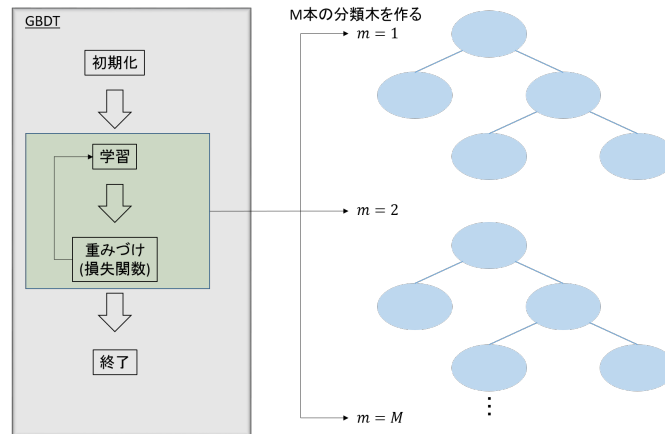


図2 Gradient Boosted Decision Tree の概略図

以上のようなGradient Boosted Decision Tree を用いて、3.2で議論した相槌の機能9つについて決定要因を分析した。先にも述べたとおり、相槌の個々の機能ごとに決定要因があるのではなく、9つの機能がどのような変数に強く影響され決定されるのかという点にのみ着目をする。説明変数としては、3.1.2で述べた定義のうちの形式5個、3.1.3で述べた相槌の直前の語の品詞14個、同項でラベル付けをした区切れダミー、新情報ダミー、重なり、相槌の相槌、相槌の重なりの合計24個とする。

### 3.4.2 分析結果

分析には統計ソフトの R、およびパッケージの Xgboost を利用した。Xgboost では Gradient Boosted Decision Tree を調整するためのいくつかのパラメーターが存在するが、その詳細は付録Bに記載する。モデルの評価方法には  $k$ -fold cross validation を用いた。 $k$ -fold cross validation では、学習データをランダムに  $k$  分割しその一つを使いモデルを構築、残りのデータで正解率を求め全ての平均をモデルの精度とする手法である。ここでは、 $k = 60$  として分析を行った。<sup>\*12</sup>

分析の結果、in sample で74.7%を正しく判別できた。これは、このデータに対しこのモデルの精度が74.7%であることを意味している。このとき、付録Bの(21)式によって求められた説明変数の重要度(Gain)を表3に示し、Gain をグラフ化したものを図3に記載する。表中の feature は説明変数、Gain は説明変数の重要度、Cover はその変数で説明できるデータに対する割合を示す。表に関してはすべての値に対し小数点第4位で四捨五入をした。なお、表に示したのは24個の説明変数のうち feature の値に関して上位20個である。図3では縦軸に説明変数名、横軸に Gain をとり、 $k$ -means 法によるクラスタリング<sup>\*13</sup>によって Gain の高い変数と低い変数に分類をして色分けをしている。

表3および図3に示した通り、相槌の機能を決定するのに最も強く影響しているのは新情報ダミーである。つまり、相槌の直前に示されている情報が聞き手にとって新しいかが相槌の機能決定に強く影響し、それが相槌の機能全体の13.2%を説明していることがわかる。次章では、Gainが0.1より大きい5つの変数に

<sup>\*12</sup>  $k$  の値がサンプルサイズと等しいとき、それは学習データから1サンプルのみを抽出してモデルを構築するのと同じである。一般に、 $k$  を大きくするとモデル精度が上がるといわれるが、その分計算コストもかかる。

<sup>\*13</sup> クラスタリングおよび  $k$ -means 法についての詳細は平井 (2012)を参照。

Feature	Gain	Cover
新情報ダミー	0.155	0.132
相づちの相づち	0.146	0.131
格助詞	0.128	0.127
助詞	0.122	0.122
動詞	0.106	0.111
係助詞	0.070	0.084
なるほど	0.057	0.041
名詞	0.047	0.057
副助詞	0.034	0.048
接続詞	0.031	0.427
重なり	0.028	0.277
繰り返し	0.027	0.021
副詞	0.016	0.019
相づちの重なり	0.015	0.015
言い換え	0.007	0.004
連体詞	0.005	0.010
終助詞	0.003	0.006
文完成	0.001	0.000
区切れダミー	0.001	0.001

表3 説明変数の重要度

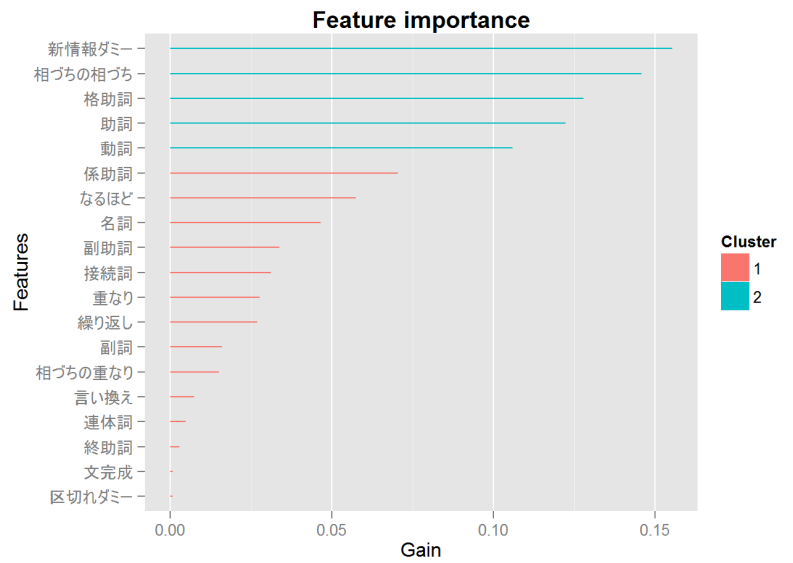


図3 説明変数の重要度

ついて、その理由などを考察する。

## 4 分析結果の考察

本章では、3.4.2で得られた結果についての考察を行う。Gain の値が0.1を超えた5つの変数は新情報ダミー、相槌の相槌、格助詞、助詞、動詞である。まず、これらの変数の特徴について考察する。新情報ダミーは相槌の直前の語が聞き手(相槌を打った者)にとって対話の中で新しい情報であるか否か、相槌の相槌はその相槌が相槌によって誘発されたものか否か、格助詞、助詞、助動詞は相槌の直前の語の品詞を表す。つまり、これらの変数は全て相槌の対話において相槌の直前の語の状態を表す変数である。直前の語を表すものではない変数は Gain が低い。例えば相槌の形式を表すものや区切れダミーなどは直前ではなく相槌そのものの状態を表す変数である。さらに、表3の Cover に注目すると5変数のCover の和は0.616である。この数字はこれらの5つの変数だけで相槌の機能の決定要因の61.6%を占めていることを意味する。Gain の高かった変数は全て相槌の直前の語の状態を表す変数であることとこれらの変数が相槌の機能の決定要因の61.6%を占めていることから、相槌の機能は相槌の直前の語の状態で決定されると考えられる。

では直前の語の状態とはどのようなものか、以下では変数別に考察を行う。

### 4.1 新情報と旧情報

相槌の直前の状態が相槌を打つ者にとって新情報である場合も旧情報である場合も、9つすべての機能が確認された。つまり、どの機能に対しても情報の新旧が要因となっていることになる。そのため、情報の新旧が直接機能の決定に影響したのか、それとも何か他の変数の影響も受けているのか判断が困難である。そのため情報の新旧は機能の決定に強く関係しているが、直接的に関係しているというよりも間接的に品詞の特性などを介して関係していると言う可能性を考えて、考察を進めることとする。本章の残りの節では、直前の情報の新旧と相槌の機能がどのように関係しているのかに注目しながら相槌の直前の状態について考察する。

### 4.2 相槌の相槌

「相槌の相槌」では、疑問以外の機能は全て現れるという結果になった。そもそも、相槌の相槌は誰かの相槌に対して自分が呼応して相槌を打つことと定義した。そのため、そこには必ず「誰かの相槌」という新情報がある。必ず新情報があるのだから、「相槌の相槌」という機能に対しては直前の情報の新旧を識別できない。よって、機能の決定に関しては、新情報ではなく「相槌が打たれた」という状況そのものが関係していると考えられる。

### 4.3 格助詞

直前の語が格助詞である相槌にみられた機能は多い順に促し(16トークン)、返答(3トークン)、同意(2トークン)であった。それぞれ1例ずつを示しながら考察する。

(27) 1 B: ええ:

2 C: トピックは(W\_ホエ|恋)の話です

3 A: <息>

- 4 B: <笑>  
 5 B: <笑>  
 6 A: ええ  
 7 A: どうぞ  
 8 C: えっ  
 9 C: 何(D\_ソツツ)  
 10 B: <笑>  
 11 C: ええ  
 12 B: <笑>  
 13 C: そっか  
 14 B: なんでこれどういう [① 仕組みにな:の  
 15 C: [① でも今(0.241)(W\_タ|だ)から:(1.1 [② 49)(W\_タ|だ)から今恋の話で  
 すって言った時点で: [③ (0.62)こう(0.293)お%とだけ聞いている人には: [④ (0.2)どっちの話かわか  
 らないから [⑤ (0.465)魚の話をしてても(0.599) 話に沿ってると思ってきっとやってたんだね  
 16 A: [② さあ  
 17 A: ん  
 18 A: [③ うん  
 19 B: [③ うん  
 20 A: [④ うん  
 21 A: [⑤ うん  
 22 B: [⑤ ふん  
 23 B: <笑>  
 24 A: そうですね:  
 25 A: まあそりゃ [⑥ そうすよ  
 26 C: [⑥ うん  
 27 C: 腹の  
 28 A: でも  
 29 B: まあでも普通に恋の [⑦ 話つって鯉を [⑧ 思い浮かべる人は少ないよね  
 30 A: [⑦ うん  
 31 A: [⑧ そう

(『3人会話コーパス』 chiba1032より)

ここで注目するのは30行目のAによる「うん」である。ここはchiba1032の会話冒頭部分で話題である恋の話からどのように話を広げるかについての会話である。1~25行目まで、「恋」の話なのか魚の「鯉」の話かという話がなされ、「恋」の話であるという結論に達している。30行目のBの発話に対するAの相槌は「聞いているから続けて」という促しの機能を持つが、どちらも相槌の直前の語が格助詞(それぞれ「の」と「を」)だということが分かる。30行目の段階でAは「鯉」ではなく「恋」の話題であることを認知しているため、Bが「普通に恋の」と言った時点でAはBが「話」という言葉を発しようとしていることを知り、「話」という言葉が続くのは予想できるから続けて」という相槌を打ったと考えられる。また、以上の

ことから、格助詞の後に現れる相槌が情報の新旧に強く依存しており、機械的に格助詞の後なら相槌が現れているのではなく、相槌を打つ者の対話の中での情報の認知状況が深く関わっているであろうことも分かる。

- (28) 1 C: トピックは(0.442)腹の立つ話です  
 2 C: (F\_\_はい)  
 3 C: どうぞ  
 4 A: おい  
 5 B: ふ [① うん  
 6 A: [① 早く言えよ  
 7 C: <笑>  
 8 A: (W\_\_クナロ|このやろう)  
 9 B: 俺が (?)\*<sup>14</sup>  
 10 C: う [② ん  
 11 A: [② お前(W\_\_ヤオ|やろ)  
 12 B: <笑>  
 13 C: <笑>

(『3人会話コーパス』 chiba0432より)

(28)も会話の冒頭部分である。ここでは10行目のCによる「うん」という相槌が返答の相槌である。Bに最初に話題の提示を求めたのは3行目のCの「どうぞ」という発言であるが、そのあとBが迷うような表現を見せずにAとCが催促していることから考えると、この時点ではBは自分が話題を振られたことにまだ気づいていない。また、Cは「早く言えよ」と発言していることからBが話題を振られていることに気づいていないことを認知しているといえる。よって、Bの「俺が(?)」という発言に対するCの「うん」という返答は旧情報に基づく相槌であるが、「返答」が起こったのはまぎれもなく質問に対してである。つまり、10行目の「うん」という相槌に情報の新旧は関係していないと考えられる。

格助詞の後に同意の相槌が打たれる例は(27)の31行目である。ここでも促しの場合と同様に、AはBが後にする発言をある程度予測していると考えられる。しかし、Bの言う内容に対してはすでに30行目の「うん」で促しをしているので、ここでは促しをするのではなく発言に同意の意を示す相槌を打ったと考えられる。

#### 4.4 助詞

助詞の後に相槌が打たれる例は3.1.3の(16)～(18)である。このうち、(16)は同意の相槌、(17),(18)は促しの相槌である。まずは同意の相槌について考察する。以下に(16)を前後の文脈が分かるように再び記載する。

- (29) (電車のトイレの前がくさいという話題)

\*<sup>14</sup> トランスクリプト上は?記号が存在しなかったが、音声資料では明らかに上昇イントネーションであったため(?)と補った。

- 1 A: [①] で[②] トイレの前に:  
2 B: [①] うん  
3 C: [②] がちゃんで開けるやつでしょ  
4 A: そうそうそ  
5 C: あ  
6 A: ボツ [③] クスがあるの  
7 C: [③] あるあるある [④] あるあるあるあるあるあるあ [⑤] るある  
8 A: [④] あるよね:  
9 B: [⑤] あー:  
10 A: あれね:そう:あれが(0.163)あそこの場所はすごい大きい  
11 C: <笑>  
12 C: あー  
13B: あ  
14 B: くさいんだ  
15 B: [⑥] やっぱ  
16 A: [⑥] うん  
17 A: やっぱ一番人気ないよ  
18 B: <笑>  
19 C: うん:  
20 C: だろうね  
21 A: <笑>  
22 C: <笑>  
23 B: <笑>  
24 A: があって開いて:こうみんな [⑦] 座ってくときに  
25 B: [⑦] あー  
26 C: <笑>  
27 B: <笑>  
28 A: <笑>  
29 B: あー  
30 A: <笑>  
31 C: 疲れてたら座るん [⑧] だよね  
32 A: [⑧] うんそうそうそう

(『3人会話コーパス』chiba0532より)

(29)で分かるように、Cが31行目の発言をしたとき、AもBも「座る」対象は電車のトイレの前にあるボックスであることを理解している。つまり、31行目のCの発言は旧情報である。よってAはCの言わんとしていることを「座る」という発言を聞いた時点で予期し、「座る」の直後に同意の相槌を打ったと考えられる。なお、ここの「んだ」は「のだ」の連用形接続による音便変化であるため、32行目の相槌を動詞の直後の相槌と考えてよく、「ん」はAがCの発言を予期することによるタイムラグであると考えられる。

次に「促し」であるが、これは(18)で考察可能である。以下に(18)を再掲する。

- (18) 1 C: でなんか(0.405)あれって思ったん [ だよ  
 2 B: [ うん

(『3人会話コーパス』chiba0832より)

この「うん」の機能は「促し」であった。「促し」は「聞いているから続けて」ということである。ここでの発話権を行使しているのはCであるので、Cが話し手でBが聞き手である。聞き手が最も必要としているのは「相手が伝えようとしている意味」であるが、1行目の「ん」は助詞であり「だ」と「よ」は終助詞であることから、Cが伝えようとしている意味は「思った」を言った時点で終結していると考えられる。したがって、2行目のBの「うん」は「私はあなたの話を聞いていて「あれって思った」ことは理解したから続けて」という意図で打たれたものだと考えられる。

#### 4.5 動詞

動詞の直後に相槌が打たれた時に現れる機能は「気づき」、「疑問」、「考える」以外の6種類である。本節では、この3種類の機能が現れなかったことについて考察する。

そもそも、3.2で述べた通りこれらの機能は堀口(1997)の「感情の表出」とメイナード(1993)の「感情を強く出す表現」を細分化したものであった。つまり、動詞の直後の相槌では感情の表出はされないといえる。これは、日本語が膠着語<sup>\*15</sup>であることによるものと考えられる。4.4で見たように、相槌を打つ者が話し手の発言に何らかの感情を持つ際、自分の対話における認知状況と照らし合わせるためタイムラグが起りやすく、その結果として動詞の直後の語、多くは助詞や助動詞の連用接続形と相槌が重なることが多いと考えられる。そのため、動詞の直後に感情を表出する相槌は現れなかったことが考えられる。また、4.1や4.4でみたように自分の理解との照らし合わせの際に参照するのは、話し手の発言が対話の中で新しい情報が存在するかどうかであることが多いため、動詞の直後の相槌の機能も動詞という品詞よりむしろ発話の情報の新旧からも影響を受けていると考えられる。

#### 4.6 考察のまとめ

4.1~4.5の考察により、以下のようなことが考えられる。相槌の機能は直前の語が相槌を打つ者にとって新しい情報を持つかどうか、相槌の相槌であるか、直前の語の品詞が何であるかによって出現するパターンがある程度決まっている。さらに、直前の語が新情報が持つかどうかと直前の語の品詞は独立ではなく、品詞によって相槌の機能が決定される場合でも、新情報が存在するかどうか、つまり相槌を打つ者の対話における情報の認知度合いによって決まる。

<sup>\*15</sup> 屈折語、孤立語とならび言語の類型論的三分類の一つ。語の順序や語形変化でなく、助詞・助動詞などの付属語によって文法的機能を示す言語。フィンランド語、トルコ語、朝鮮語、日本語などウラル-アルタイ語族はこれに属する。付着語。粘着語。漆着語。接続語(『日本国語大辞典』)。

## 5 結論と展望

### 5.1 研究のまとめと結論

本稿では、「日本語の対面会話において相槌の機能は会話のどのような要素で決定されるのか」という問いの下、『3人会話コーパス』を用いて相槌の抽出及び分析を行った。抽出にあたり、相槌の定義として大浜・西村(2005)を参考にして単語の「形式」と「条件」という二段階の定義づけを行った。さらに、先行研究で提示された機能を抽出された相槌と照らし合わせることで先行研究では機能とされていなかったものを含め9つの機能を確認することができた。そして相槌の機能の決定要因を Gradient Boosted Decision Tree を用いて分析した結果、変数としては新情報であるか否か、相槌の相槌であるか否か、相槌の直前の語の品詞として格助詞、助詞、動詞が強く影響しており、その5つの変数で相槌の機能決定の61.6%を占めることが明らかになった。また、その結果をもとにデータを振り返った結果、新情報であるか否かが、機能の決定に対して、直接的ではなく、ほかの直前の語である格助詞、助詞、動詞を媒介する形で間接的に深く関連していることが結論として得られた。

### 5.2 今後の課題と展望

まず、課題として Gradient Boosted Decision Tree の in sample に対する予測率である74.7%や、3.4.2では挙げなかったが out sample に対する予測率である45.1%を上げることが考えられる。今回の結果は、データによる制約が大きいと考えられる。2.2にも変数として用いられていた個々の相槌に対する周波数などの音響情報はソフトウェアの制限上抽出不可能であった。そもそも会話は音声によってなされるものなので、このような音響情報が取得できていた場合、予測率は改善されていた可能性がある。

さらに、今回はデータとして存在しなかったため分析の対象にできなかったが頷きなどの非言語的な情報も相槌の機能決定に強く影響していると考えられる。なぜなら、1.2で述べたように、相槌と頷きには強い共起関係があるためである。このような音響情報、非言語的情報による分析の改善は今後の課題としたい。

本研究の展望として、分析結果の日本語教育への応用が挙げられる。例えば、本研究で明らかになった新情報と相槌の機能の関係を応用可能である。日本語学習者が相手の話を聞いているとき、相手の発言が自分にとって新しい情報かどうかによってどのような相槌を打てばいいのかの重要な指標となりえる。さらに、本研究は基礎研究に過ぎないが相槌の機能決定に関する研究が進めば、現在研究が盛んである機械による対話システムや人工知能に相槌を打たせる際の指標としても、今回の研究成果を応用できる可能性がある。

## 付録A トランスクリプトの記号一覧

表4 トランスクリプトの記号

[, [①, ]	音が重なっていることを示す。論述の都合によっては丸文字を付けた。短い発話については [ のみを付けたが、長い音の重なりについては ] で閉じることで重なりを終点を示した。
A, B, C	会話参加者
(0.334)	0.1秒以上の発話単位内休止（数字は秒）
( )	筆者による説明
(.) 0.1	0.1秒未満の発話単位内休止
(F_あの)	フィラー
(R_木村)	固有名を仮名（かめい）に置き換えたもの
(D_スハ)	言いさし（意図された語が不明）
(W_ジュオー 授業)	言い誤りや非標準的な発音
:	非語彙的な音の引き伸ばし
%	非語彙的な音の詰まり
?	上昇イントネーション
-	語の中断
<笑>	発話を伴わない笑い
<息>	呼気・吸気
太字	用例内で注目すべき箇所であることを示す。

## 付録B Gradient Boosted Decision Tree アルゴリズム

3.4で用いた Gradient Boosted Decision Tree について、そのアルゴリズムを述べ、最後に Xgboost パッケージで用いたパラメーターの値について記載をする。

### B.1 決定木(Decision Tree)

Gradient Boosted Decision Tree は決定木を弱学習器とするため、まず決定木について下川他 (2013)を参考にして説明をする。ただし、3.4では回帰木ではなく分類木を用いているため、分類木のみ説明する。決定木のアルゴリズムとして CART 法が有名であり、Xgboost パッケージでもこのアルゴリズムが用いられている。

CART 法のために、データ集合を  $\mathcal{L} = \{(y_n, \mathbf{x}_n), n = 1, \dots, N\} = \{y_n, \mathbf{x}_n\}_{n=1}^N$  とする。ここで、 $N$  は標本サイズであり、 $y_n$  と  $\mathbf{x}_n$  はそれぞれデータ集合  $\mathcal{L}$  内の目的変数と  $P$  次元の説明変数ベクトル  $\mathbf{x}_n = (x_{1n}, \dots, x_{Pn})$  である。CART 法では、次の三つの段階により分類木が構築される。ここで、p.21図1のような木における条件式が記されている部分を「ノード」と呼ぶこととする。

1. 木の成長
2. 木の剪定
3. 最適な木のサイズの決定

#### B.1.1 木の成長

応答変数  $y_n, n = 1, \dots, N$  は分類指標であり、 $K$  個のクラス  $\{1, \dots, k, \dots, K\}$  のいずれかの値をとるとする。このとき、ノード  $t$  において変数  $y_n$  がクラス  $k$  をとる確率は

$$pr(k|t) = \sum_{n \in t} \frac{\mathbf{1}(y = k)}{N(t)} \quad (1)$$

ここで、 $\mathbf{1}(\cdot)$  は括弧内の条件が真なら1、偽なら0をとるインディケーター関数である。このとき、ノード  $t$  における応答変数のクラス予測値は、多数決方式で決まり、

$$\hat{k}(t) = \arg \max_k pr(k|t) \quad (2)$$

である。ノード  $t$  にはデータが含まれているが、その不均一性測度(不純度)を  $R(t)$  とする。不均一性測度の評価には多くの場合ジニ係数が用いられる。ジニ係数は以下の式で表される。

$$r(t) = \sum_{k \neq k'} pr(k|t)pr(k'|t) = \sum_{1 \leq k \leq K} pr(k|t)(1 - pr(k|t)) \quad (3)$$

そして、データの不均一性測度は

$$R(t) = pr(t)r(t) \quad (4)$$

と定義される。

次にノードからその下のノードである子ノードへの分岐ルールを考える。ノード $t$ において、ある分岐規則を $s_t$ と、全ての $s_t$ の集合を $S_t$ とする。このとき、 $s_t$ によって二分割される子ノードをそれぞれ $t_L, t_R$ とする。また、分岐対象のノードは「終結ノード」( $\tilde{T}$ )と呼ぶ。今、説明変数 $x_p$ は名義尺度なので分岐規則 $s_t$ は $s_t \in c?$ ( $c?$ は変数 $x_p$ における分岐候補)である。このとき、ノード $t$ に属するデータ $n$ は分岐規則 $s_t$ を満たせば $t_L$ へ、そうでなければ $t_R$ に送られる。

分岐規則 $s_t$ に対する分岐測度は、不均一性測度の減少量

$$\Delta R(s_t, t) = R(t) - R(t_L) - R(t_R) \quad (5)$$

で測られる。このとき、ノード $t$ で選択される分岐規則 $s_t^* = \{x_{p(t)} \in c(t)?\}$ は、 $S_t$ の中で最大の減少量を与える分岐規則

$$s_t^* = \arg \max_{s_t \in S_t} \Delta R(s_t, t) \quad (6)$$

として選択される。全ての分岐対象のノードにおいて、以上のような分岐が実行され新しい子ノードが生成される。このようにして、木 $T$ は根幹ノード $t_1$ から始まり、分岐を繰り返し通常以下の三つの停止基準に達するまで成長を続ける。

Stop1  $\Delta R(s_t, t)$ が予め規定した値より小さくなる。

Stop2 全ての終結ノード $t \in \tilde{T}$ 内の個体数が予め規定した値よりも小さくなる。

Stop3 終結ノード数 $|\tilde{T}|$ が予め規定した値よりも大きくなる。

### B.1.2 木の剪定

与えられた説明変数 $\mathbf{x}$ に対し、木に基づく予測クラスは $\mathbf{x}$ がたどり着く終結ノードの代表値である。そこで、木 $T$ のデータへの適合度(正解率)は、終結ノード $\tilde{T}$ の不均一性測度の総和となり、これを

$$R(T) = \sum_{t \in \tilde{T}} R(t) \quad (7)$$

とする。 $T$ が増加すれば $R(T)$ は小さくなり正解率が上がるが、木が複雑すぎると過学習を起こす。つまり、終結ノード数 $|\tilde{T}|$ が大きすぎると学習データにのみ正解率が高く予測データへの正解率が著しく低くなる。これを避けるため、木の剪定が必要になる。

木 $T$ に含まれるノード $t$ を根幹とする枝を $T_t$ 、木 $T$ から枝 $T_t$ を刈ることで得られる部分木を $(T - T_t) \cup t$ とする。このとき、木 $T$ の不均一性測度に複雑度を加えることで複雑度コスト $R_\alpha(T)$ を、木 $T (\leq T_{max})$ の複雑度を終結ノード数 $|\tilde{T}|$ とおき、複雑度のパラメーターを実数 $\alpha (\geq 0)$ とすることにより

$$R_\alpha(T) = R(T) + \alpha |\tilde{T}| \quad (8)$$

と定義できる(Breiman et al. (1984))。ここで、 $T \leq T_{max}$ は $T$ が $T_{max}$ の部分樹木であることを表す。この $R_\alpha(T)$ に基づき、任意の実数 $\alpha$ に対して

$$T(\alpha) = \operatorname{argmin}_{T \leq T_{max}} R_\alpha(T) \quad (9)$$

で表される部分木 $T(\alpha)$ を見つけることができる。つまり、任意の実数 $\alpha$ に対して $R_\alpha(T)$ を最小にするような部分木系列

$$T_{max} \succeq T_0 \succ T_1 \succ \cdots \succ T_J = \{t_1\}$$

を構成できる。ここで、 $T_J$ は $\alpha$ が $\alpha_j \leq \alpha < \alpha_{j+1}$ の範囲にあるときの $R_\alpha(T)$ を最小にする $T_{max}$ の最小部分木を表し、 $J$ は $\{T_j\}_{j=0}^J$ の最後の番号を意味する。すなわち、 $T_J$ は根幹ノード $\{t_1\}$ である。

以上より、木 $T_{max}$ に適用される式(9)は次のようにまとめられる。

$$T_{max}(\alpha) = \begin{cases} T_0, & \text{if } \alpha < \alpha_1 \\ T_j, & \text{if } \alpha_j \leq \alpha < \alpha_{j+1} \quad (1 \leq j < J) \\ \{t_1\}, & \text{if } \alpha_J \leq \alpha \end{cases} \quad (10)$$

### B.1.3 最適な木のサイズの決定

木の剪定を行ったら、部分木列 $\{T_j\}_{j=0}^J$ の中で最良の木を決定する。木 $T_J$ の不均一性測度 $R(T_J)$ は分岐と期待損失量の2乗から得られた偏りがあるため、 $R(T_J)$ 自身は木 $T_J$ の期待リスク

$$R^*(T_J) = E[L(y, d_{\tilde{T}_J}(\mathbf{x}))]$$

のよい推定値ではない(Breiman et al. (1984))。ここで、 $L(y_n, a)$ は $y = y_n$ の時に $a$ の行動をとる損失関数であり、分類木では $L(y_n, a) = \mathbf{1}(y_n \neq a)$ である。また、 $d_{\tilde{T}_J}(\mathbf{x}_n)$ は $\mathbf{x}_n$ があるノード $\tilde{T}_J$ に基づいて構築される $y_n$ の予測子である。

CART 法では通常、 $V$ -fold cross validation<sup>\*16</sup> によって $R^*(T_J)$ の良い推定値を構成し、予測性能の良い木を決定する。CART 法の $V$ -fold cross validation では、データ全体 $\mathcal{L} = \{(y_n, \mathbf{x}_n), n = 1, \dots, N\}$ を $V$ 個の部分データ $\mathcal{L}_{sub}^1, \dots, \mathcal{L}_{sub}^V$ にランダムに分割し、 $\mathcal{L}_{cv}^{(v)} = \mathcal{L} - \mathcal{L}_{sub}^v$ を $v$ 番目の学習用データとして使用し、 $\mathcal{L}_{sub}^v$ を検証用データとする。各学習用データ $\mathcal{L}_{cv}^{(v)}$ によって木 $T_{max}^{(v)}$ を構築し、その剪定を行うことで $T_{max}^{(v)}$ の部分木係列 $\{T_j^{(v)}\}$ を得る。ある複雑度パラメーター $\alpha'_j$ によって指定される木 $T_{max}^{(v)}(\alpha'_j)$ に対応する予測子を $d_j^{(v)} = d_{\tilde{T}_{max}^{(v)}(\alpha'_j)}$ , ( $j = 0, 1, \dots, J$ )とする。ここで $\alpha'_j$ は、データ全体 $\mathcal{L}$ を用いる木の成長過程で得られた複雑度パラメーター列 $\{\alpha'_j\}_{j=0}^J$ の隣接するものの幾何平均 $\alpha'_j = \sqrt{\alpha_j \alpha_{j+1}}$ である(Breiman et al. (1984))。以上より、 $\mathcal{L}_{cv}^{(v)}$ を用いて構築された予測子 $d_j^{(v)}$ に、 $d_j^{(v)}$ の構築に使用されていないデータ $\mathcal{L}_{sub}^v$ をあてはめることにより、 $R^*(T_J)$ の交差検定推定値

$$R^{cv}(T_j) = \frac{1}{N} \sum_{v=1}^V \sum_{(y_n, \mathbf{x}_n) \in \mathcal{L}_{sub}^v} L(y_n, d_j^{(v)}(\mathbf{x}_n)), \quad j = 0, 1, \dots, J \quad (11)$$

となる。ここで、予測子 $d_j^{(v)}(\mathbf{x}_n)$ はノード内の最多数クラス $d_j^{(v)}(\mathbf{x}_n) = \hat{k}_t$ である。

予測性能の良い最適な木は、部分木係列 $\{T_j\}_{j=0}^J$ の中から $R^{cv}(T_j)$ が最小値

$$R^{cv}(T_{j_0}) = \min_j R^{cv}(T_j) \quad (12)$$

を持つ木 $T_{j_0}$ が選択される。

以上の過程により、データ $\mathcal{L}$ から最適な分類木 $T$ が構築された。

## B.2 Gradient Boosted Decision Tree

Friedman (2001)を参考に、Gradient Boosted Decision Tree を構築する。Gradient Boosted Decision Tree では、損失関数を最小化するように弱学習器を追加していく。

<sup>\*16</sup> 3.4.2では $k$ -fold と記したが、ここでは $k$ はクラス名を表すので $V$ を用いる。

まず、学習の初期値として

$$F_0(x) = \arg \min_{\gamma} \sum_{i=1}^n L(y_i, \gamma) \quad (13)$$

が最小になるような定数 $\gamma$ をとる。この損失関数は、分類するクラスが $k = 1, \dots, K$ の時、

$$L = - \sum_{k=1}^K y_k \log p_k(\mathbf{x}) \quad (14)$$

である。ここで、

$$y_k = 1(\text{class} = k) \in \{0, 1\}$$

$$p_k = Pr(y_k = 1|\mathbf{x})$$

である。

次に、 $m = 1, \dots, M$ を学習器の番号としたときに $m$ をランダムに並び替える。すなわち、

$$\{\pi(i)\}_i^N = \text{rand - perm}\{i\}_i^N \quad (15)$$

とする。

次に、 $i = 1, \dots, N'$  ( $N' < N$ ) に対して重み $r_{\pi(i)km}$ を以下の式によって算出する。

$$r_{\pi(i)km} = - \left[ \frac{\partial L(y_{\pi(i)}, F_k(\mathbf{x}_{\pi(i)}))}{\partial F_k(\mathbf{x}_{\pi(i)})} \right]_{F_k(\mathbf{x})=F_{k,m-1}(\mathbf{x})} \quad (16)$$

この $r_{\pi(i)km}$ を目的変数(応答変数)としてB.1で述べた手順に従い分類木を構築し、その終結ノードを $R_{jkm}$ , ( $j = 1, 2, \dots, J_{km}$ )とする。このとき、 $j = 1, 2, \dots, J_{km}$ に対して、

$$\arg \min_{jk} \sum_{\mathbf{x}_{\pi(i)} \in R_{jkm}} L(y_{\pi(i)}, F_k(\mathbf{x}_{\pi(i)}) + \gamma_{jkm}) \quad (17)$$

となるように $\gamma_{jkm}$ をおく。すなわち、

$$\gamma_{jkm} = \arg \min_{\gamma} \sum_{\mathbf{x}_{\pi(i)} \in R_{jkm}} L(y_{\pi(i)}, F_{k,m-1}(\mathbf{x}_{\pi(i)}) + \gamma_{jkm}) \quad (18)$$

を計算する。最後にこの $\gamma_{jkm}$ を用いて、予測関数を

$$F_{km}(\mathbf{x}) = F_{k,m-1}(\mathbf{x}) + \sum_{j=1}^{J_{km}} \gamma_{jkm} \mathbb{1}(\mathbf{x}_{\pi(i)} \in R_{jkm}) \quad (19)$$

のように更新することで木を追加していく。

式(15)~(19)の手順を $k = 1, \dots, K$ と $m = 1, \dots, M$ に対して繰り返すことで、最終的な予測モデルを

$$F_k(\mathbf{x}) = F_{kM}(\mathbf{x}) \quad (20)$$

とする。このとき $m$ 番目の木を構築するに際し、すべての説明変数を使用するのではなく、 $n$ 個の変数のうち一部をランダムに抽出することで過学習を防いでいる(小林・高本 (2010))。

また、Xgboost パッケージでは、最終モデル  $F_{kM}(\mathbf{x})$  を構築する中で一つの木に現れる変数の数を指定できる。そのため、木を逐次的に構築する中で加えられていく重みによって木ごとに現れる変数と現れない変数が分かれる。ある説明変数  $\mathbf{x}_{\pi(i)}$  の重要度は、構築された木全体  $M$  のうちいくつかの木にその説明変数が現れているかで説明ができる。すなわち、ある説明変数  $i$  が最終的に  $H (H \leq M)$  個の木に現れたとすると、説明変数  $i$  の重要度は

$$\frac{H}{M} \quad (21)$$

によって求められる。

以上の過程により、弱学習器を決定木とした Gradient Boosted Decision Tree モデルを構築することができる。

### B.3 Xgboost パッケージで用いたパラメーターの値

関数は `xgb.cv` および `xgb.importance` を用いた。

表5 xgb.cv関数のパラメーター

パラメーター	意味	値
eta	過学習に対するペナルティ	0.4
max_depth	木の深さ(階層数)の最大値	6
subsample	木を作る際ランダムに抽出するデータの割合	0.5
nrounds	木の数	500
nfold	クロスバリデーションの回数	60

## 使用データおよび用例出典

伝康晴・榎本美香(2014) 『千葉大学3人会話コーパス』

## 参考文献

- Breiman, Leo, Jerome Friedman, Charles J Stone, and Richard A Olshen (1984) *Classification and regression trees*: CRC press.
- Chafe, Wallace (1987) “Cognitive constraints on information flow,” *Coherence and grounding in discourse*, Vol. 11, pp. 21-51.
- Clancy, P. M., S. A. Thompson, Ryoko Suzuki, and Hongyin Tao (1996) “The conversational use of reactive tokens in English, Japanese, and Mandarin,” *Journal of Pragmatics*, Vol. 26, pp. 355-387.
- Dittmann, Allen T. and Lynn G Llewellyn (1968) “Relationship between vocalizations and head nods as listener responses,” *Journal of Personality and Social Psychology*, Vol. 9, No. 1, pp. 79 - 84.
- Duncan, Starkey (1974) “On the Structure of Speaker-Auditor Interaction during Speaking Turns,” *Language in Society*, Vol. 3, No. 2, pp. 161-180.
- 榎本美香・石本祐一 (2009) 「「うん」の音響的系譜～応答・承認・相槌の自動抽出に向けて～」, 『研究報告音声言語情報処理 (SLP)』, 第23号, 1-6頁.
- Friedman, Jerome H (2001) “Greedy function approximation: a gradient boosting machine,” *Annals of statistics*, pp. 1189-1232.
- Fries, Charles Carpenter (1952) *The structure of English : an introduction to the construction of English sentences*: Harcourt, Brac.
- 後藤正幸・小林学 (2014) 『入門パターン認識と機械学習』, コロナ社.
- 畠弘巳 (1988) 「外国人のための日本語会話ストラテジーとその教育」, 『日本語学』, 第7巻, 第3号, 110-117頁.
- 林大 (1964) 「ダとナノダ」, 『講座日本語』, 第6巻, 282-289頁.
- 平井有三 (2012) 『はじめてのパターン認識』, 森北出版,152-174頁.
- 堀口純子 (1988) 「コミュニケーションにおける聞き手の言語行動 (聴解の指導)」, 『日本語教育』, 第64号, 13-26頁.
- (1997) 『日本語教育と会話分析』, くろしお出版.
- Iwasaki, Shoichi (1997) 「The Northridge earthquake conversations: The floor structure and the ‘loop’ sequence in Japanese conversation」, *Journal of Pragmatics*, 第28巻, 第6号, 661-693頁.
- Kendon, Adam (1967) “Some functions of gaze-direction in social interaction,” *Acta Psychologica*, Vol. 26, pp. 22-63.
- 小林淳一・高本和明 (2010) 「確率勾配ブースティングを用いたテレコムの契約者行動予測モデルの紹介 (KDD Cup 2009 での分析より)」, 『データマイニングと統計数理研究会 (第 12 回)』.
- 国立国語研究所 (1951) 『現代語の助詞・助動詞：用法と実例』, 国立国語研究所.
- 松田陽子 (1988) 「対話の日本語教育-あいづちに関連して-」, 『日本語学』, 第7巻, 第13号, 59-66頁.

- メイナード泉子 K. (1993) 『会話分析』, くろしお出版.
- Miller, Laura (1991) “Verbal listening behavior in conversations between Japanese and Americans,” *The Pragmatics of International and Intercultural Communication*. Amsterdam: John Benjamins Publishing Company, pp. 111-130.
- 水谷信子 (1988) 「あいづち論」, 『日本語学』, 第12巻, 第7号, 4-11頁.
- 任榮哲・李先敏 (1995) 「あいづち行動における価値観の韓日比較」, 『世界の日本語教育.日本語教育論集』, 第5巻, 239-251頁.
- 大浜るい子 (2006) 『日本語会話におけるターン交替と相づちに関する研究』, 溪水社.
- 大浜るい子・西村史子 (2005) 「日英のターン交替と相づち使用の実相：日本人学生とニュージーランド学生の比較を通して」, 『社会言語科学』, 第7巻, 第2号, 78-87頁.
- 岡崎敏雄 (1987) 「談話の指導—初～中級を中心に」, 『日本語教育』, 第62号, 165-178頁.
- 下川敏雄・杉本知之・後藤昌司 (2013) 『樹木構造接近法』, 共立出版.
- 友定賢治 (2015) 『感動詞の言語学』, ひつじ書房.
- White, Sheida (1986) *Functions of Backchannels in English: A cross-cultural analysis of Americans and Japanese*: University Microfilms.
- Yngve, Victor H (1970) “On getting a word in edgewise,” in *Chicago Linguistics Society, 6th Meeting*, pp. 567-578.
- 楊晶 (2002) 「調査報告日本語の相づちに関する意識における中国人学習者と日本人との比較」, 『日本語教育』, 第114号, 90-99頁.

## 辞典

『日本国語大辞典』(2002), 第二版, 小学館.