

# 対話エージェントのうなずきタイミングが 発話長に及ぼす影響分析

Cultural Study on Speech Duration and Perception of Virtual Agent's Nodding

貴志悠 神田智子

Haruka Kishi, and Tomoko Koda

大阪工業大学 情報科学部

Department of Information Science and Technology, Osaka Institute of Technology

**Abstract :** Nodding plays an important role in HAI as well as human-human communication. We compare speech durations and perceptions in cross-cultural human-agent interactions. The results indicates the importance of cultural adaptation of timing and frequency of virtual agent's nodding.

## 1. はじめに

近年, HAI (Human-Agent Interaction) 研究の発展により, 人とコンピュータエージェントとのコミュニケーションが注目を集めている[1]. 特に, 人同士の対話と同様にユーザと対面コミュニケーションを行えるコンピュータエージェントを対話エージェントと呼ぶ. 対話エージェントが人とより自然なコミュニケーションを成立させる要素として, 身振りやうなずき等の身体表現によるノンバーバルコミュニケーション能力を兼ね備えることが挙げられる. ノンバーバルコミュニケーションの中でもうなずきという動作は, 話の聞き手から話し手に対する視覚的フィードバックとしてコミュニケーションを円滑に進める上で非常に重要な役割を果たしている. うなずきはあいづちの一つの形態であり, 非言語表現としてのあいづちに属する[2]. 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送る短い表現のことを指し, 「うん」「ふうん」等の言語表現とうなずき等の非言語表現に分類される. あいづちの言語表現と非言語表現は伴って使われることもある. また, あいづちには文化差があるとされており, 特

に日本とアメリカではあいづちの文化差が頻度とタイミングに顕著に表れ, 日本人はアメリカ人の2倍も多くあいづちを打つとされている. タイミングに関しても日本人は文節や文末などにあいづちを打つものに対して, アメリカ人は文末のみに集中するとされている[2]. コミュニケーションにおけるあいづちの頻度とタイミングの重要性について, 例えば日本人とアメリカ人があいづちの文化差を考慮せずにコミュニケーションを行った場合, 聞き手の日本人による日本の頻度とタイミングでのあいづちは, 頻繁にあいづちを打たれることに慣れていない話し手のアメリカ人にとっては話の邪魔をされていると感じるなど, 誤解を招く可能性も考えられる[3]. 一方で, 聞き手のアメリカ人によるアメリカの頻度とタイミングでのあいづちは, 頻繁にあいづちを打たれることに慣れている話し手の日本人にとっては話を聞いていない, または理解していないと感じるなど, あいづちの文化差によって結果的に良好な人間関係の構築が困難になる可能性も考えられる. そこで, 人同士が円滑なコミュニケーションを成立するためには, 話し手となる人に文化適応した頻度とタイミングで話の聞き手となる人があいづちを打つことが重

要だと言える。

本研究では人同士のコミュニケーションと同様に、人と対話エージェントのコミュニケーションにおいても話し手となる人に文化適応した頻度とタイミングで話の聞き手となるエージェントがあいづちを打つことが重要であるとし、言語表現を伴わないうなずきのみに着目して実験を行うことで検証をした。実験では、『日本人実験参加者は、日本の頻度とタイミングでうなずくエージェントとの対話の方が、アメリカの頻度とタイミングでうなずくエージェントよりも発話長が長くなり、よりリラックスできる』と仮説を立て、日本の頻度とタイミングでうなずくエージェント（以下、JA）、アメリカの頻度とタイミングでうなずくエージェント（以下、AA）、および、うなずかないエージェント（以下、NA）を用いて発話長の比較を行った。

## 2. 関連研究

本章では、あいづちの定義・日米文化差について述べる。

### 2.1 あいづちとうなずきの定義

「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送る短い表現のことを指し、短い表現のうち話し手が発話権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、「聞き役中」の聞き手の反応に焦点をあてたものをあいづちと定義する[2]。あいづちは、言語表現と非言語表現に分類される。言語表現として、「うん」「ふうん」等の短い表現がある。また、会話の中で話者が交わす笑いや笑いに似た発声も、会話の進行上重要な役目を果たすと考えられ、会話という相互作用の中で一定の機能を果たしていることからあいづちの一つの形態とされている。非言語表現としては、うなずき(頭の縦ふり動作)が主に上げられる。非言語表現は言語表現を伴って使われることもある。非言語表現のみで使われるうなずきは「身振りのあいづち」として言語表現のあいづちと同じ機能を持ちうるとされている[4]。よって、本研究では言語表現を伴わない

うなずきのみでのあいづちに着目する。

### 2.2 あいづちの文化差

本研究では泉子・K・メイナードの会話分析[2]を基にあいづちの文化差が頻度とタイミングに顕著に表れている日本とアメリカに着目した。[2]より、日本とアメリカのあいづちの頻度を表1に、あいづちのタイミングを表2に示す。なお、日本とアメリカのあいづちデータの収集方法には、日本人同士とアメリカ人同士それぞれ二人一組各20組の日常会話を録画した計120分間のビデオデータが使用されている。

表1から日本人とアメリカ人では日本人の方があいづちを打つ回数が2倍以上多いことがわかる。また表2からあいづちを打つタイミングに関しても、日本人は文末のポーズ付近や、終助詞・間投助詞付近などいたるところであいづちを打っていることが分かる。言語体系の違いにもよるが、特に日本人は、話し手の頭の動きに聞き手もつられてあいづちを打つことが多い。これに対してアメリカ人のあいづちはあいづち総数のうち80%以上を文末のポーズ付近で占めており、日本人のようにいたるところであいづちを打たないのは明確である。

本研究では、このあいづちの頻度とタイミングの違いを日本とアメリカのあいづちの文化差とし実験を行なった。

表1. 日米のあいづちの頻度[2]

日米、各20組計120分の日常会話データ	日本語	英語
合計	948回	364回

表2. 日米のあいづちのタイミング[2]

日米、各20組計120分の日常会話データ	日本語	英語
文末のポーズ付近	351回	309回
終助詞・間投助詞付近	281回	—
付加疑問付近	54回	26回
話し手の頭の動き付近	262回	29回

## 3. 実験

本章では、JA、AA、NAとの対話における発話長の比較実験についての詳細を記述する。

### 3.1 関連研究

本実験はMatarazzoらによる面接官のうなずきと

被面接者の発話長の関係を調べた実験[5]を参考にし  
て行った。これは面接場面における面接官のうなず  
きを統制し、それによって被面接者の発話長がど  
のように影響されるのか検討したものである。具体  
的には、警察官と消防士の採用試験の場で、60人  
の男性志願者に1人3テーマ(職歴, 教育歴, 家族歴)  
について1テーマ15分を目安に話してもらう面接を  
行った。実験条件として通常グループ(20人)では  
面接官は3テーマとも自然に応答するが、実験グル  
ープ1, 実験グループ2(各20人)では、面接官が  
2テーマ目のみ被面接者が話している間は継続的  
にうなずき、話し終えるまで繰り返す。3グル  
ープとも面接官の発言は統制され、1回あたりの  
発言は5秒以内と決められている。さらに微笑な  
どの承認的な行動は極力避けるように配慮され  
た。面接官の見かけの効果を考慮して、実験グル  
ープ1, 実験グループ2はそれぞれ別の人物が  
面接官の役割を担当した。図1は各グループの  
被面接者の一文あたりの平均の発話長である。実  
験グループ1, 実験グループ2ともにほとんどの  
被面接者がテーマ1からテーマ2にかけて発話長  
が増加し、テーマ2からテーマ3にかけて発話長  
が減少した。通常グループでは発話長にこのよ  
うな変化はなかった。この結果から、面接官の  
うなずきによって被面接者の発話長が増加したと  
言える。

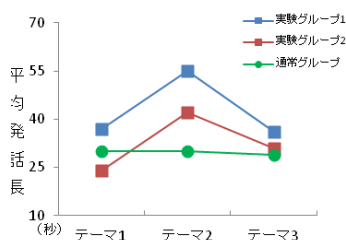


図1. 被面接者の平均発話長の推移[5]

### 3.2 使用エージェント

本実験で使用する対話エージェントには、先行研  
究[6]で開発された聞き手としての傾聴感が最も高  
いと思われる回数往復2回、角度垂直方向に20°速  
度往復0.6秒/回でうなずき動作を行なう動物型の対

話エージェントを用いた。対話エージェントの開発  
環境にはMicrosoft Visual C++ [7]およびDirectX  
SDK[8]、モデリングにはMetasequoiaLE[9]が使用  
されている。エージェントのうなずき動作の例を図2  
に示す。

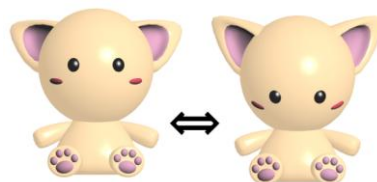


図2. エージェントによるうなずき動作の例

### 3.3 実験手順

本実験では対話エージェントのうなずき動作を制  
御するのに実験参加者にはシステムが存在してい  
るのように見せかけ、実際は実験者がシステムを操  
作するWizard of Oz法を用いた。音声は、ビデオ撮  
影をすることで収録した。

実験参加者に実験手順について説明した後、1テ  
ーマ5分を目安に均等に話せる3テーマを選択して  
もらった。テーマとして“芸能人, ペット, 食べ物,  
本・漫画, 映画, バイト, ファッション, 音楽, ス  
ポーツ, 勉強, ニュース, TV番組, 旅行, ギャン  
ブル, ゲーム, 車・バイク, 料理, その他”の計18項  
目を用意し、その他のみ重複を許した。その場合、  
話す内容は違う内容に限った。

実験参加者には1テーマ毎に話し終えた時点で  
「以上です」と言ってもらい、その後アンケートに  
回答してもらう(「以上です」は発話長の対象にしな  
い)。3テーマ終了時点で実験は終了する。尚、エ  
ージェントの提示順およびテーマの提示順はランダ  
ムとした。

実験参加者は大阪工業大学の学生29名(男性17  
名, 女性12名)である。

### 3.3 主観評価アンケート

実験参加者には、各テーマ終了後に、アンケート  
に回答してもらった。

アンケートの内容は、①対話そのものに対するス  
トレス(10項目)、②対話のスムーズさ(4項目)、

③エージェントの積極性 (3項目), ④エージェントに対する好感度 (2項目), ⑤エージェントの見かけの親身度 (5項目) に関する質問で計 24 項目, 7 段階 (7:非常にそう思う - 1:まったくそう思わない) である。

## 4. 結果

本章では, JA, AA, NA との対話における発話長の比較実験の結果を示す。

### 4.1 発話長の比較実験の結果

#### 4.1.1 発話長の計測結果

録画したビデオデータの音声部分を Any Audio Converter(<http://www.any-audio-converter.com/jp/>) で波形として抽出し, 無音区間を除去することで発話長を計測した。その際, 雑音が発話長に含まれないようにした。実験参加者ごとの発話長の計測結果を表 3 に示す。尚, 実験参加者には 5 分を目安として話をしてもらっていたため, 無音区間を除去して 5 分 (300 秒) を超えるデータは省いた。その結果, 被験者は 20 名 (男性 11 名, 女性 9 名) となった。

表 3. 実験参加者ごとの発話長

実験参加者	実験参加者の発話時間 (秒)		
	JA	AA	NA
1	77(20)	36(5)	28
2	38(12)	28(8)	22
3	237(47)	207(20)	151
4	105(26)	93(17)	58
5	97(24)	80(10)	77
6	122(21)	112(13)	78
7	43(12)	31(5)	25
8	93(17)	81(6)	64
9	208(23)	80(16)	118
10	115(27)	45(10)	56
11	166(21)	76(15)	152
12	299(45)	239(15)	242
13	58(7)	34(4)	47
14	105(24)	59(8)	60
15	71(17)	83(9)	148
16	105(23)	190(17)	56
17	30(10)	42(4)	37
18	73(16)	133(12)	104
19	40(14)	44(8)	28
20	28(10)	29(5)	33
平均	105.5(20.8)	86.1(10.35)	79.2

()内はエージェントのうなずき回数

表 3 より, 20 名の実験参加者のうち 14 名 (表 3 の実験参加者 1~14) において, JA の場合が最も発話長が長いことが分かる。さらに, 8 名 (表 3 の実験参加者 1~8) において, JA, AA, NA の順に発話長が長くなっていることが分かる。うなずき回数に関しては, JA は平均 20.8 回, AA は平均 10.35 回で

あった。

#### 4.1.2 発話長の多重比較結果

実験参加者の発話長を JA - AA - NA 間で多重比較を行った結果を図 3 に示す。

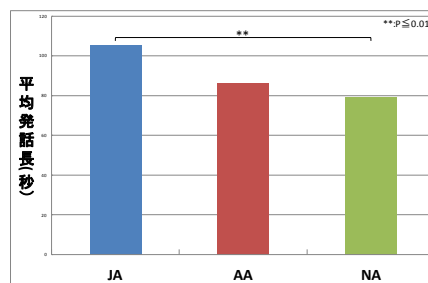


図 3. 実験参加者の平均発話長の比較

図 3 より, JA - NA 間で有意差 ( $p \leq 0.01$ ) が見られた。JA - AA 間, AA - NA 間では有意差は見られなかった。

### 4.2 主観評価に関する分析結果

#### 4.2.1 対話そのものの主観評価の分析結果

対話そのものに対する主観評価として, 対話そのものに対するストレス (10 項目) 及び対話のスムーズさ (4 項目) について, それぞれ 3 水準の一元配置分散分析を行った結果を図 4 に示す。尚, 対話そのものに対するストレスは, ストレスが低い方が評価が高くなるように, 回答を反転している。

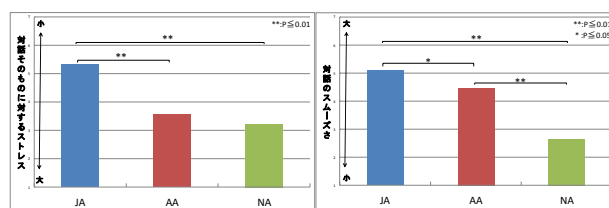


図 4. 対話そのものに対するストレス(左), 対話のスムーズさ(右)の分析結果

図 4 より, 対話そのものに対するストレスについて, JA - AA 間, JA - NA 間で有意差 ( $p \leq 0.01$ ) が見られた。

対話のスムーズさについて, JA - NA 間, AA - NA 間で有意差 ( $p \leq 0.01$ ) が見られた。また, JA - AA 間で有意傾向 ( $p \leq 0.05$ ) が見られた。

#### 4.2.2 エージェントの印象の主観評価の分析結果

エージェントの印象に対する主観評価として、エージェントの積極性（3項目）、エージェントに対する好感度（2項目）、エージェントの見かけの親身度（5項目）について、それぞれ3水準の一元配置分散分析を行った結果を図5、図6に示す。

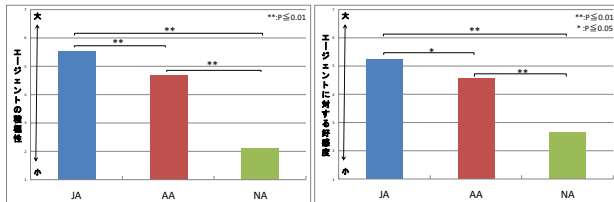


図5. エージェントの積極性(左), エージェントに対する好感度(右)の分析結果

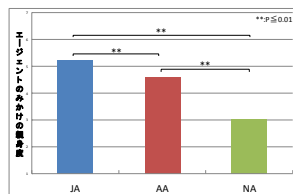


図6. エージェントの見かけの親身度の分析結果

図5より、エージェントの積極性について、すべてのエージェント間で有意差( $p \leq 0.01$ )が見られた。また、エージェントに対する好感度について、JA - NA間、AA - NA間で有意差( $p \leq 0.01$ )が見られた。JA - AA間で有意傾向( $p \leq 0.05$ )が見られた。図6より、エージェントの見かけの親身度について、すべてのエージェント間で有意差( $p \leq 0.01$ )が見られた。

## 5. 考察

本章では、実験結果の考察について述べる。

### 5.1 発話長の比較実験の考察

表3より、20名の実験参加者のうち14名において、JAの場合が最も発話長が長く、さらに8名の実験参加者において、JA, AA, NAの順に発話長が長くなっていることが分かる。うなずき回数に関しては、JAは平均20.8回、AAは平均10.35回で、日本のうなずき頻度はアメリカのうなずき頻度の約2倍であり、泉子・K・メイナードの会話分析の日米のあいづちの頻度[2]と同様の結果である。発話長の増減

とうなずきの関係について、図3よりJA - NA間で有意差( $p \leq 0.01$ )が見られたことから、対話エージェントによるうなずきが実験参加者の発話長を増加させる要因になっていたと言える。また、JA - AA間、AA - NA間では有意差は見られなかった要因として、発話長のばらつきが挙げられる。本実験では1テーマ5分を目安としたが、実験参加者によって、対話の得意さや、選んだテーマの中でも話しやすさにはばらつきがあったことが原因だと考えられる。

### 5.2 主観評価に関する考察

#### 5.2.1 対話そのものに対する主観評価の考察

図4より、対話そのものに対するストレスについて、JA - AA間、JA - NA間で有意差( $p \leq 0.01$ )が見られたことから、日本人実験参加者にとって、対話エージェントが日本の頻度とタイミングでうなずくことで、最もストレスが軽減されることが分かる。

対話のスムーズさについては、JA - NA間、AA - NA間で有意差( $p \leq 0.01$ )、JA - AA間で有意傾向( $p \leq 0.05$ )が見られたことから、対話エージェントが全くうなずかない場合より、うなずき動作があることで対話がよりスムーズになることが分かる。また、日本の頻度とタイミングでうなずきがアメリカの頻度とタイミングのうなずきよりも対話をスムーズにする傾向がある。

これらから、日本人実験参加者にとって、聞き手が日本の頻度とタイミングでうなずくことで対話そのもののストレスが軽減され、聞き手がうなずき動作を行うことで対話をよりスムーズに行うことができるということを示唆しており、実験参加者の文化に合わせたうなずきが対話をする上で重要であることが分かる。

#### 5.2.2 エージェントの印象の主観評価の考察

図5より、エージェントの積極性について、すべてのエージェント間で有意差( $p \leq 0.01$ )が見られたことから、対話エージェントが全くうなずかない場合よりもうなずき動作がある方がエージェントに積極性が感じられ、さらに、日本の頻度とタイミングでうなずく方がアメリカの頻度とタイミングでうなず

く場合よりもエージェントに積極性が感じられることが言える。

エージェントに対する好感度について、JA - NA 間、AA - JA 間で有意差( $p \leq 0.01$ )、JA - AA 間で有意傾向( $p \leq 0.05$ )が見られたことから、対話エージェントが全くうなずかない場合より、うなずき動作がある方が、エージェントに好感を持てることが分かる。また、日本の頻度とタイミングでうなずきがアメリカの頻度とタイミングのうなずきよりもエージェントに好感を持てる傾向にあることが言える。

図 6 より、エージェントの見かけの親身度について、すべてのエージェント間で有意差( $p < 0.01$ )が見られたことから、対話エージェントが全くうなずかない場合よりもうなずき動作があることで、エージェントが親身に感じられ、さらに、日本の頻度とタイミングでうなずく方がアメリカの頻度とタイミングでうなずく場合よりもエージェントが親身に感じられる。

これらから、日本人実験参加者にとって、聞き手が日本の頻度とタイミングでうなずくことで、エージェントの積極性、好感度、親身度が上がることが言え、実験参加者の文化に合わせたうなずきをすることで、話し手により印象を与えられるということが言える。

## 6. おわりに

本研究では、人と対話エージェントのコミュニケーションにおいてあいづちの文化差を考慮する重要性を検証するために、日本語と米語それぞれの頻度とタイミングでうなずくエージェント、うなずかないエージェントを用いて発話長の比較実験を行った。その結果、実験参加者の 7 割が日本の頻度とタイミングでうなずくエージェントとの対話で最も発話長が長くなっていた。また、アンケートの結果より、対話エージェントが日本の頻度とタイミングでうなずくことで対話そのもののストレスが軽減され、対話をよりスムーズに行うことができ、エージェントに対してよい印象を持つことが分かった。このこと

から、人と対話エージェントのコミュニケーションにおいて、話し手となる人に文化適応した頻度とタイミングで話の聞き手となる対話エージェントがうなずくことの重要性が示唆された。

本実験では、実験参加者は日本人のみだったため、今後の展望として、同様の実験をアメリカ人など異文化の参加者で行うことで、より各文化に適応したノンバーバルコミュニケーション能力を備えた対話エージェント開発の足がかりになると期待している。また、ストレスは主観的なアンケートのみであったが、皮膚電気反射(GSR)等の生理指標を用いることで、ストレスを客観的に捉えることができると考えている。

## 謝辞

本研究の一部は科学研究費補助金（基盤研究© 2011-2013 年）によるものである。

## 参考文献

- [1] 山田誠二,2007,人とロボットの<間>をデザインする,東京電機大学出版局,pp1-22
- [2] 泉子・K・メイナード,1993,会話分析,くろしお出版
- [3] Lebra,T.S.,1976,Japanese patterns of behavior, Honolulu: University of Hawaii Press
- [4] 杉戸清樹,1989,ことばのあいづちと身ぶりのあいづち—言語行動における非言語的表現—,日本語教育 67, pp.48-59
- [5] Matarazzo,JD,Saslow,G,Wiens,AN,Weitman,M,Allen,BV,1964,Interviewer head nodding and interviewee speech durations,Psychotherapy,Theory, Research and Practice,1,pp.54-64
- [6] 脇坂昌志,会話エージェントの自然なうなずきの頻度,角度,速度に関する分析,大阪工業大学情報科学部 2009 年度卒業論文
- [7] Visual C++デベロッパーセンター,<http://msdn.microsoft.com/ja-jp/visual/default.aspx>
- [8] DirectX デベロッパーセンター,<http://msdn.microsoft.com/ja-jp/direct/default.aspx>
- [9] metaseq.net,<http://www.metaseq.net/>