

ロボット受付嬢 SAYA を用いたリアルなうなづきに関する研究*

橋 本 卓 弥^{*1}, 平 松 幸 男^{*2}
辻 俊 明^{*3}, 小 林 宏^{*4}Research on Realistic Nod with Receptionist Robot SAYA That
Has Human-like AppearanceTakuya HASHIMOTO^{*5}, Sachio HIRAMATSU,
Toshiaki TSUJI and Hiroshi KOBAYASHI^{*5} Department of Mechanical Engineering, Tokyo University of Science,
1-3 Kagurazaka, Shinjyuku-ku, Tokyo, 162-8601 Japan

In face-to-face communication, the speaker as well as the listener nods during conversation. We refer to such behavior as "Speaker's nod". In this paper, a "Speaker's nod" is realized as a natural interpersonal behavior in the receptionist robot SAYA, which has human-like appearance. First, we investigated the timing, angle and angular velocity of the human "Speaker's nod". Based on the results, we developed a "Speaker's nod" function for SAYA. As a result, we found that "Speaker's nod" appears at the end of utterance frequently and appropriate nod-angle and angular velocity improve the humanness of SAYA. We then confirmed the effectiveness of "Speaker's nod" by a subjective experiment. The results indicate that appropriate nod timing, angle and angular velocity improve not only the humanness but also communicative effectiveness of SAYA in face-to-face communication with humans.

Key Words: Receptionist Robot, Human-robot Interaction, Nod, Face-robot, Android

1. 結 言

人間同士のような自然な対面コミュニケーションには、言語情報だけでなく非言語情報が大きな役割を果たしていると言われている⁽¹⁾⁽²⁾。そこで、人間と人工物とのコミュニケーションにおける非言語情報の有効性に注目し、従来から HCI(Human-Computer Interaction)の分野では、ディスプレイ上の仮想エージェントに非言語機能を持たせることにより、遠隔地のユーザや情報端末とのコミュニケーションにおける自然な対話型インタフェースを実現する試みがなされている⁽³⁾⁽⁴⁾。また、ロボットの分野においても、ロボットの身体性を積極的に活用し、視線やうなづき、ジェスチャーなどの非言語表現により人間とのコミュニケーションを円滑にする試みがなされている⁽⁵⁾⁽⁶⁾⁽⁷⁾。さらに、人間と情緒的・感性的に係わり合うことを志向したロボットの研究開発も行われており^{(8)~(10)}、これらのロボットは、将来的に人間と空間を共有し、オフィスや日常生活の

中で人間を支援することが期待される⁽¹¹⁾⁽¹²⁾。

この様に、これまでに多くのコミュニケーションロボットが開発されてきたが、その外観は非常に機械的なものであった。一方で、一見しただけでは人間なのか機械なのか見分けがつかないようなロボットの開発も行われている⁽¹³⁾⁽¹⁴⁾。小林らは、コミュニケーションにおける表情の重要性に着目し、人間にそっくりな顔ロボットを用いて自然な表情表出の実現を目指している⁽¹³⁾。また、石黒らは、全身型のアンドロイドを用いて自然な動作生成や認知科学的な評価を試みている⁽¹⁴⁾。このようなアンドロイドと呼ばれる人間にそっくりなロボットを用いることの大きな利点は、あたかも人間と接しているような感覚でロボットと接することが出来るということである。擬人化しただけの機械的なロボットでは人間のような豊かな非言語情報は表現しきれないと考えられるが、人間にそっくりな見た目を持つロボットが、人間と同じような自然な動作で反応を返すことができれば、人間とロボットがごく自然にインタラクションできるようになると考えられる。

そこで本研究では、人間同士の対話中に見られる“話し手のうなづき”に注目し、人間のような外観を持つロボット受付嬢 SAYA⁽¹⁵⁾(図 1)に、人間のような自

* 原稿受付 2007 年 3 月 19 日。

^{*1} 正員、東京理科大学大学院工学研究科(〒162-8601 東京都新宿区神楽坂 1-3)。^{*2} 東京理科大学大学院工学研究科。^{*3} 正員、埼玉大学工学部(〒338-8570 さいたま市桜区下大久保 255)。^{*4} 正員、東京理科大学工学部。
E-mail: tak@kobalab.com

然なうなづき動作を実装することを目的とする。一般的にうなづきは、相槌に用いられる聞き手側の動作として考えられがちであるが、話し手もその発話中に頻繁にうなづいており、人間同士の対話における自然な動作だと言える。

うなづきなどの身体動作に関する代表的な研究として渡辺らの報告⁷⁾がある。この研究で開発されたインタラクティブロボットは、話し手の音声の on-off 情報を基にロボットがうなづいたり身体を動かしたりするというものであるが、この場合、文脈や言語情報、感情情報などは考慮されていない。一方、話し手がその発話中にうなづく場合、言葉や意図を強調したり、発話リズムや会話の流れを調整したりするような働きがあるとされており¹⁶⁾、むやみに表出されるものではないと推測される。また、人間らしいうなづき動作を実現するためには、うなづき時の動作の大きさや速さも重要な要素であると思われる。しかしながら、“話し手のうなづき”に関しては、表出位置に基づいた機能の分類に関する研究に留まっており¹⁶⁾、うなづき時の頭部動作の大きさや速さに関しては新たに検討する必要がある。

そこで本研究では、人間同士の対面コミュニケーションをモーションキャプチャシステムやビデオを用いて分析し、人間の発話中に見られるうなづきの表出タイミングやうなづき角度・角速度の分析を行う。次に、この分析で得られた人間のうなづき特性に基づいて作り込んだうなづきをリアルなうなづき動作とし、SAYA に実装する。最後に、被験者を用いた主観評価により対面コミュニケーションにおけるリアルなうなづきの効果を検証し、“話し手のうなづき”の動作設計の指針とその意義を明確にする。

2. ロボット受付嬢 SAYA

2.1 ロボット受付嬢 SAYA の概要 SAYA は大学構内において受付システムとして稼働しており、キャンパス案内や研究室紹介、事務室や研究室内の所在地・内線番号などの案内業務を行っている。

SAYA の頭部には、筆者らが開発している人間そっくりな表情を持つ顔ロボット¹³⁾¹⁷⁾を用いており、これをマネキンの身体に取り付けることにより、一見すると本物の女性のような外観を実現している(図 1(a))。

顔表情は人間同士のコミュニケーションにおいて非常に重要な要素であり¹⁸⁾、顔ロボットでは、人間のような表情表出を可能にするため、Ekman らが提案した FACS(Facial Action Coding System)¹⁹⁾を基に顔面上に 19ヶ所の制御点を設け、これらの動きを組み合わせる

ことで様々な表情表出を実現している(文献¹³⁾¹⁷⁾を参照)。特に、万国共通に認識・表出ができると言われる 6 基本表情(嫌悪、怒り、恐怖、悲しみ、驚き、幸福)の認識率は平均 90%以上である¹⁷⁾。

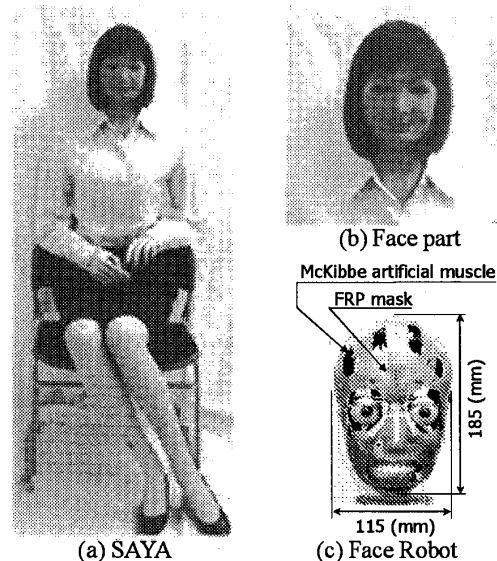


Fig.1 Receptionist Robot SAYA

2.2 ハードウェア構成 表情表出には空気圧で駆動する McKibben 型人工筋肉²⁰⁾²¹⁾を用いており、これを全て骨格内部に納めている。眼球は 2 自由度あり、DC モータによりヨーとピッチ方向の眼球運動が可能である。また、片眼には小型の CCD カメラが内蔵されている。

首の機構にはコイルばねを用いており(図 2)、人間のようにしなやかで柔軟なロール回転(Roll)、ピッチ回転(Pitch2)が可能である。また、頭部と首の付け根部分には、ヨー回転(Yaw)の 1 自由度とピッチ回転(Pitch1)の 1 自由度を設けている。ロール回転とピッチ回転には McKibben 型人工筋肉を用いており、ヨー回転には DC モータを用いている(図 3)。これらの可動域は図 2 に示す通りで、人間の可動域を十分満たしている²²⁾。

皮膚には人間と同じような質感を持たせるため、軟質ウレタン樹脂を用いている。

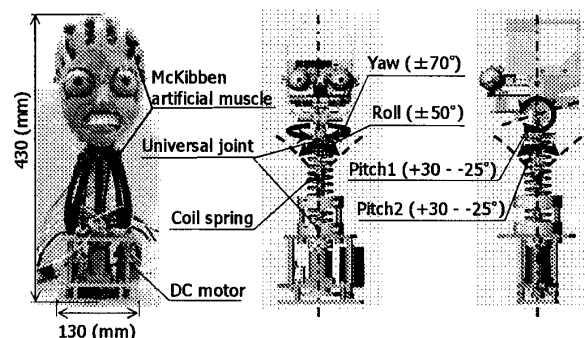


Fig.2 Internal structure of SAYA

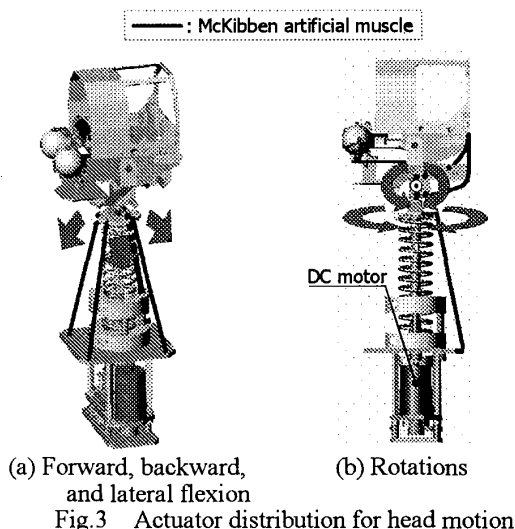


Fig.3 Actuator distribution for head motion

2.3 システム構成 SAYA のシステム構成を図4に示す。McKibben 型人工筋肉へ供給する圧縮空気の制御には、電空レギュレータを用いている。また、眼球内に埋め込まれた小型 CCD カメラを用いて肌色の認識を行い、人のいる方向に視線を向けるように眼球と首を制御している(本報には直接関係ないため詳細は省略)。さらに、音声認識と録音音声の再生により言葉による受応えも可能である。

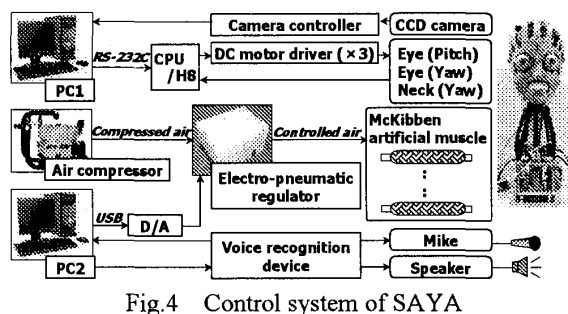


Fig.4 Control system of SAYA

対話者: こんにちは	P1(●)	P2	P3	P4	P5	P6	P7(●)
SAYA: こんにちは 大学 の こと わたし サヤ について ご 紹介 でき ます	P8	P9	P10(●)				
対話者: 大学について教えてください	P11	P12	P13	P14	P15	P16	
SAYA: それでは ご 希望 の 研究 室 名 または 事務 室 名 の どちら か を	P17(●)						
対話者: 小林研究室をお願いします	P18(●)						
SAYA: 小林 研究 室 です ね							
対話者: はい	P19	P20	P21	P22	P23	P24(●)	
SAYA: それでは 所在地 内線 番号 研究 室 紹介 を ご 案内 でき ます が	P25						
対話者: 所在地を教えてください	P27	P28	P29(●)				
SAYA: 小林 研究 室 は 3 号 館 7 階 です							

Fig.5 Conversation example

2.4 対話シナリオ 受付案内における SAYA と対話者との会話例を図5に示す(図中の区切りや記号に関しては後述)。この会話例に示すように、基本的に会話の主導権は SAYA 側にあり、SAYA からの質問に対話者が答える形で会話が進められる。音声認識には、

東芝 IT コントロールシステム株式会社製音声認識・合成ユニットを用いており、ワードスポッティング方式による簡単なキーワード認識を行っている。また、SAYA の音声には女性の録音音声を用いている。

3. “話し手のうなづき”の分析

3.1 分析方法 実験者と被験者を向かい合わせに座らせ(図6)、自由に会話をしてもらった。この時、実験内容に関する説明は行わなかった。また、モーションキャプチャシステム(Northern Digital Inc. 製 Optotrak Certus)により被験者の頭部の3次元位置の時間変化を記録し、同時に、ビデオカメラを用いて映像と音声の両方を記録した。モーションキャプチャのマーカー位置は図7に示す通りである。実験に参加した被験者は20代の男子学生5人で、実験者は1人に固定した。実験時間は被験者1人当たり5分程度であった。

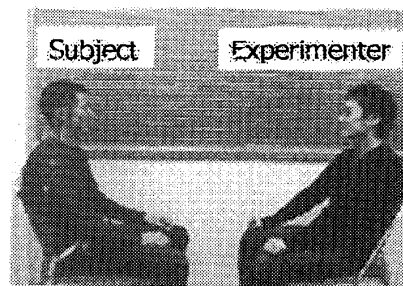


Fig.6 Overview of face-to-face conversation

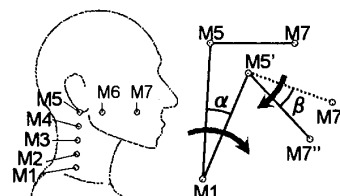


Fig.7 Marker position of motion capture

3.2 分析結果と考察

3.2.1 “話し手のうなづき”が表出するタイミング 撮影したビデオから、被験者の発話とうなづきとの関係を調査した。ここで、被験者のうなづきを、「実験者への相槌」、「発話途中」、「発話末」、「その他」の4つに分類することにした。このとき、自然発話には句読点などの明確な区切りがないため、200ミリ秒以上の無音区間を区切りとして分析を行った⁽²³⁾⁽²⁴⁾。被験者5人分の分析結果を図8に示す。

この結果から、被験者は、相槌以外にも自分の発話中に多くうなづいており、その中でも特に、発話末にうなづく頻度が極端に高いことが分かる。この発話末に見られるうなづきに関してメイナード⁽¹⁶⁾は、発話末や発話順番の終了を明確化するものだとしており、話者交替において重要な役割を果たすものとしている。

そこで本研究では、有意なうなづきタイミングとして発話末に注目する。ここで、文末が発話末になることが比較的容易に推測できるため、発話文の最後にうなづきを表出させればよいと考えられる。つまり、図5に示した SAYA の発話文の場合では、P7「ます」、P10「か」、P17「ください」、P18「ね」、P26「か」、P29「です」が発話されるタイミングに合わせてうなづきを表出させればよいと考えられる。

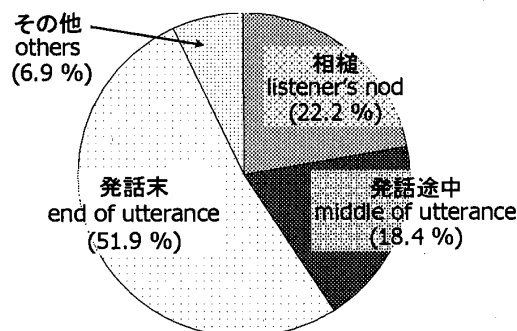


Fig. 8 Classification of "Speaker's nod"

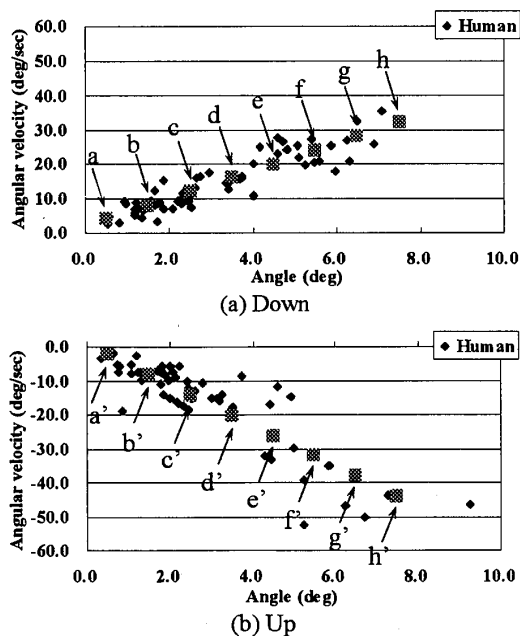


Fig. 9 Angle and angular velocity of "Speaker's nod"

3.2.2 “話し手のうなづき”の角度と角速度 モーションキャプチャとビデオの一連のデータから、うなづきが現れた区間を切り出し、被験者5人から各10～15サンプル、合計58サンプルを抽出した。そして、頭部が下降する方向を正の方向とし、角度と角速度を求めた(図9)。このとき、うなづきがほとんど頭部の回転だけで行われていたため、図7の β をうなづき角度とした。図9(a)(b)はそれぞれ、頭部が下降するときと上昇する(正面に頭を戻す)ときの角度と平均の角速度(以下、単に角速度と表記する)の関係を示している。

この結果より、うなづき角度・角速度共に広範に分布しているものの、角度が大きくなると、それに伴って角速度も大きくなる傾向があることが分かった(後の4.2に詳述するが、図9中a～h, a'～h'は、人間のうなづき角度・角速度を7等分したものである)。

4. “話し手のうなづき”の実装

4.1 うなづき表出タイミングの実験的検討

4.1.1 実験方法 3.2では、“話し手のうなづき”が発話末に高い頻度で現れることに注目し、発話文の最後に合わせてうなづきを表出させればよいと予測された。しかし、自然発話では文末以外にも発話の区切れが存在し、さらには発話の途中にもうなづきが現れる場合もあるため、うなづきの表出位置を明確に予測するのは難しい。そこで、図5に示した SAYA の発話において、実際にどういったタイミングでうなづきが現れるかを、被験者を用いて検討した。

実験では、受付の状況を想定し、図10に示す“Receptionist”の位置に被験者を座らせ、“Visitor”の位置に実験者を立たせた。このとき、被験者と実験者の距離はデスクを隔てて1.5m程度とした。この距離は社会距離⁽¹⁾に相当し、受付の状況としてこの程度が適切であると判断した。被験者には SAYA 役として図5中の SAYA の台詞を発話してもらい、実験者には来客者役として対話者の台詞を発話してもらった。この時の様子をビデオカメラにより撮影した。実験を行うに当たり、自然な会話ができるように、被験者と実験者にはあらかじめ会話の流れと台詞を覚えてもらい、台詞に慣れるまで練習してもらった。被験者は20代の学生20人(男性17人、女性3人)で、各被験者2回ずつ実験を行った。このとき、実験者は1人に固定した。

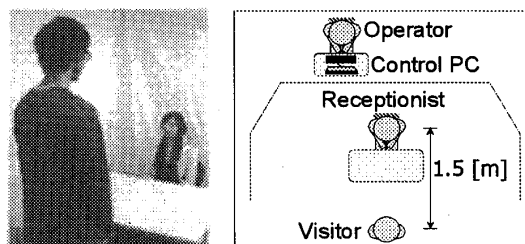


Fig. 10 Overview of experiment

4.1.2 実験結果と考察 上記の実験により、20人×2回の計40回分の音声と映像のサンプルが得られる。これを解析するに当たり、“話し手のうなづき”と言語情報との関係を見るために、SAYA の発話文を図5に示すように品詞に分解した。次に、40サンプル中でうなづきが現れた位置を抽出し、それぞれに番号を付けた。そして、各番号の位置でうなづいた被験者の数

を累計し、被験者の総数を1としたときの番号毎の被験者の割合を求めた(図11)。このときのうなづき回数は、平均で約10.5回(標準偏差5.07)であった。次に、図11の縦軸を0.1間隔で区切り、各区分内にある棒グラフの先端の数を縦軸にとったグラフを作成した(図12)。これに判別分析を行ったところ、 $0.5 < \leq 0.6$ を境に2つに分けられることが分かった。そこで、この境界以上の範囲に注目し、P1「こんにちは」、P7「ます」、P10「か」、P17「ください」、P18「ね」、P24「が」、P26「か」、P29「です」の8ヶ所が抽出できた。

この8ヶ所について詳しく見てみると、会話初めのP1「こんにちは」は、日常的な挨拶の状況であり、うなづきというよりもお辞儀に近い。P7「ます」やP10「か」、P17「ください」、P18「ね」、P26「か」、P29「です」は、日本語によく見られる文末表現で、なおかつ発話末になっていた。P24「が」は文の途中ではあるが、多くの被験者がこの直後に一拍置いており、ここでも発話が区切られると考えられる。この8ヶ所は、文末に多く分布しており、前節の予測結果ともほぼ一致している。また、多くの被験者が共通にうなづいていることから、個人差の影響が比較的小さいと考えられる。これより、「話し手のうなづき」を実装する際は、発話文の末および文中でも発話が区切れるような位置にうなづかせるように設計すれば良いと言える。以上を踏まえ、SAYAでは、ここで抽出された8ヶ所にうなづきを表出させるようにした。

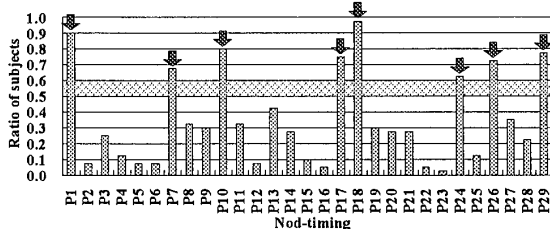


Fig.11 Ratio of subjects who nodded at each nod-timing

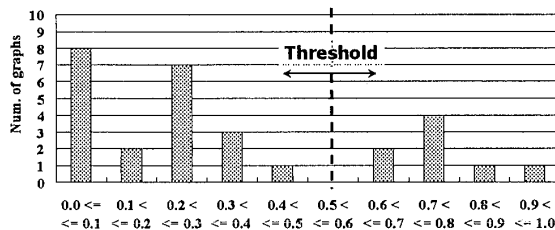


Fig.12 Result of discriminant analysis

4.2 うなづきの角度・角速度に関する実験的検討

4.2.1 実験方法

図9に示したように、人間の場合のうなづき角度・角速度は広範に変化している。しかしながら、ロボットで人間の動きを完全に再現することは難しく、ロボットの動きが自然に感じられる範

囲は人間の場合とは異なる可能性がある。そこで、SAYAのうなづき動作が自然に感じられるような角度・角速度の範囲を実験により検証した。

実験に際して、図9に示すように、人間のうなづき角度・角速度を7等分する $a \sim h$, $a' \sim h'$ を定めた。そして、人間はほとんどの場合、うなづいた分だけ頭を戻していたため、 a と a' を組み合わせで動作Aとし、残りの $b \sim h$, $b' \sim h'$ に関しても同様に組み合わせ、表1に示すようなA($a-a'$) \sim H($h-h'$)の8つのうなづき動作を作成した。また、表1には、SAYAが各うなづき動作を5回試行したときの平均値()内は標準偏差が示されており、概ね目標値を実現できているのが分かる。次に、このA \sim Hのうなづき動作を、発話音声に合わせてそれぞれ実装し、8つの動作パターンを作成した。このとき、4.1で定めた8ヶ所の位置に基づいてうなづきを実装した。実験では、図10の“Receptionist”の位置にSAYAを座らせ、“Visitor”の位置に被験者を立たせた。この状況で、A \sim Hの各パターンで図5中の台詞を順に発話しているSAYAの様子を被験者に順次提示し、『この動作が「人間らしい」、「自然」と感じるか、「はい(感じる)」か「いいえ(感じない)」で解答してください』と指示した。このとき、各パターン間にどのような差異があるかは説明しなかった。また、順序効果の影響を排除するため、被験者の半数にはA \sim Hの順に提示し、残りの半数にはH \sim Aと逆順に提示した。さらに、うなづき動作のみの評価を行うため、SAYAの口の動きや瞬きは各パターンに渡って統一した。このときの口の動きは、音声に合わせて作り込んだ動作を用い、瞬きは2秒に1回の間隔で動作させた。被験者は20代の男子学生21人、女子学生4人の計25人であった。

Table1 Angle and angular velocity of each nod-pattern

Nod-pattern		Angle [deg]		Angular velocity [deg/sec]	
		Target	SAYA	Target	SAYA
A	a	0.50	0.55 (0.03)	3.94	4.13 (0.24)
	a'	-0.50	-0.56 (0.03)	-2.26	-2.80 (0.16)
B	b	1.50	1.51 (0.04)	7.93	8.13 (0.68)
	b'	-1.50	-1.47 (0.11)	-8.25	-7.94 (0.88)
C	c	2.50	2.52 (0.11)	11.93	11.51 (0.70)
	c'	-2.50	-2.43 (0.17)	-14.25	-13.51 (0.61)
D	d	3.50	3.50 (0.11)	15.93	15.50 (0.68)
	d'	-3.50	-3.44 (0.05)	-20.24	-20.59 (0.31)
E	e	4.50	4.50 (0.08)	19.92	19.32 (0.35)
	e'	-4.50	-4.44 (0.14)	-26.23	-26.58 (0.85)
F	f	5.50	5.53 (0.04)	23.92	23.72 (0.16)
	f'	-5.50	-5.41 (0.15)	-32.23	-32.40 (0.92)
G	g	6.50	6.58 (0.05)	27.92	28.23 (0.21)
	g'	-6.50	-6.49 (0.07)	-38.22	-38.86 (0.41)
H	h	7.50	7.48 (0.01)	31.91	32.09 (0.05)
	h'	-7.50	-7.47 (0.10)	-44.22	-44.72 (0.62)

4.2.2 実験結果と考察

A \sim Hの各パターンにおいて、「はい(感じる)」と解答した被験者の数を集計

した結果を図 13 に示す。この結果から、D が最も好まれ、次いで E、C、F、B の順に評価が高く、A、G、H は極端に評価が低かった。

この結果から、8 割以上の高い評価を得た C～E の間では十分に自然な動作が実現できていると考えられる。これに対して、A、B、F、G、H で評価が低かったのは、人間とロボットとの構造や自由度の違いが影響しているからだと考えられる。また、うなづきの大きさは文脈によっても影響されると考えられるが、4・1 において、8 ヶ所のうなづき位置における被験者のうなづき角度は広範に分布しており、関係性は見られなかった(図 14)。図 14 には、参考として被験者 10 人分のうなづき角度の分布の様子を示す。このため、SAYA には、C、D、E の大きさのうなづきを、4・1 で得られた 8 ヶ所にランダムに実装することとした。

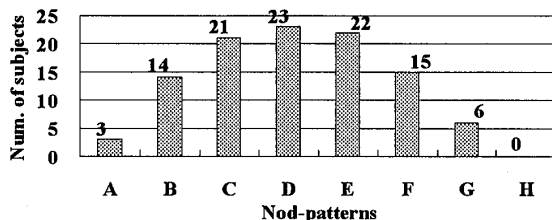


Fig. 13 Evaluation result about humanity and reality of each nod-pattern

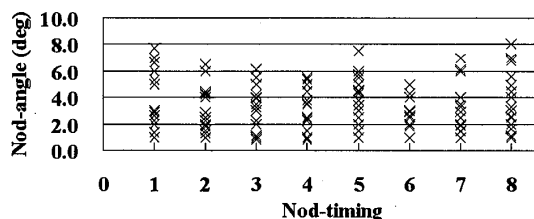


Fig. 14 Variability of nod-angle at each nod-timing

5. 評価実験

SAYA に実装した“話し手のうなづき”に関して、その表出タイミングや角度・角速度が対人コミュニケーションに及ぼす効果を検証する。

5.1 実験 1: うなづきタイミングに関する実験

5.1.1 実験方法 本節では、SAYA がうなづくときのタイミングが、SAYA の人間らしさや対人コミュニケーションに及ぼす影響を検証する。具体的な実験手順は以下の通りである。

[実験条件]

うなづきの表出タイミングの違いによる影響を比較するため、以下に示す 3 条件 (T1～T3) を用意した。

T1: うなづきなし。

T2: 4・1 節で得られた 8 ヶ所以外でランダムにうなづく。ただし、挨拶である「こんにちは」では

必ずうなづき、1 つの台詞内のうなづき回数は T3 と同じ 8 回とした。

T3: 4・1 節で得た、通常、人間がうなづく 8 ヶ所でうなづく。

ここで、T2、T3 のうなづき動作として、4・2 節で評価が高かった C、D、E をランダムに実装した。また、前章と同様、口の動きは音声に合わせて作り込んだ動作を用い、瞬きは 2 秒に 1 回の間隔で動作させた。

[実験環境]

図 10 の“Receptionist”の位置に SAYA を座らせ、“Visitor”の位置に被験者を立たせた。SAYA と被験者の距離は 1.5m 程度とし、音声認識を行うために、被験者の前にマイクを配置した。

[実験手順]

被験者にはあらかじめ、『SAYA に“こんにちは”と話しかけ、〇〇研究室の所在地(または内線番号)を聞いてください』というタスクを与え、図 5 に示すような会話を行ってもらった。各被験者には動作 T1～T3 を実装した SAYA と 1 回ずつ対話してもらい、各動作での対話実験終了後、後に示すアンケートに回答してもらった。このとき、半数の被験者は 1～3 の順に、残り半数の被験者は 3～1 の順に実験を行うことにより順序効果の影響を排除している。

[評価方法]

質問紙を用いた主観評価により評価を行った。ここで、“話し手のうなづき”が対人コミュニケーションに有効に働いた場合に、対話しやすさや親和性を向上させるかどうか、また、その振る舞いや動作が人間らしく自然であるかどうかに関して評価を行った。そこで、人間とロボット等とのコミュニケーションに関する文献⁽⁹⁾⁽¹⁰⁾⁽¹²⁾⁽²⁵⁾で用いられている評価項目を基に、情報伝達、親和性、動作の自然さに関して以下の 8 つの項目を用意した。被験者には、この 8 つの質問に対して-3 から 3 までの 7 段階で評価してもらった。

・情報伝達に関する項目

- (1) 対話しやすい
- (2) 受応えが丁寧
- (3) (伝えたいという)意図を感じる

・SAYA の親和性に関する項目

- (4) 親しみやすい
- (5) 好感が持てる
- (6) 煩わしさを感じない

・振る舞いや動作の自然さに関する項目

- (7) 人間らしい
- (8) 自然

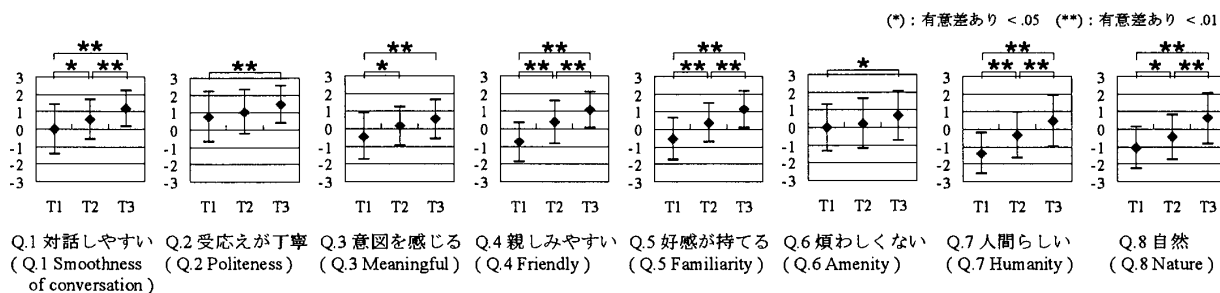


Fig.15 Result of comparison among 3 nod-patterns (T1, T2, T3)

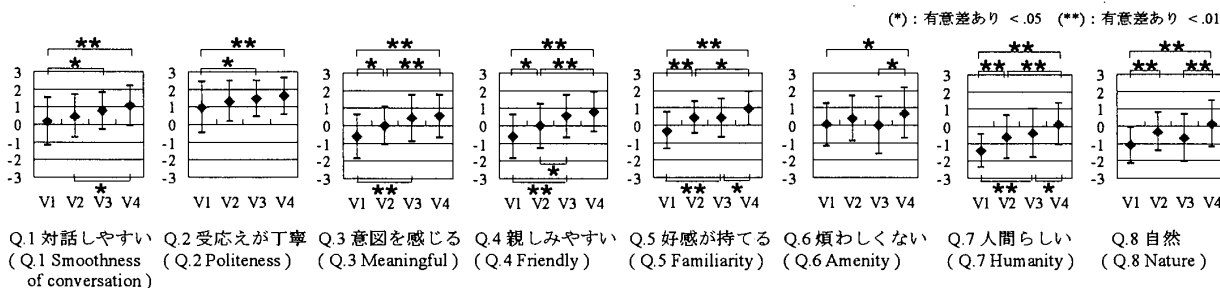


Fig.16 Result of comparison among 4 nod-patterns (V1, V2, V3, V4)

[被験者]

SAYA にあまり馴染の無い大学生および大学院生 50 人(男性 43 人, 女性 7 人)に参加してもらった。

5・1・2 実験結果と考察 評価実験の結果を図 15 に示す。分散分析の結果、全ての質問項目において有意水準 1%または 5%で有意差が認められた。さらに LSD による多重比較を行ったところ、Q.3「意図を感じる」において $T3 > T2 > T1$ なる関係に有意差が見られ、うなづきがあることで話し手の意図が伝わりやすくなることが示された。また、Q.1「対話しやすい」、Q.4「親しみやすい」、Q.5「好感が持てる」、Q.7「人間らしい」、Q.8「自然」の 5 項目において $T3 > T2 > T1$ の関係に有意差が見られ、Q.2「受応えが丁寧」、Q.6「煩わしくない」において $T3 > T1$ の関係に有意差が見られた。これにより、ただ単にうなづかせるのではなく、適切なタイミングでうなづかせることにより、人間らしさや親和感が向上することが示された。さらに、対話のしやすさや受け答えの丁寧さなどにも影響を与えることが確認された。

今回は、被験者に実際に SAYA の発話文を発話してもらうことで、言語情報や感情情報等を含めたうなづきタイミングを定めることができた。今後、相槌のタイミングに関する研究⁽⁷⁾⁽²⁶⁾等を拡張させることによって、発話音声に基づいて自然な表出タイミングを自動で生成することも可能になると考えられる。

5・2 実験 2：うなづきの角度・角速度に関する実験

5・2・1 実験方法 本節では、うなづきの角度・角速度が、SAYA の人間らしさや対人コミュニケーションに及ぼす影響を検証する。実験方法は前節 5・1 とほぼ同じため、以下には変更点のみを示す。

[実験条件]

この実験では、「話し手のうなづき」に関して、うなづきの角度・角速度の違いによる比較を行うため、以下に示す 4 条件 (V1~V4) を用意した。

V1 : うなづきなし。

V2 : 4.2 の A, B をランダムに実装。

V3 : 4.2 の F, G をランダムに実装。

V4 : 4.2 で評価が高かった C, D, E をランダムに実装。V2~V4 でのうなづきのタイミングは、4・1 節で得られた 8 ヶ所とし、口の動きや瞬き、視線の動きの条件は前節 5・1 と同じである。

[実験手順]

各被験者には動作 V1~V4 と 1 回ずつ対話してもらった。このとき、前節と同様、半数の被験者は 1~4 の順に、残り半数の被験者は 4~1 の順に実験を行うことにより順序効果の影響を排除している。

[被験者]

前節 5・1 の被験者とは別に、SAYA にあまり馴染の無い大学生および大学院生 50 人(男性 44 人, 女性 6 人)に協力してもらった。

5.2.2 実験結果と考察 実験の結果を図 16 に示す。前節と同様、分散分析を行った結果、すべての質問項目において有意水準 1% または 5% で有意差が見られた。また、LSD による多重比較を行った結果、Q.1「対話しやすい」、Q.2「受答えが丁寧」、Q.3「意図を感じる」、Q.4「親しみやすい」で、V1 と他の V2, V3, V4 とに有意な差が見られたことから、うなづきがあることで対話しやすさや親和感が向上することが示された。さらに、Q.5「好感が持てる」、Q.7「人間らしい」において $V4 > V2 = V3 > V1$ といった関係に有意差が見られ、適切な範囲内の角度・角速度でうなづかせることにより、SAYA の人間らしさや親和感を向上させることが確認され、対人コミュニケーションにおいて好ましい効果をもたらすことが確認された。また、Q.6「煩わしくない」、Q.8「自然」において、 $V4 > V3$, V1 の関係に有意差が見られたことから、あまり大きくうなづかせると自然さが欠け、相手に煩わしさを与える可能性があると考えられる。

6. まとめ

本研究では、人間の発話中に見られる“話し手のうなづき”に注目し、リアルなうなづき動作を実現するための設計指針を明確にすることを目的とした。そこで、発話中にうなづく場合の好ましい表出タイミング、角度・角速度に関して分析を行った。この結果、発話文の末や、文中であっても発話が区切れるような位置に合わせてうなづかせれば良いことが分かった。また、適切な角度・角速度でうなづかせることにより、SAYA の人間らしさや動作の自然さを向上させることが分かった。これにより、表出タイミング、角度・角速度が、リアルなうなづき動作を実現するための重要な要素であることが確認された。

さらに、SAYA に実装した“話し手のうなづき”に関して評価実験を行ったところ、コミュニケーションにおける情報伝達や親和性の面においても好ましい効果が得られた。これにより、“話し手のうなづき”の対人コミュニケーションにおける有意性が確認された。

人間同士の対面コミュニケーションにおいて、対話相手の非言語行動の僅かな違いに違和感を抱くのと同様に、SAYA のように人間に似た外観を持つロボットの場合でも、非言語行動の表現方法や表出するタイミングの僅かな違いが相手に違和感を与えられられる。そのため、被験者による主観評価において、人間のうなづき特性に基づいた表出タイミング、角度・角

速度でうなづかせた場合に、ランダムにうなづかせたり、うなづき角度を極端に小さく、または大きくした場合と比較して SAYA の人間らしさが向上したと考えられる。

その他、評価実験の自由記述における被験者のコメントとして、「もっと目の動きがあると良い」「口の動きが不自然」「腕の動きがあると良い」「もっと表情が豊かだと良い」といった意見があった。またその一方で、「思っていたより人間らしかった」「うなづきがあることで対話しやすく、好感が持てた」といった意見もあり、更なる技術的、構造的な改善により、人間らしさや親和感を向上させられると期待できる。

本研究では、主にうなづきの表出タイミングや角度・角速度に注目したが、韻律情報や言語情報、または感情情報との関係性も重要だと考えられる⁽²⁶⁾。さらに、うなづき以外の非言語行動、例えば表情や視線、身体は無意識な揺らぎ等と組み合わせて身体全体で協調的な動作を行うことにより、SAYA の人間らしさがさらに向上すると考えられる。今後、さらに人間らしい身体動作や表情表出を実現するために、コミュニケーション中における頭部の動きや表情の変化、また、静止しているときの無意識の身体の揺らぎなどを詳細に解析し、再現していきたいと考えている。

文 献

- (1) Matsuo, T., *Psychology of Communication, Nakanishiya Publishers*, (1999), pp.41-68. (in Japanese)
- (2) Takao, K., *Non-verbal Interface, Ohmsha, Ltd.*, (1994), pp.1 - 99. (in Japanese)
- (3) Watanabe, T., et al., *An Embodied Interaction System Based on Speech by Using InterActor, Journal of Human Interface Society*, vol.2, No.2 (2002), pp.21-29. (in Japanese)
- (4) Morishima, S., *Anthropomorphic Dialog Agent Development Tool Using Facail Image Synthesis and Lip Synchronization, Human-Computer Interaction, Theory and Practice (Part II)*, vol.2 (2003), pp.518-522.
- (5) Matsusaka, Y. et al, *Conversation Robot Participating in Group Conversation, Journal of the Institute of Electronics, Information and Communication Engineers D-II*, Vol.J84-D-II, No.6 (2001), pp.898-9028. (in Japanese)
- (6) Kanda, T. et al, *Embodied Cooperative Behavior for a Humanoid Robot that Communicate with Humans, Journal of the Robotics Society of Japan*, Vol.23, No.7 (2005), pp.808-909. (in Japanese)
- (7) Watanabe, T. et al, *An Embodied Interaction Robots System Based on Speech, Transactions of the Japan Society of Mechanical Engineers, Series C*, Vol.66, No.648 (2000), pp.251-258. (in Japanese)
- (8) Wada, K. et al, *Psychological and Social Effects in Long-Term Experiment of Robot Assisted Activity to Elderly people at a Health Service Facility for the*

- Aged, *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS 2004)*, (2004), pp.3068-3073.
- (9) Ogata, T. and Sugano, S., Emotional Communication between Humans and the Autonomous Robot WAMOEBA-2 (Waseda Amoeba) which has the Emotion Model, *Transactions of the Japan Society of Mechanical Engineers, Series C*, Vol.65, No.633 (1999), pp.1900-1906. (in Japanese)
 - (10) Nakata, T. et al, Generating Familiarity by Robot Behavior toward a Human Being, *Journal of the Robotics Society of Japan*, Vol.15, No.7 (1997), pp.1088-1074. (in Japanese)
 - (11) Matsui, T. et al, A Speech Dialogue System of the Office Mobile Robot, *Jijo-2, Journal of the Robotics Society of Japan*, Vol.18, No.2 (2000), pp.300-307. (in Japanese)
 - (12) Kanda, T., et al, An Evaluation on Interaction between Humans and an Autonomous Robot Robovie, *Journal of the Robotics Society of Japan*, Vol.20, No.3 (2002), pp.93-101. (in Japanese)
 - (13) Kobayashi, H. et al, Study on Face Robot for Active Human Interface, *Journal of the Robotics Society of Japan*, Vol.12, No.1 (1994), pp.155-163. (in Japanese)
 - (14) Ishiguro, H., Android Science -Toward a new cross-interdisciplinary framework-, *Proc. of International Symposium of Robotics Research*, (2005).
 - (15) Hashimoto, T. et al, Development of the Interactive Receptionist System by the Face Robot, *4th JSME Annual Conference on Robotics and Mechatronics*, 2A1-H-45(1)-45(3) (2004). (in Japanese)
 - (16) Maynard, S. K. Anlysis of conversation, *Kuroshio Publishers*, (1993), pp.23-179. (in Japanese)
 - (17) Hashimoto, T. et al, Development of the Face Robot SAYA for Rich Facial Expressions; *SICE-ICASE International Joint Conference 2006*, (2006), pp.5423-5428.
 - (18) J. Cole, *About Face*, MIT Press,(1998).
 - (19) P. Ekman and W.V. Friesen, The Facial Action Coding System, *Consulting Psychologists Press*, (1978).
 - (20) Chou, C.P. and Hannaford, B., Measurement and Modeling of McKibben Pneumatic Artificial Muscle, *IEEE Transactions on Robotics and Automation*, vol.12 (1996), pp.90-102.
 - (21) Schulte H. F., The characteristics of the McKibben artificial muscle, *In The Application of External Power in Prosthetics and Orthotics, National Academy of Sciences-National Research Council*, Publication 874, (1961), pp.94-115.
 - (22) Kapandji, I. A. et al, The Physiology of the Joints Vol.3 -The Trunk and the Vertebral Column-, *Ishiyaku Publishers, Inc.*, (1974), pp.162-245. (in Japanese)
 - (23) Horiuchi, Y., et al, Software Tools for the Creation and Analysis of Dialogue Corpora, *Reports of the Meeting of Special Interest Group on Spoken Language Understanding and Dialogue Processing*, SIG-SLUD-9902-5 (1999), pp.23-28. (in Japanese)
 - (24) Koiso, H., et al, 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 (1998), pp.291-317.
 - (25) Yamamoto, M. and Watanabe, T., Time Lag Effects of Utterance to Communicate Actitons on Human-Robot Greeting Interaction, *Journal of human Interface Society*, Vol.6, No.3 (2004), pp.87-94. (in Japanese)
 - (26) Kitaoka, N. et al, Response Timing Detection Using Prosodic and Linguistic Information for Human-friendly Spoken Dialog System, *Journal of Japanese Society for Artificial Intelligence*, Vol.20, No.3, SP-E (2005), pp.220-22