

## 自然対話における聞き手の反応と話し手のうなずき・ 言語情報・韻律情報との関係に関する予備的検討

堀内 靖雄† 庵原 彩子†  
西田 昌史† 市川 薫†

自然対話でのマンマシンインタフェースの実現を考えたとき、人間のように自然な対話を実現するためには、音声言語情報ばかりでなく、ジェスチャーなどのマルチモーダルな情報を利用することが有効であると考えられる。本研究では日本語の自然対話をマルチモーダルな環境で収録し、人間同士の自然対話におけるジェスチャーの分析を行なった。日本語の自然対話ではとくにうなずきが多用される。収録した対話を分析した結果、うなずきにはあいづち同様の機能だけでなく、話し手の発話末でうなずきが生じるケースが非常に多いことが示され、また、発話末のうなずきに対し、聞き手がうなずきやあいづちなどの反応を返す現象が数多くみられた。一方、過去の研究から発話末の言語情報、韻律情報などが聞き手の反応と関係があることが示唆されていることを踏まえ、発話末の話し手のうなずき・言語情報・韻律的特徴に対し、聞き手の反応がどのような傾向を示すか分析した。結果として、発話末のうなずきは言語情報や韻律情報以上に聞き手の反応をうなずき機能があることが示唆された。

### Analysis of Speaker's Nod, Word, Prosody and Listener's Reaction in Japanese Spontaneous Speech

YASUO HORIUCHI,† SAIKO IHARA,† MASAFUMI NISHIDA†  
and AKIRA ICHIKAWA†

People use gestures like nods for smooth communication in spontaneous speech dialogue. Usual dialogue continues exchanging interlocutor's information with each other using speech and gestures and therefore it is supposed that there is correlation between speakers' gestures and listeners' reaction. In this paper, we focused the relationship between the speaker's nods, syntactic features and prosodic features in the final part of each utterance and the listener's reaction like nods or backchannels. 4 dialogues by four pairs of good friends were recorded, where they can look at each other via two prompters. The prompter can record the interlocutor's gesture on videotape and project the partner's image through a half mirror. As a result, it was suggested that speakers' nod is caused frequently in the final part of utterances and that listeners show reaction like nods or backchannels frequently when the speaker nodded, said some typical words, or spoke with the typical intonation in the final part of utterance.

#### 1. はじめに

近年の計算機能力の急速な発展により、人間と対話するコンピュータシステムの研究が盛んになってきている。ユーザフレンドリーなマンマシンインタフェースを目指すためには、人間らしい受け答えが可能となることが重要である。そのため人間同士が行なっている自然対話における現象を分析し、システム上に実装する必要がある。しかしながら、文献<sup>1)</sup>などで指摘されているように、とくに日本語の場合、話し手と聞き手がお互いに密接に協調しながら、対話を行なっていることが示唆されている。たとえば、日常の会話では、聞き手は話し手の発話を聞きながら、適切なタイミングであいづちやうなずきを返したりしている。このような聞き手のあいづちやうなずきなどの反応がないと話し手は相手が自分の発話を理解しているのか不安に

なってしまう。

このようにあいづちなどの現象は自然対話によるマンマシンインタフェースを考える上で非常に重要である。日本語のあいづちに関しては古くからさまざまな研究が行なわれており、興味深い結果が得られている<sup>1)~9)</sup>。また、分析して得られた知見を用いて、音声対話システムが適切なタイミングであいづち発話を生成する研究も数多く行なわれている<sup>10)~17)</sup>。

さらに、あいづちなどの音声現象だけでなく、うなずきや視線などの非言語情報が対話でどのような機能を有しているのかを調べるため、マルチモーダルな状況における対話収録が行なわれ<sup>18)~22)</sup>、うなずきなどのジェスチャーの研究も盛んになってきている<sup>1),23)~31)</sup>。うなずきはあいづち同様、日本語の自然対話において頻出する現象であり、人間とのマンマシンインタフェースを考えると非常に重要なものとなる。

これまでのあいづちに関する研究では、おもに言語的な情報や韻律的な情報により、どのようなタイミングであい

† 千葉大学  
Chiba University

づちがうたれるのか、というようなことが中心に分析されてきている。一方、うなずきに関する研究ではうなずきと話者交替の関係などが論じられている。このように、過去の多くの研究では、あいづちやうなずきが別個に扱われていたが、うなずきもあいづちの機能を有することがあり、文献 1) ではうなずきもあいづちとして分析されている。そこで本研究では、日本語の自然対話において、うなずきがどのような状況で発現し、それぞれがどのような機能を持っているかを分析し、これまであいづち研究で言われてきた言語的な情報や韻律的な情報との関係も明らかにすることを目的とする。

## 2. マルチモーダル対話収録

我々はうなずきなどのジェスチャーを含めたマルチモーダルな自然対話の収録を行なってきた<sup>22)</sup>。マルチモーダルな状況での自然対話の収録に際して、両者の分離したクリアな音声を録音するとともに、正面からの顔および上半身の画像収録を行なうため、独立した二つの防音室において、相手のプロンプタを介した対話の収録を行なった。

各対話者には別々の独立した防音室内に入って対話を行なってもらうことにより、二人の話者の音声を分離して収録することが可能となった。対話者同士の音声のやり取りはイヤホンとピンマイクを用いた。

アイコンタクトやうなずきといった顔の情報を記録するため、それぞれの被験者の正面にプロンプタを配置した。プロンプタは下部に設置された液晶ディスプレイから、相手話者の画像をハーフミラーを介して呈示すると同時に、ハーフミラーの背面から話者の正面画像を二台のカメラで収録することが可能である (図 1 参照)。

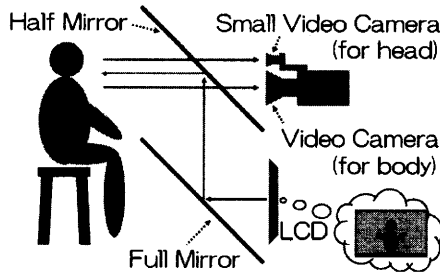


図 1 プロンプタ

プロンプタのハーフミラーの背面に設置された下側のカメラで上半身全体の映像、上側の小型カメラで顔の映像を収録することができる。相手話者には、下側のカメラで撮影された上半身全体の映像が呈示される。記録した映像の例を図 2 に示す。

実験参加者は、日頃から仲のよい友人同士 (大学生) で、男性同士のペア 2 組、女性同士のペア 2 組の計 4 組である。

本研究では活発な会話を自然に行なえるようにするため、ディベート形式の対話を収録することとした。対話収録に先立ち、アンケート調査を実施し、被験者二人に意見が分



図 2 対話収録画像例

かれそうな二者択一の質問にいくつか答えてもらい、二人の意見が分かれたテーマに関して、自由に討論するよう指示した<sup>\*</sup>。途中で議論が収束するか、あるいは、約 5 分経過した段階で実験者が対話を打ち切り、対話の終了とした。

このように別室におけるプロンプタを介した対話という特殊な収録環境により、対話における自然さがそこなわれる可能性があるが、実際に収録を行なってみたところ、大学生の実験参加者は、そのような環境にもすぐ適応し、比較的自然な会話が収録された<sup>\*\*</sup>。

## 3. 分析

### 3.1 分析データ

本研究では、上記収録により得られた 4 対話中から約 13 分の対話データを分析対象とした。分析には上述の図 2 のような四画面が合成された音声付動画像を用いた。

画面上方に呈示されている二台のカメラ映像は顔の拡大画像が映されており、視線やうなずきの分析が容易に行なえるようになっている。また、手や体のジェスチャーなどは下側の二台のカメラ映像を用いて分析が可能である。

うなずきなどのジェスチャーのアノテーションは動画像アノテーションツール ANVIL<sup>19)</sup> を用いて行なった。また、音声情報の書き起こしは ANVIL では困難であるため、千葉大学で開発された書き起こし支援ツール<sup>32)</sup> を用いて行ない、ANVIL のデータ形式に変換した。

### 3.2 発話単位

自然発話には句点や読点などの明確な区切りがなく、発話の単位を決定するのは難しい。そこで本研究では 200 ミリ秒以上の無声区間で分割された音声区間を分析単位 IPU (Interpausal Unit)<sup>8),32)</sup> とした。

また、自然発話には笑いが多く表われるが、本研究では笑いは分析対象外とした。しかしながら、笑いはあいづちなどと似た機能がある、という指摘もあるため、今後の検討課題とした。

### 3.3 うなずき

本研究で対象とするジェスチャーはうなずきである。う

<sup>\*</sup> 質問の例としては、「男と女どっちが得だと思いますか?」「神様の存在を信じますか?」「機械が知能を持つことに賛成ですか?」「男女間に友情は成り立つと思いますか?」など。

<sup>\*\*</sup> 文献 24) でも同様の傾向が示されている。

うなずきとは顔の縦方向の運動で典型的には下降上昇する動きとした。また、うなずきが一連の動作として、連続して複数回おこなわれた場合にはそれらを一つのうなずきとして分析を行なった。うなずきのタグ付けは二人の熟練した作業者が共同で行ない、不明な点については両者相談の上、決定した。

### 3.4 分析1：うなずきはいつ起こるのか

まずはじめに日本語対話におけるうなずきがいづつ起こるのか検討してみたい。うなずきが発話手、聞き手のどのような状況で発現するのかを調査した結果を表1に示す。ここで、発話直後とはそれぞれ話し手が発話の直後にうなずいた場合、聞き手が相手の発話の直後にうなずいた場合を表わしている。また、無音区間中とは、それぞれ、話し手が発話終了後しばらくたってから無音区間でうなずいた場合と聞き手が相手の発話終了後しばらくたってからうなずいた場合を表わしている。なお、表中、重複発話とは二人が同時に話している状況であり、うなずきが各話者の発話とどのような関係があるのかが不明なため、分析対象外とした。

表1 話し手/聞き手の状況とうなずきの出現頻度 (表中の数値は出現回数。太字は状況ごとの合計。)

	話し手	聞き手
発話途中	173	50
発話末	196	69
あいづち発話と同時	—	79
発話中のうなずき	<b>369</b>	<b>198</b>
発話直後	8	15
無音区間中	12	8
無音中のうなずき	<b>43</b>	
重複発話中のうなずき	<b>47</b>	
うなずきの合計	<b>657</b>	

この表からほとんどのうなずきはどちらかの発話中に行なわれていることがわかる。また、一般的には、うなずきはあいづちなどと同様、聞き手のときに表出されることが多いと考えられがちであるが、この表から、話し手のうなずきの方が多いことが示唆される。さらに、話し手のうなずきの中でも、話し手の発話末のうなずき頻度が高いことがうかがえる。何故なら、もし話し手が発話中、無作為にうなずきを行なっているとすると、発話末よりも発話途中の方がうなずきは多く発生すると考えられるが、それに比べると表1の発話末のうなずきの頻度は高い。すなわち、話し手は発話末にピンポイントでうなずきを発生させていることが示唆される。この発話末のうなずきはメイナードによるうなずきのカテゴリー分類<sup>1)</sup>では3および4に相当する。文献1)ではこのようなうなずきは発話末を明確にしたり、発話順番の終了を示すのではないかと、というような考察が与えられているが、本研究では、この発話末のうなずきに着目し、その機能について、より詳細な分析を試みる。

### 3.5 分析2：発話末のうなずきと聞き手の反応

前節で示したように、話し手の発話末のうなずきの頻度

は非常に高い。日本語の会話において、話し手の発話末のうなずきはどのような役割を果たしているのであろうか。そこで、発話末に話し手のうなずきがある場合とない場合に対し、聞き手の反応(あいづち、または、うなずき)がどのように行なわれるのかについて分析を行なった。

表2 話し手の発話末のうなずきに対する聞き手の反応(うなずき・あいづち)の出現頻度

話し手のうなずき	話し手と聞き手の関係	聞き手の反応			反応なし	合計
		うなずき あいづち	うなずきのみ	あいづちのみ		
あり	交替	—	35 (44%)	—	45 (56%)	80
	継続	31 (27%)	40 (35%)	11 (10%)	32 (28%)	114
なし	交替	—	20 (33%)	—	40 (67%)	60
	継続	13 (10%)	17 (13%)	11 (8%)	91 (69%)	132

表2に話し手の発話末のうなずきに対する聞き手の反応を示す。表から、話し手の発話末にうなずきがある場合とない場合を比較すると、話者交替と話者継続、どちらの状況でも、うなずきがない場合よりもうなずきがある場合の方が聞き手の反応の頻度が高い。このことから、話し手の発話末のうなずきは聞き手への何らかの働きかけの機能があり、聞き手はそれに対して反応しているように見える。また、話者交替に比べ、話者継続時の方がより一層、その傾向が強いことがわかる。

### 3.6 分析3：言語情報・韻律情報と聞き手の反応

前節の分析より、話し手の発話末のうなずきには、話し手と聞き手の会話制御に関する何らかの機能が含まれていることが示唆されたが、ここで、あいづちに関する過去の研究<sup>1),7)</sup>を考慮すると、言語情報や韻律情報も発話末のうなずきと同様に聞き手の反応との関係が深いと考えられる。

そこでまず、言語情報として、メイナード<sup>1)</sup>らが指摘しているように、「ね」「さ」「よ」などの終助詞や間投助詞、あるいは、「じゃない」「でしょ(う)」などの助動詞の文末表現などがあいづちの直前のコンテキストに表われやすいということを考慮して、発話末の言語情報と聞き手の反応について分析を行なった。

表3に発話末にこれらの言語情報がある場合とない場合の聞き手の反応の違いを示す。なお、表中、「言語情報あり」とは、終助詞・間投助詞として「さ」「ね」「の」「よ」「な」「じゃん」、助動詞として「でしょ」と「だ」を分析対象とした\*。

表3から、ある特定の助詞や助動詞が発話単位末に出現する場合としない場合を比較すると、ここでも上述の発話末のうなずき同様、話者交替と話者継続、どちらの状況でも、言語情報がない場合よりもある場合の方が聞き手の反応の頻度が高い。また、話者交替時に比べ、話者継続時の

\* これら8個の助詞と助動詞は、今回の分析データに出現したもののみを列挙しており、それ以外に該当する助詞や助動詞は今回収録したデータには見受けられなかった。

表3 話し手の発話末の言語情報に対する聞き手の反応（うなずき・あいづち）の出現頻度

	話し手	話し手	聞き手の反応			反応なし	合計
			うなずき あいづち	うなずき のみ	あいづち のみ		
発話末の言語情報	あり	交替	—	28 (47%)	—	32 (53%)	60
		継続	17 (24%)	22 (32%)	6 (9%)	24 (35%)	69
	なし	交替	—	27 (34%)	—	53 (66%)	80
		継続	27 (15%)	35 (20%)	16 (9%)	99 (56%)	177

方がその傾向が強い。

次に発話末の韻律的特徴について検討を行なう。以前行なった研究<sup>5)~8)</sup>では、発話末の韻律的特徴が話者の交替や継続、あいづち現象などと関係があることが示唆された。そこで本研究では、韻律的特徴の中でもとくに顕著な山型イントネーション（尻上がりイントネーション）と発話末のモーラの音引き現象（山型イントネーションは除く）に焦点をあて、これらの韻律的特徴と聞き手の反応について分析を行なった。山型イントネーションとは最後のモーラが音引きされ、上昇下降調の独特のイントネーションを形成するものである。音引き現象とは発話単位の最後のモーラがそれ以前の平均的なモーラ長に比べ、引き伸ばされる現象である。おもに助詞などにその現象が見られるが、名詞の最後のモーラを伸ばす現象なども多々見られた。

表4に発話末の韻律的特徴（山型イントネーション・音引き現象・特徴なし）と聞き手の反応の関係を示す。

表4 話し手の発話末の韻律情報に対する聞き手の反応（うなずき・あいづち）の出現頻度

	話し手	話し手	聞き手の反応			反応なし	合計
			うなずき あいづち	うなずき のみ	あいづち のみ		
発話末の韻律的特徴	山型	交替	—	7 (32%)	—	15 (68%)	22
		継続	21 (32%)	18 (27%)	4 (6%)	23 (35%)	66
	音引	交替	—	7 (39%)	—	11 (61%)	18
		継続	5 (14%)	4 (11%)	3 (8%)	24 (67%)	36
	なし	交替	—	41 (41%)	—	59 (59%)	100
		継続	18 (13%)	35 (24%)	15 (10%)	76 (53%)	144

表4を見ると、山型イントネーション、音引き現象が生じた場合、話者交替に比べ、話者継続の頻度が高くなっていることがわかる。このような現象は従来研究の結果と一致している。

また、山型イントネーションが生じた場合には聞き手の反応が非常に強くなっており、聞き手の反応をうながしていることが示唆される。一方、音引き現象に関しては韻律的特徴がない場合とほとんど差が見られなかった。これは、今回の分析では、音引き現象を1パターンのみとしてまと

めて扱ってしまったためではないかと考えられる。文献7)、8)などで指摘したように、音引き現象も様々なパターンがあるため、それらをいくつかのカテゴリーに分けて分析した方が傾向が出たのではないかと考えられる。

### 3.7 分析4：話し手のうなずき・言語情報・韻律情報の相互的關係の分析

上述の分析では話者交替と話者継続をともに扱ったが、話者交替時に生じるうなずきはメイナード<sup>1)</sup>のうなずきのカテゴリー2に相当し、聞き手が次に発話権を獲得し、発話を行なうことを示すものである。一方、話者継続時のうなずきはメイナード<sup>1)</sup>のカテゴリー1に相当し、あいづちと同様に扱われる。すなわち、話者継続時のうなずき、あいづちは聞き手のあいづち的反応ととらえることができる。ところで、これまでの分析から、発話末の話し手のうなずき、特定の言語情報、特定の韻律的特徴がそれぞれ聞き手の反応をうながすことが示唆された。では、これら三つの要因間の相互的關係はどのようになっているのだろうか。そこで、話者継続時における上記三要因と聞き手の反応を表5にまとめた。なお、ここでは分析の見通しをよくするため、うなずき、あいづちの区別をせず、聞き手の反応として、うなずきかあいづちのどちらかがあった場合を反応あり、どちらもなかった場合を反応なしとして記述することとする。

表5から、おおむね各要因が存在するときに聞き手の反応が顕著であることがわかる。

表5にはすべての要因が書かれているが、各要因ごとの違いを明らかにするため、ここでは二要因ずつ別々に分析することとする。まずはじめに、表6に発話末の話し手のうなずき、言語情報と聞き手の反応を示す。

表6 話し手のうなずき・言語情報と聞き手の反応の關係。各欄は反応のあり/なしの頻度と反応ありの比率

	話し手	うなずき		合計
		あり	なし	
言語情報	あり	34/13 (72%)	11/11 (50%)	45/24 (65%)
	なし	48/19 (72%)	30/80 (27%)	78/99 (44%)
合計		82/32 (72%)	41/91 (31%)	123/123 (50%)

表6から、話し手の発話末にうなずきがある場合には、非常に高い頻度で聞き手が反応していることがわかる（31%→72%）。一方、言語情報については、言語情報なしの44%に比べ、言語情報ありの場合、65%となっているため、聞き手の反応に影響を与えていることは示唆されるが、うなずきほど強い要因とはなっていないようである。また、うなずきがある場合にはともに72%となっていることから、うなずきと同時に発現してもその効果は薄れてしまうのかもしれない。逆にうなずきが生じていなければ、27%から50%へと向上しており、やはり何らかの影響を及ぼしていると考えるのが妥当であろう。

次に、表7に発話末の話し手のうなずき、韻律情報と聞

表5 話し手のうなずき・言語情報・韻律情報と聞き手の反応の関係。各欄は反応のあり/なしの頻度と反応ありの比率

		韻律情報				合計	
		あり		なし		うなずき	
		あり	なし	あり	なし	あり	なし
言語情報	あり	13/5 (72%)	9/6 (60%)	21/8 (72%)	2/5 (29%)	34/13 (72%)	11/11 (50%)
	なし	22/11 (67%)		23/13 (64%)		45/24 (65%)	
韻律情報	あり	26/7 (79%)	8/29 (22%)	22/12 (65%)	22/51 (30%)	48/19 (72%)	30/80 (27%)
	なし	34/36 (49%)		44/63 (41%)		78/99 (44%)	
合計		39/12 (77%)	17/35 (33%)	43/20 (68%)	24/56 (30%)	82/32 (72%)	41/91 (31%)
		56/47 (54%)		67/76 (47%)		123/123 (50%)	

き手の反応を示す。

表7 話し手のうなずき・韻律情報と聞き手の反応の関係。各欄は反応のあり/なしの頻度と反応ありの比率

		うなずき		合計
		あり	なし	
韻律情報	あり	39/12 (77%)	17/35 (33%)	56/47 (54%)
	なし	43/20 (68%)	24/56 (30%)	67/76 (47%)
合計		82/32 (72%)	41/91 (31%)	123/123 (50%)

話し手の発話末にうなずきがある場合には、非常に高い頻度で聞き手が反応しているが、韻律情報については、聞き手の反応にはほとんど影響を及ぼしていないように見受けられる。しかしながら、うなずきがある場合に注目すると、韻律情報がない場合は68%であるのに対し、韻律情報が付加されると77%まで上昇している。このことは韻律が補助的に聞き手の反応をうながしている可能性を示唆しているが、それほど大きな差ではないので、今後、より詳細な検討が必要であると考えられる。

では、言語情報と韻律情報を比べるとどうなるであろうか。表8に発話末の話し手の言語情報、韻律情報と聞き手の反応を示す。

表8 話し手の言語情報・韻律情報と聞き手の反応の関係。各欄は反応のあり/なしの頻度と反応ありの比率

		韻律情報		合計
		あり	なし	
言語情報	あり	22/11 (67%)	23/13 (64%)	45/24 (65%)
	なし	34/36 (49%)	44/63 (41%)	78/99 (44%)
合計		56/47 (54%)	67/76 (47%)	123/123 (50%)

表8を見ると、どちらもそれほど高い要因とはなっていないが、お互いに補完しあって、影響を強めているようにも見える。また、その程度の違いから、韻律情報よりも言語情報の方が聞き手の反応に与える影響はわずかに大きいようである。

### 3.8 考察

これらの結果から、話し手のうなずきに関して、重要な示唆を得ることができた。まず第一に、話し手のうなずき

にはメイナードの指摘<sup>1)</sup>にもあるように、さまざまな機能があり、その使われるコンテキストにより、その機能は異なると考えられる。とくに、今回の分析結果を見ると、話し手の発話末のうなずきは発話単位の区切り、発話権の放棄の機能だけではなく、積極的に聞き手に反応をうながす機能を持っていることが推測される。また、発話末のうなずきは従来のあいづち研究などで言われてきた言語情報や韻律情報などと補いあって、聞き手の反応をうながしているが、その中でも、発話末のうなずきはその要因が強いということが示唆される。

最後にひとつ検討事項が残っている。今回の分析では話者交替時についての詳細な分析を行っていないが、話者交替時にも同じように話し手の発話末にうなずきが生じ、それを受けて、うなずいてから発話を開始する場面が非常に多く見受けられた。すなわち、発話末のうなずきは話者継続、話者交替の状況を問わず、ともに話し相手(聞き手)への積極的な会話への参加をうながすシグナルになっているのではないとも考えられるが、今回のデータからでは、これ以上のことは言えないので、今後、より詳細な分析を進めていきたいと考えている。

## 4. おわりに

本稿では日本語の自然対話における話し手のうなずきに注目し、話者継続時における発話末の話し手のうなずきと聞き手のあいづちやうなずきとの関連性を分析した。また、従来のあいづち研究などから指摘されてきた言語情報や韻律情報との比較・検討を行った。結果として、話し手の発話末のうなずきは聞き手の反応を強くうながす機能を有していることが示唆された。

なお、今回は視線の分析は行っていないが、視線はマルチモーダル対話において、非常に重要な役割を果たしているため、今後は検討していきたいと考えている。また、考察で指摘したように、発話末のうなずきの後、聞き手のうなずきに続いて話者交替するケースなども含め、さまざまな状況におけるうなずきについて、より詳細な分析を行っていく予定である。今回の分析では8人分のデータを用いているが、うなずきやあいづちなどは若干の個人差が見受けられたため、今後はそれらの影響も考慮し、収録データを増やしながらか、さらに深く検討していきたいと考えている。

## 参 考 文 献

- 1) 泉子 K メイナード. 会話分析. くろしお出版, 1993.
- 2) 小坂直敏. あいづちを中心とした会話音声の呼応関係の分析. 電子情報通信学会 信学技報 SP87-107, pp. 61-66, 1987.
- 3) 島津明, 川森雅仁, 小暮深. 対話の分析: 間投詞的応答に着目して. 電子情報通信学会 信学技報 NLC93-9, pp. 567-574, 1993.
- 4) 片桐恭弘, 川森雅仁, 島津明. あいづちの分散システムモデル. 言語処理学会第 1 回年次大会予稿集, 1995.
- 5) 小磯花絵, 堀内靖雄, 土屋俊, 市川薫. 下位発話単位の音声的特徴と「あいづち」との関連について. 人工知能学会研究会資料 SIG-J-9501-2, pp. 9-16, Dec. 1995.
- 6) 堀内靖雄, 小磯花絵, 土屋俊, 市川薫. 自発的音声対話における話者交替の制御に関わる発話末の統語的・韻律的特徴. 情報処理学会 研究報告, Vol. 96, No. 21, pp. 45-50, Feb. 1996. (96-SLP-10-9).
- 7) 小磯花絵, 堀内靖雄, 土屋俊, 市川薫. 先行発話断片の終端部分に存在する次発話者に関する言語的・韻律的要素について. 電子情報通信学会 信学技報, Vol. 95, No. 600, pp. 25-30, Mar. 1996. (NLC95-72).
- 8) Hanae Koiso, Yasuo Horiuchi, Syun Tutiya, Akira Ichikawa, and Yasahuru Den. An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese map task dialogues. *Language and Speech*, Vol. 41, No. 3-4, pp. 291-317, 1998.
- 9) 野口広彰, 片桐恭弘, 伝康晴. 尤度付きあいづち生起文脈コーパスの提案. 人工知能学会研究会資料 SIG-SLUD-A101-6, pp. 25-32, 2001.
- 10) 菊池英明, 工藤彦男, 小林哲則, 白井克彦. 音声対話インタフェースにおける発話権管理による割込みへの対処. 電子情報通信学会論文誌 D-II, Vol. J77-D-2, No. 8, pp. 1502-1511, 1994.
- 11) Nigel Ward. In Japanese a low pitch means "back-channel feedback please". 情報処理学会 研究報告 SLP-11-2, pp. 7-12, 1996.
- 12) 平沢純一, 川端豪. わかってうなずくコンピュータの試作. 情報処理学会 研究報告 SLP-19-20, pp. 131-138, 1997.
- 13) 平沢純一, 川端豪. 音声対話システム nobby: ユーザ発話途中でのうなずき・相槌生成. 情報処理学会 研究報告 SLP-20-9, pp. 51-52, 1998.
- 14) 岡登洋平, 加藤佳司, 山本幹雄, 板橋秀一. 韻律情報を用いた相槌の挿入. 情報処理学会論文誌, Vol. 40, No. 2, pp. 469-478, 1999.
- 15) 向井理朗, 関進, 中沢正幸, 綿貫啓子, 三吉秀夫. 力学系モデルに基づくマルチモーダル対話システムの試作. 1999 年度人工知能学会全国大会予稿集, 1999.
- 16) 向井理朗, 関進, 中沢正幸, 綿貫啓子, 三吉秀夫. 感情を持つマルチモーダル対話エージェントシステムの提案. 人工知能学会研究会資料 SIG-SLUD-9902-3, 1999.
- 17) M. Takeuchi, N. Kitaoka, and S. Nakagawa. Timing detection for realtime dialog systems using prosodic and linguistic information. In *Proceedings of Speech Prosody 2004*, pp. 1177-1180, 2004.
- 18) Satoru Hayamizu, Shigeki Nagaya, Keiko Watanuki, Masayuki Nakazawa, Shuichi Nobe, and Takashi Yoshimura. A multimodal database of gestures and speech. In *Proceedings of the 6th European Conference on Speech Communicational Technology*, pp. 2247-2250, 1999.
- 19) Michael Kipp. Anvil a generic annotation tool for multimodal dialogue. In *Proceedings of the 7th European Conference on Speech Communicational Technology*, 2001.
- 20) Tony Bigbee, Dan Loehr, and Lisa Harpr. Emerging requirements for multi-modal annotation and analysis tools. In *Proceedings of the 7th European Conference on Speech Communicational Technology*, 2001.
- 21) Ulrich Turk. The technical processing in smartkon data collection: a case study. In *Proceedings of the 7th European Conference on Speech Communicational Technology*, 2001.
- 22) 前田真季子, 堀内靖雄, 市川薫. 音声対話コーパスにおける画像情報のアノテーション手法の検討. 情報処理学会研究報告, Vol. 2001, No. 38, pp. 21-28, 2001.
- 23) 坂本憲治, 綿貫啓子, 外川文雄. マルチモーダル対話解析. 人工知能学会研究会資料 SIG-FAI-9401-6, pp. 39-46, 1994.
- 24) 綿貫啓子, 関進, 三吉秀夫. ヒューマンインタフェースのための人間の振舞いの解析: マルチモーダル対話データの解析. 情報処理学会研究報告 HI-84-5, 1999.
- 25) 綿貫啓子, 関進, 三吉秀夫. 発話時の人間の振舞い: マルチモーダル対話データの解析. 人工知能学会研究会資料 SIG-SLUD-9902-9, pp. 49-54, 1999.
- 26) 綿貫啓子, 関進, 三吉秀夫. 対話過程における身体の動きと音声との関係. 2000 年度人工知能学会全国大会予稿集, 2000.
- 27) Keiko Watanuki, Susumu Seki, and Hideo Miyoshi. Turn taking and multimodal information in two-person dialog. In *Proceedings of International Conference on Spoken Language Processing*, pp. 657-660, 2000.
- 28) 前田真季子, 堀内靖雄, 市川薫. 話者交替における視線とうなずきの分析. 人工知能学会研究会資料 SIG-SLUD-A201-09, pp. 53-58, Jun. 2002.
- 29) 前田真季子, 堀内靖雄, 市川薫. 自然対話におけるジェスチャーの相互的関係の分析. 情報処理学会 研究報告, Vol. 2003, No. 9, pp. 39-46, Jan. 2003.
- 30) 前田真季子, 西田昌史, 堀内靖雄, 市川薫. 自然対話における発話者のうなずきに対する聞き手の反応. 人工知能学会研究会資料 SIG-SLUD-A302-07, pp. 35-42, Nov. 2003.
- 31) 矢野博之, 善本淳. 合意形成対話における同意表現の言語・非言語情報の分析. 人工知能学会研究会資料 SIG-SLUD-A203-07, pp. 41-46, 2003.
- 32) 堀内靖雄. 音声対話コーパス作成, 分析ツールについて: 日本語地図課題対話コーパス作成の経験から. 人工知能学会研究会資料 SIG-SLUD-9902-5, pp. 23-28, Oct. 1999.