

# 自律型アンドロイドによる対話における同調的笑いの生成

勝見 久央<sup>1</sup> 井上 昂治<sup>2</sup> 中村 静<sup>2</sup> 高梨 克也<sup>2</sup> 河原 達也<sup>2</sup>

**概要：**人間に近い外見を備えた自律型アンドロイドによる、対話の聞き手応答として、同調的笑いの生成を試みる。同調的笑いとは、対話相手の笑いに誘われて笑うことを指し、人間どうしの対話では、しばしば観察される。このような同調的笑いは、話を聞いていることを相手に示したり、対話の場を和ませたりなど、対話を円滑に進める上で重要な役割を果たすと考えられる。本研究では、はじめに、オペレータにより遠隔操作されたアンドロイド ERICA と、被験者との初対面対話データを用いて、同調的笑いの生起および関係する特徴を分析した。分析に基づき、韻律的、意味論的、相互行為的という観点から特徴量を設計した。これらを用いて、被験者の笑いの直後に、ERICA が同調的笑いをを行うか否かの予測を、ロジスティック回帰によりモデル化した。実験結果より、提案手法はチャンスレートを上回る予測精度を達成した。また、韻律的特徴量のうちの発話との共起に基づく分類、相互行為的特徴量のうちの直前発話末尾の品詞と表層形、これらが予測精度の向上に寄与することがわかった。

**キーワード：**自律型アンドロイド、音声対話システム、同調的笑い

## 1. はじめに

従来の音声対話システムは、飛行機やバスの検索システム、街の案内システムなどに代表されるように、データベース検索による情報検索のインターフェースとしてのタスク遂行型のもが主流であった。一方で、近年、このようなタスク遂行型システムでない柔軟な対話システムの研究が行われている。さらに近年、人間の外見をした自律型アンドロイドの研究が盛んになりつつある。このようなアンドロイドは、従来の音声対話システムと異なり、人間に似た外見やふるまい、表情を用いて非言語情報を表現でき、研究室案内や傾聴対話といったタスクが検討されている。自律型アンドロイド ERICA[1] もその一つである。

これまでに、人間どうしの対話をモデル化して自律型アンドロイドに取り入れる様々な試みがなされている。具体的には、システムが人間の会話に相槌を打ったり、相手の発話を聞いて内容に応じた適切な質問を返すことによって対話相手の発話を促す研究が行われている。これらの相槌や質問は、対話をより自然なものにするために欠かせないものであり、また、相手の発話をより引き出す役割がある。

同様に、人間どうしの対話において笑いも重要な要素の一つである。特に、笑いの一種である同調的笑いは、対話において、ある対話参加者が笑うと、他の参加者もそれに誘わ

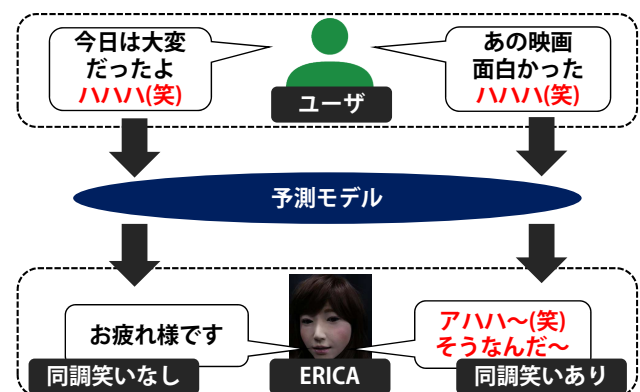


図 1 システム概要図

れて一緒に笑うものであり、相手に対して友好的な態度を示したり、会話の場を和ませて円滑にしたりといった、さまざまな対話上の効果を持っている。そこで、本研究ではアンドロイドにおいて適切なタイミングで同調的笑いを生成することを目指す。この同調的笑いがアンドロイドによって生成できれば、より自然な対話を実現でき、対話の場を和ませたり、ユーザのシステムへの親近感を強められると期待できる。自律型アンドロイドにおいて同調的笑いを生成するシステムの概要を図 1 に示す。システムは対話相手であるユーザの笑いを特徴ベクトル化したものを入力し、それに同調してシステムが笑うべきかどうかの 2 値を出力する。

本稿の構成を以下に示す。2 章では、はじめに本研究で使用するために収録した ERICA と被験者との対話データに

<sup>1</sup> 京都大学 工学部情報学科

<sup>2</sup> 京都大学 大学院情報学研究所

ついて述べ、次に、そのデータを対象に行なったアノテーションや分析、予測の際に使用する特徴量の抽出について述べる。3章では、2章の分析を踏まえてロジスティック回帰モデルによる同調的笑いの予測を行った結果について述べる。

## 2. データ分析とアノテーション

### 2.1 対話収録

本研究では2016年9月下旬から10月初旬に大阪大学において行われた、ERICA[1]と被験者1名の約10分間の初対面対話46対話のうち、正常に収録された39対話を分析と予測に用いる。これは遠隔対話であり、別室でERICA役を担う劇団員などの女性オペレータが対話相手の音声とのカメラの映像に基づいて、自身の声で言語的に応答するとともに、上半身の簡単な動作や視線などの非言語行動をコントローラによって操作することによって対話を行っている。自然な初対面対話をシュミレートするため、対話の状況としては研究室の秘書役のERICAが来訪者に応対する場合を設定した。オペレータにはERICAが教授の秘書であるものとして振る舞ってもらい、来訪者役の被験者には対話の収録前に、ワークシートへの記入を通じて来訪目的や社会的属性に関する設定を行い、研究室見学にきた学生や営業にきた業者といった来訪者の役になりきってもらった。やりとりされる来訪目的に関する質問応答や教授が戻ってくるまでの時間つぶしのための雑談などの10分程度の自由な対話をマイクとカメラ、Kinectによって収録した。収録した対話データに対して以下のような情報のアノテーションを行った。

### 2.2 笑いと同調的笑い

人手で作成された発話内容の書き起こしには笑っていると知覚できる区間を判定した「笑い」のタグも含まれている。「笑い」のタグは次の2種類である。

- (L こんにちは L)  
発話と笑い声が混ざって聞こえる笑い。笑いながら発された発話を括弧で囲んで表記。
- {LAUGH}  
発話を伴わない笑い声のみの笑い。

ただし、2秒以内に連続して知覚された笑いの箇所はまとめて1つの笑いとした。また、同調的笑いの定義については、ある笑いの発生からその直後4秒以内にもう一方の会話参加者による笑いが発生したとき、その後続の笑いを「同調的笑い」とした。分析対象とした全39対話中では被験者の「笑い」は合計269回生起していた。これらについてさらにタグを付与して細分した。タグ付けはGuptaら[2]を参考にし、観測された被験者の笑いに対して次の3種類のタグ付けを行った。

- stand-alone subject's laughter: St-S

表 1 全39対話中の3種類の被験者笑いの回数

種類	St-S	Sh-S	Sh-E	計
回数	186	63	20	269

被験者の笑いのうち、その直後4秒以内にERICAの同調的笑いが発生しなかったもの

- subject-lead shared laughter: Sh-S  
被験者の笑いのうち、その直後4秒以内にERICAの同調的笑いが発生したもの
- ERICA-lead shared laughter: Sh-E  
被験者の笑いのうち、その直前4秒以内にERICAの笑いが発生したもの

この分類結果を表1に示す。また、この3種類のうち、St-SとSh-Sは被験者がERICAに同調したのではなく主体的に発した笑いであるため、被験者主導の笑いといえる。この被験者主導の笑いは92.6%  $(=(186+63)/269)$ と被験者の笑いのうちほとんどを占めている。さらに、Sh-Eは同調的笑いであるが、先にERICAが笑い、それに被験者が後から同調したものであるため、被験者の笑いにERICAが同調すべきかを予測するという本研究の目的から外れる。したがって、以降ではさらにこれら計249  $(=186+63)$ 回のSt-SとSh-S(以下、被験者主導の笑い)から様々な特徴量を抽出し、機械学習によって両者の識別を行うことを通じて、同調的笑いの生起予測を行う。

### 2.3 基本的な韻律的特徴量

第2章で述べたように、Guptaら[3]は笑いがpositive, neutral, negativeのいずれかであるかを予測するための韻律的特徴として基本周波数とパワーを使用している。また、山口ら[4]も同様に、相槌の生起予測として相槌の直前発話の基本周波数とパワーを特徴量として使用している。このように韻律的特徴は音声による自然なコミュニケーションを分析する上で重要なものである[5]。また、これらの韻律的特徴はリアルタイムに取得しやすいものである。そこで、本研究でもまず基本的な韻律的特徴を特徴量として使用する。ただし、各セッションの被験者間で個人差が見られるため、個人差の影響を受けにくい対数基本周波数と対数パワーのレンジを使用した。

### 2.4 笑いに固有の韻律的特徴量

前述の2つの特徴量は対話における韻律的特徴として一般的なものであるが、笑いに特化したものではない。これに対して、本研究では、笑いを分析対象としていることから、笑いに固有の以下の韻律的特徴も用いる。

#### 2.4.1 発話との共起

観測された全ての被験者主導の笑いをTianらの研究に基づき次のような2つに分類した。  
発話との共起に基づく笑いの分類

- speech laughter:  
言語的な発話をしながら笑っているように聞こえるもの

• stand-alone laughter:  
言語的な発話を伴わない単独の笑いに聞こえるもの  
両者の分類は基本的には2.2で述べた2種類のタグを使用した  
したが、同一の笑い区間に2種類のタグが混在するものは、  
前者の speech laughter に分類した。この2種類の分類と  
前節で述べた直後の同調的笑いの分類との関係は表2のよ  
うになる。これから、speech laughter の32.5% (=50/156)  
で直後に同調的笑いが生起しているのに対し、stand-alone  
laughter は14.0% (=13/93)しか直後に同調的笑いが生起  
していない。したがって、speech laughter はstand-alone  
laughter よりも同調的笑いを誘発しやすい傾向にあること  
がわかる。

### 2.4.2 発声法

笑いの中には小さく「フッフ」と笑うような愛想笑いや  
かすかな笑いも多いが、このような笑いが相手の同調的笑  
いを誘うことはあまりない。一方で、面白い話を聞いたり  
したときの「ハハハ」というような笑いは同調的笑いを誘  
いやすい。そこで、このような笑いの聞こえ方が同調的笑  
いの生起頻度に影響を与えている可能性を考慮し、大原 [8]  
を参考にし、笑いをその発声法によって次の2種類に分類  
した。

#### 発声法に基づく笑いの分類

- obvious laughter:  
「アハハハ」、「フッフ」というように声の繰り返し  
構造、強い震えが聞き取れる笑い
- breath laughter:  
obvious laughter には分類されない笑いであり、息を  
吸う音、息を吐く音、鼻笑音で笑い声の大部分が構成さ  
れているような笑い。大原の研究における「引き笑い  
声」、「吸気音」、「鼻笑呼気音」

この2種類の分類と直後の同調的笑いの生起の関係は、表  
3のようになる。obvious laughter の30.7% (=27/88)で直  
後に同調的笑いが生起しているのに対し、breath laughter  
は22.4% (=36/161)しか直後に同調的笑いが生起してい  
ない。したがって、obvious laughter はbreath laughter より  
も同調的笑いを誘発しやすい傾向にある。

また、発話との共起に基づいた分類と笑い声の発声法  
に基づいた分類を表4に示す。ここから、全ての被験者  
主導の笑いのうち、全般的にbreath laughter かつ speech  
laughter である笑いが多い。さらに、obvious laughter かつ  
speech laughter である笑いの43.2% (=19/44)で直後  
に同調的笑いが発生しており、8種類の笑いのうちで最も直  
後に同調的笑いが発生している。一方で、大部分を占める  
breath laughter かつ speech laughter である笑いのうちで  
は、26.8% (=30/112)で直後に同調的笑いが発生している。

表2 発話との共起に基づいて分類した被験者主導の笑いの回数

	St-S	Sh-S	計
speech laughter	106	<b>50</b>	156
stand-alone laughter	80	<b>13</b>	93
計	186	63	249

表3 笑い声の発声法に基づいて分類した被験者主導の笑いの回数

	St-S	Sh-S	計
obvious laughter	61	<b>27</b>	88
breath laughter	125	<b>36</b>	161
計	186	63	249

表4 笑い声の発声法による分類と発話との共起に基づく分類

笑い声の発声法	発話との共起	St-S	Sh-S	計
obvious laughter	speech	<b>25</b>	<b>19</b>	<b>44</b>
	stand-alone	36	8	44
breath laughter	speech	<b>82</b>	<b>30</b>	<b>112</b>
	stand-alone	43	6	49
計		186	63	249

### 2.5 意味論的特徴量

日常的に頻繁にみられる同調的笑いは、会話中に先に一  
方が笑い始め、それに誘われてもう一方が笑うものである  
が、必ずしも常に相手が笑えば、それに同調して笑うことが  
適切なわけではない。例えば、

話者A: 「Bさん、運動は得意ですか?」

話者B: 「運動ですか。まあ、見たと通り得意じゃない  
ですね(笑)。」

話者A: 「そうなんですか。」

といったやり取りのように、相手が自虐的な発話などのネ  
ガティブな内容と共に笑った場合、それに同調して笑って  
しまうと無礼な印象を相手に与えてしまうことがある。こ  
れは自己卑下に対しては不同意が選好的である [9] が、笑い  
は同意を表していると解釈されやすいためであり、特に初  
対面での対話の場合には、相手の発話内容により慎重に配  
慮して同調的笑いを生成する必要がある。こういった点を  
考慮するため、Gupta ら [3] を参考にし、観測された被験者  
の笑いを次の3種類に分類し、タグ付けを行った。

#### 極性に基づく笑いの分類

- positive laughter  
ポジティブな発話の直後に発生した笑い
- negative laughter  
ネガティブな発話の直後に発生した笑い
- neutral laughter  
上記の2種類のどちらにも当てはまらない、両者の中  
間的笑い

この分類は、笑い自体ではなく、その直前の言語的な発話の

表 5 極性に基づいて分類した被験者主導の笑いの回数

	St-S	Sh-S	計
positive laughter	70	18	88
neutral laughter	77	30	107
negative laughter	39	15	54
計	186	63	249

意味内容に基づくものであり、次の手順で行っている。

(1) 対象の笑いの直前の発話を両参与者のものを合わせて5発話まで遡り、それぞれに positive、neutral、negative いずれかの極性を付与する。(なお、発話単位としては 200msec 以上のポーズで区切られたの間休止単位 (IPU) を用いる。これ以降についても同様)

(2) 極性から笑いの直前の対話の極性を次のように決定する。

- 5 発話の極性が neutral の 1 種類のみからなる場合、その極性を neutral とする。
- 上記にあてはまらない場合、neutral な発話を取り除き、時間的に最も笑いの発生箇所に近い発話の極性を、この箇所の極性とする。

各発話の極性は、MeCab[10] と CaboCha[11][12] を用いて日本語極性辞書 [13][14] にエントリーが存在する名詞と用言を抽出し、それらの極性のスコア平均とした。被験者主導の笑いの極性と直後の同調的笑いの生起の関係を表 5 に示す。positive laughter のうち 20.5% (=18/88), neutral laughter のうち 27.1% (=29/107), negative laughter のうち 27.8% (=15/54) で直後に同調的笑いが生起している。

## 2.6 相互行為的特徴量

### 2.6.1 談話行為

同調的笑いは聞き手反応の一種であるため、対話における聞き手反応に関する特徴量として談話行為 (以下、DA) を用いる。具体的には、全ての発話単位 (LUU[15]) に対して [16] で以下の 4 種類の DA タグを人手で付与し、このうち被験者主導の笑いが発生した直前の被験者の発話の DA タグと直前の ERICA の発話の DA タグを特徴量として用いる。

使用する談話行為タグとその定義

- Question (Q タグ)  
情報要求機能を持つ発話
- Statement (S タグ)  
情報提供機能、行為交渉機能、行為拘束機能を持つ発話や指図など
- Response (R タグ)  
Q と S への応答
- Other (O タグ)  
挨拶、御礼などのその他の発話

### 2.6.2 文末表現

文末に出現する助動詞や終助詞は、対話相手への働きかけなどを表す表現であるため、聞き手応答を分析する際にも有用であると考えられる。そこで、本研究ではこれらに關しても特徴量として使用する。具体的には次の 4 つを特徴量として用いる。

文末表現に関する特徴量

- 被験者主導の笑いの直前の被験者の発話の末尾の品詞
- 被験者主導の笑いの直前の ERICA の発話の末尾の品詞

ただし、末尾の品詞が助動詞もしくは終助詞であるときは、末尾の品詞名に加えて末尾から 3 つの形態素の表層形も特徴量として用いる。

### 2.6.3 オーバーラップの有無

対話相手への働きかけは、その発話のタイミングにも表れる。そこで、被験者の笑いを直前の ERICA の発話とのタイミングに基づき、次の 2 種類に分類する。

笑いの直前発話へのオーバーラップ区間の有無

- オーバーラップあり  
被験者の笑いとその直前の ERICA の発話が重なるとき
- オーバーラップなし  
被験者の笑いとその直前の ERICA の発話が重ならないとき

## 3. 同調的笑いの予測実験

### 3.1 実験条件

本実験の目的は、収録データで観測された被験者の笑いを特徴ベクトル化したものから、その直後に ERICA の同調的笑いが発生したかどうかを予測することである。使用したデータは 2.1 節で述べた初対面対話 39 セッションで、この中には、被験者主導の笑いは全部で 249 回含まれており、そのうち 63 回では直後に ERICA の同調的笑いが発生していた (表 1)。予測にはロジスティック回帰モデルを使用し、249 個の被験者の笑いの特徴ベクトルうち 80% の 199 個をランダムに選んでモデルの学習のための訓練データとし、残りの 50 個をテストデータとした。また、ロジスティック回帰では笑いの入力ベクトルそれぞれに対し、同調的笑いが生起する事後確率と生起しないと予測する事後確率の 2 つが算出される。このとき本実験では、

$$threshold = \left( \frac{\text{訓練データのうち同調的笑いを生じた回数}}{\text{訓練データの総数 (=199)}} \right)$$

として出力の閾値  $threshold$  を定め、予測モデルの出力を

$$\begin{cases} \text{同調的笑いが生起する : } & (threshold \leq (\text{同調的笑いの事後確率})) \\ \text{同調的笑いが生起しない : } & (threshold > (\text{同調的笑いの事後確率})) \end{cases}$$

とする。

表 6 使用した特徴量

A) 韻律的特徴量	a) 基本的な韻律的特徴量	1. 対数基本周波数レンジ 2. 対数パワーレンジ 3. 笑いの持続時間
	b) 笑いに固有の韻律的特徴量	4. 発話との共起に基づく分類タグ (speech laughter / stand-alone laughter) 5. 笑い声の発声法に基づく分類タグ (breath laughter / obvious laughter)
B) 意味論的特徴量		6 極性に基づく分類タグ (positive laughter / neutral laughter / negative laughter)
C) 相互行為的特徴量		7. 直前発話の DA タグ (Q / S / R / O) 8. 直前発話の文末の品詞, 表層形 9. 直前発話へのオーバーラップ

被験者の笑いから抽出した特徴量には第 3 章で述べた特徴量を使用した。(表 6)

### 3.2 予測結果

予測実験の結果は表 7, 8, 9 のようになった。表 7 に示すチャンスレートは、全ての箇所において同調的笑いが生起するものと予測するモデルである。ただし、これらの数値は各行に応じた特徴量を予測モデルに入力し、F 値、適合率、再現率を求めるという実験を 1000 回繰り返して行い、それらの平均値を記載したものである。したがって、各行の F 値は同じ行の適合率と再現率から求まる F 値とは一致しない。

まず表 7 より、本研究の予測モデルはチャンスレートを上回っていることが確認できる。同調的笑いは、間違っただけで生成してしまうと、相手からの信頼感を損なったり、不信感を与えたりするリスクが大きく、適合率も重要である。

次に表 8 に示すように、各特徴量がどれだけ予測精度の向上に寄与しているかを特定するため、最も基本的かつ容易に抽出が可能である対数 F0 レンジ、対数パワーレンジ、笑いの持続時間からなる基本的な韻律的特徴 (Aa) のみを用いたモデルを基準とし、これに他の 3 種類の特徴量を加えた予測モデルと精度を比較している。

その結果、韻律的特徴量と笑いに固有の韻律的特徴量を併用した予測モデルと韻律的特徴量と相互行為的特徴量を加えたモデルにおいて予測精度が向上することがわかった。

最後に、表 9 のように、全ての特徴量を使用した予測モデルの精度を基準とし、構成特徴量を一つずつ取り除いた場合との精度の比較を行った。すると、Ab) 笑いに固有の韻律的特徴量のうち特に特徴量 4 の発話とその共起による笑いの分類タグ (speech laughter と stand-alone laughter) を使用しなかったモデルでは、全ての特徴量を使用したモ

表 7 全特徴量による予測結果とチャンスレートとの比較

	適合率	再現率	F 値
チャンスレート	<b>0.2519</b>	1	<b>0.3993</b>
全特徴量を使用した予測モデル	<b>0.3310</b>	0.5691	<b>0.4113</b>

表 8 使用した特徴量と予測精度

	適合率	再現率	F 値
Aa) 基本的な韻律的特徴量のみを使用	0.2784	0.4703	0.3425
Aa) 基本的な韻律的特徴量と Ab) 笑いに固有の韻律的特徴量を使用	<b>0.3186</b>	<b>0.5988</b>	<b>0.4078</b>
Aa) 基本的な韻律的特徴量と B) 意味論的特徴量を使用	0.2800	0.4644	0.3424
Aa) 基本的な韻律的特徴量と C) 相互行為的特徴量を使用	<b>0.3014</b>	<b>0.5154</b>	<b>0.3743</b>
全ての特徴量を使用	0.3310	0.5691	0.4113

デルと比べて予測精度が大幅に低下した。これは 2.4.1 節でも示したように、同調的笑いが speech laughter の直後の方が stand-alone laughter の直後に比べて生起しやすいことによる。また、C) 相互行為的特徴量のうち、8 の直前発話末尾の品詞と表層形を使用しなかったモデルも、若干の予測精度の低下が観察された。このことから、先行笑いと言話の共起や、笑いの直前の発話の末尾の言語的特徴が同調的笑いの発生の有無に影響を及ぼしているといえる。

## 4. おわりに

本研究では、自律型アンドロイドによる同調的笑いを生成するため、対話コーパスを収録し、それをを用いて同調的笑いの生起の予測を行った。予測にはロジスティック回帰モデルを用い、入力特徴量として、笑い声の韻律的特徴量、意味論的特徴量、相互行為的特徴量の 3 種類のを数個ずつ用いて、予測実験を行った。韻律的特徴量としては抽出が容易な笑いの基本周波数、パワー、持続時間だけでなく、笑い固有の韻律的特徴として直前発話との共起による分類結果や発声法による分類結果を加えた。また、意味論的特徴量として直前発話の positive や negative などの極性を用い、相互行為的特徴量としては談話行為タグと直前発話の末尾語の品詞と表層形、直前発話へのオーバーラップの有無を用いた。予測実験の結果、同調的笑いの予測には、笑いが発話とともに発されたものか、単独で発されたものかどうかを分類した、発話との共起による分類が有用な特徴量であり、直前発話の末尾語の品詞や表層形についても有用であることがわかった。今後は予測モデルを改良したり、音声情報以外から得られるマルチモーダルな特徴量についても検討する必要がある。

謝辞 本研究は、JST ERATO 石黒共生ヒューマンロボットインタラクションプロジェクト (課題番号: JPM-JER1401) の一環で行われた。

表 9 全ての特徴量を使用した場合と各特徴量を使用しなかった場合の比較

		適合率	再現率	F 値
全ての特徴量を使用したとき		0.331	0.5691	0.4413
使用しなかった特徴量		適合率	再現率	F 値
Aa) 基本的な 韻律的特徴量	1. 対数 F0 レンジ	0.3332	0.5700	0.4131
	2. 対数パワーレンジ	0.3298	0.5762	0.4128
	3. 笑いの持続時間	0.3372	0.5885	0.4220
Ab) 笑いに固有の 韻律的特徴量	4. 発話との共起に基づく 分類タグ	<b>0.2899</b>	<b>0.4881</b>	<b>0.3570</b>
	5. 笑い声の発声法に基づく 分類タグ	0.3164	0.5770	0.4023
B) 意味論的特徴量	6. 極性に基づく笑いの 分類タグ	0.3281	0.5780	0.4116
C) 相互行為的特徴量	7. DA タグ	0.3301	0.5800	0.4141
	8. 直前発話末尾の品詞, 表層形	<b>0.3056</b>	<b>0.5757</b>	<b>0.3926</b>
	9. 直前発話へのオーバーラップ の有無	0.3364	0.5678	0.4159

参考文献

- [1] 井上昂治: 自律型アンドロイド Erica のための音声対話システム, *SIG-SLUD*, Vol. 75, pp. 21–24 (2015).
- [2] Gupta, R., Chaspari, T., Georgiou, P., Atkins, D. and Narayanan, S.: Analysis and Modeling of the Role of Laughter in Motivational Interviewing Based Psychotherapy Conversations, *INTERSPEECH* (2015).
- [3] Gupta, R., Nath, H., Agrawal, T., Georgiou, P., Atkins, D. and Shrikanth, N.: Laughter Valence Prediction in Motivational Interviewing based on Lexical and Acoustic Cues, *INTERSPEECH* (2016).
- [4] 山口貴史, 井上昂治, 吉野幸一郎, 高梨克也, Ward, N., 河原達也: 傾聴対話システムのための言語情報と韻律情報に基づく多様な形態の相槌の生成., *人工知能学会論文誌*, Vol. 31, No. 4, pp. C–G31.1 (2016).
- [5] 市川熹: 対話のことばの科学-プロソディが支えるコミュニケーション, *日本語の研究*, Vol. 8, No. 2, p. 99 (オンライン), 入手先 (<http://ci.nii.ac.jp/naid/110009517560/en/>) (2011).
- [6] Kawahara, H., Masuda-Katsuse, I. and De Cheveigne, A.: Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds, *Speech communication*, Vol. 27, No. 3, pp. 187–207 (1999).
- [7] Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T. and Banno, H.: TANDEM-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation, *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, IEEE, pp. 3933–3936 (2008).
- [8] 大原察: 対話音声の笑い声や笑い方についての分析, 奈良先端科学技術大学院大学修士論文 (2005).
- [9] Pomerantz, A.: Agreeing and Disagreeing with Assessments: Some Features of Preferred/Dispreferred Turn Shapes, *Structures of Social Action: Studies in Conversation Analysis*, pp. 57–101 (online), available from (<http://ci.nii.ac.jp/naid/10009701891/en/>) (1984).
- [10] Kudo, T., Yamamoto, K. and Matsumoto, Y.: Applying Conditional Random Fields to Japanese Morphological Analysis., *EMNLP*, Vol. 4, pp. 230–237 (2004).
- [11] 工藤拓, 松本裕治: チャンキングの段階適用による日本語係り受け解析, Vol. 43, No. 6, pp. 1834–1842 (2002).
- [12] Taku Kudo, Y. M.: Japanese Dependency Analysis using Cascaded Chunking, *CoNLL 2002: Proceedings of the 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)*, pp. 63–69 (2002).
- [13] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: 意見抽出のための評価表現の収集, *自然言語処理*, Vol. 12, No. 3, pp. 203–222 (2005).
- [14] 東山昌彦, 乾健太郎, 松本裕治: 述語の選択選好性に着目した名詞評価極性の獲得, *言語処理学会第 14 回年次大会論文集*, pp. 584–587 (2008).
- [15] Japanese Discourse Research Initiative: 発話単位ラベリングマニュアル, Vol. 2.
- [16] Bunt, H., Alexandersson, J., Carletta, J., Choe, J.-W., Fang, A. C., Hasida, K., Lee, K., Petukhova, V., Popescu-Belis, A., Romary, L. et al.: Towards an ISO standard for dialogue act annotation, *Seventh conference on International Language Resources and Evaluation (LREC'10)* (2010).