

ラポールを誘発する会話エージェントの言動の検討

Designing the Behaviors of a Conversational Agent who can Stimulate Rapport of the Users

澁澤 紗優美^{1*} 黄 宏軒² 林 勇吾² 川越 恭二²
Sayumi Shibusawa¹ Hung-Hsuan Huang² Yugo Hayashi² Kyoji Kawagoe²

¹ 立命館大学大学院 情報理工学研究科

¹ Graduate School of Information Science and Engineering, Ritsumeikan University

² 立命館大学情報理工学部

² College of Information Science and Engineering, Ritsumeikan University

Abstract: The elderly who live alone are increasing rapidly in these years. For their mental health, it is useful to maintain their social life with others. This project is aiming to develop a companion agent who can engage long-term relationship with the elderly users. This paper presents our first step in the work to explore the rapport occurred in human-human communication which is considered to be essential in keeping reliable relationship with others. We analyzed the corpus collected in human-human experiment from three view points, the speaker (potential user), the listener, and the third person. Encouraging results that may provide the hints of agent development are found in the analysis: the mood of the speaker can be potentially observed by another person, and the attitude of the conversation can have influence on the speaker's mood.

1 はじめに

近年の急速な高齢化により、一人暮らしの高齢者が増加している。一人暮らしの高齢者は、地域での孤立化が顕著なため [1]、高齢者が信頼関係を築けるパートナーが、コミュニケーションを継続的に取ることが重要である。対話相手として、介護事業所のホームヘルパーが挙げられるが、対応できる時間には限りがある。学術的な支援の例として、酒井ら [4] による認知患者のための語りかけエージェントの研究がある。また、気軽に対話を行える方法として、Skype で行う傾聴ボランティアもあるが、高齢者と信頼関係を形成することは簡単ではない。

我々は、エージェントの態度を話し手にとって好ましくすることで、高齢者と信頼関係を構築できないかと考えている。このような、親身な言動から生じる信頼関係を「ラポール」と定義する。エージェントのラポールを誘発する言動は、以下のような仕組みで実現できると考えている。

前提として (1) 聞き手の態度が話し手の気分に影響を与える (2) 話し手の気分は話し手の態度に現れる (3) 話し手の態度の変化を第三者が検知できる、という3つの仮説が正しいと仮定する。エージェントの態度が話し

手の気分に影響を与えれば、話し手の気分は話し手の態度に表れる。話し手の態度をエージェントが検知し、話し手の気分に応じた適切な言動でラポールを誘発する仕組みを考えている。

仮説の検証方法として、会話データを収集し、会話時の態度の印象を互いに評価したものを分析するものがある。本研究は高齢者支援が目的のため、高齢者同士のデータを収集して分析すべきだが、高齢者には認知機能が低下している者もいるため、互いの態度の印象評価を行うことは困難である。そこで、会話に参加していない第三者が会話時の高齢者の態度の評価をすることで、態度評価（自己、会話相手）と同様の評価を得られないかと考えた。

本稿では、上記の仕組みでエージェントを開発することが可能かを検証する目的で、学生同士の会話データの収集び会話時の態度の評価（自己、対話相手）を行った。また、第三者による態度評価が会話者による態度評価と比べて差があるか検証するために、第三者も会話者の態度評価を行った。第三者とは、話し手でも聞き手でもなく会話に不参加だった者を指す。エージェントは将来的に、Skype で行う傾聴ボランティアの代替になりうると考え、会話は Skype で行った。会話データを収集し分析することで、エージェントの適切な言動の指針を検討する。

*連絡先：立命館大学大学院情報理工学研究科
〒525-8577 滋賀県草津市野路東1丁目1-1
E-mail: is021083@ed.ritsueki.ac.jp

2 関連研究

高齢者の心的ケアに焦点を当てたエージェントの分野では, Bickmore ら [7] が提案した Hospital Buddy がある. 病院に入院している患者, 特に高齢者の不安や孤独, 退屈といった心理的負担の軽減を目的とし, 患者の話し相手となるエージェントであり, 今後の進展が期待される.

コミュニケーションロボットの分野では, 加納ら [2] の Babyloid があり, 社会福祉法人の個室に設置し動作させた結果, 抑うつ度を軽減できる可能性が示唆されている. 小林ら [3] は, 卓上インタフェースロボット ApriPoco を対話インタフェースとして用い, 対話コミュニケーションを通じて高齢者の生活を豊かにすることを目指している. ロボットとの対話はコミュニケーションの活性化につながることを確認されている. しかし, Babyloid と ApriPoco は会話相手の感情を考慮していない.

また, Huang ら [6] が研究している Rapport Agent は Tichle-Degnen ら [5] が提案した三要素(明るさ, 自然な注視, 協調)に基づいた会話エージェントを実装している. エージェントと人間の二者間の対話において, 人間の発話の有無, 視線, 頷き, 笑顔等をリアルタイムに分析してエージェントの反応をパターン化している. 相槌, 話者交代等の動作にいくつかのパターンを設定している. しかし, 幅広い人間との間にラポールを形成することを目指しているため, 対象が広い. その為, 高齢者に特化しているとはいえない.

我々は, 感情や気分の変化を考慮した上で, 高齢者との会話に特化したラポールを誘発するエージェントを作成を目指す. 将来的には, Skype の傾聴ボランティアの代替となるような性質を持たせたいと考えている.

3 コーパス収集実験

3.1 実験目的

ラポールを誘発する会話エージェントの態度モデルの検討と, 第三者による評価の有効性の確認のために, コーパス収集実験を行った. ラポールの誘発にあたり, 我々は気分と態度が重要と考える. なお, 本稿では気分と態度を以下のように定義する.

気分: ある状況下で生じる, 内面の感情・心情

態度: 気分から生じる言動・表情・動作で, 外から観察できるもの

ラポールを誘発する為には, 相手の気分を心地よくする必要があり. 対話において, 話し手が得る情報は, 主に聞き手の言動・表情・動作であるため, これら聞き手の態度の情報が話し手の気分及びラポールの形成に大

きな影響を与えていると考えた. そこで我々は, 以下の3つの仮説を立てた.

- (1) 聞き手の態度が話し手の気分に影響を与える
- (2) 話し手の気分が話し手の態度に影響を与える
- (3) 話し手の態度の変化を第三者が検知できる

上記に挙げた3つの仮説が正しければ, 聞き手の態度が話し手の気分に影響を与え, 話し手の気分が態度に影響を与え, 態度の変化を第三者が検知できることを示せる. 第三者の役割をエージェントに置き換えれば, エージェントが話し手の態度の変化を検知し, 話し手の気分を心地よくするような態度をとれば, 話し手の気分が心地よくなることを示すことができる. この仮説の条件下であれば, 聞き手の態度の分析が, ラポールを誘発する言動の分析につながると考える. 3つの仮説を検証するためには, 会話時の話し手の態度の自己評価及び話し手が聞き手の態度を評価したデータと, 第三者が話し手の態度を評価したデータを収集し, 分析する必要がある. エージェントによる話し手の態度の検知に関しては, 表情や身体の深度情報を用いて判断できると考えている.

次に, 第三者による評価の有効性の確認について述べる. ラポールを誘発する会話エージェントは, 将来的には認知症高齢者の傾聴を目指している. そのため, 先述した3つの仮説を検証するためには, 高齢者同士の対話における互いの言動の印象の評価及び気分の評価を収集すべきである. しかし, 認知症等を患う高齢者は, 会話時の態度の印象の評価を行うことが困難である. そのため, 高齢者の態度の評価(自己, 会話相手)を, 第三者による評価で代替できないかと考えている.

3.2 実験方法

実験は2段階に分けて行った. まず2者の会話を録画した. その後, 録画した会話データを使い, 態度の自己評価, 他者評価(会話相手, 第三者)を行った.

3.3 会話

(1) 実験環境

実験は, 同性ペアの学生4組(男性3組, 女性1組)を対象に行った. 年齢は21~24歳. 平均年齢は22.1歳である. 別々の部屋で, Skypeを通じて対話を行い, 対話時の被験者の言動をビデオカメラで録画した. また, 赤外線式の深度センサーとマイクロホンアレイが内蔵されている Kinect センサー¹を用いて被験者の上半身の深度情報を取得した. 実験環境の配置を図1, 図2に示し, 実験

¹<http://www.xbox.com/ja-JP/KINECT>

の風景を図 3, 図 4 にて示す。それぞれの部屋には 2 台のビデオカメラがあり, 1 台は話し手及び聞き手の映像を直接録画し, もう 1 台は Skype に映る相手の映像を撮る。将来エージェントと話し手が会話する状況を推定すると, 話し手と聞き手が同じ部屋で向かい合って話す状況と比較して, 映像を見て対話する Skype 会話の方が想定している状況を再現できると考えたため, 会話相手の Skype 映像を録画した。また, Kinect で得た深度情報と会話時の言動のタイミングを正確に把握する為, Kinect の映像も録画した。被験者の下半身の動きを制限するために, 動かしにくいパイプ椅子を使用した。話し手の部屋では, デスクトップ PC の片方は Kinect と接続し, もう片方は Skype 用に使用している。また, 大型スクリーンで等身大に聞き手役の被験者を写し, 目線の高さもできる限り合わせた。聞き手の部屋では, 実験環境の制約上, スクリーンではなく通常のディスプレイを用いたが, 目線の高さは話し手の映像と合うように調整した。

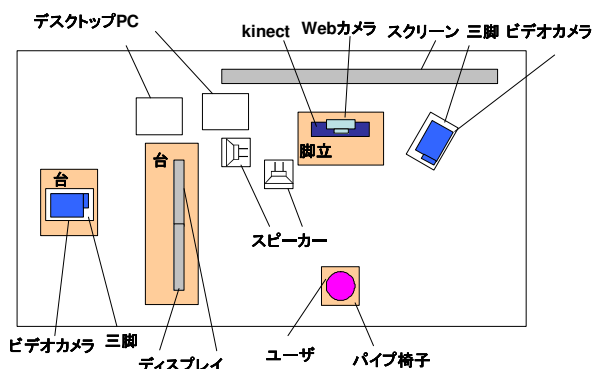


図 1: 話し手役の被験者がいる部屋

(2) 対話方法

1 セッションあたりの時間は酒井らの研究の設定 [4] を参考に, 被験者が十分に実験環境に馴染んで落ち着き, 有効に会話の様子が観察でき, 実験も合理的な時間内で終わると考えられる 7 分間に設定した。1 組につき 4 セッションの会話を行い, 被験者は, セッション毎に話し手と聞き手を交互に担当した。話し手は話題を主導する役割を担い, 聞き手は相槌, 質問等で話し手をフォローする役割を担う。話題は, 高齢者及び学生双方が話しやすいものであり, 話の及ぶ範囲をある程度限定するため, 「家族」を題材にした。また, 気分がよくなる場合, 悪くなる場合双方のデータを収集する

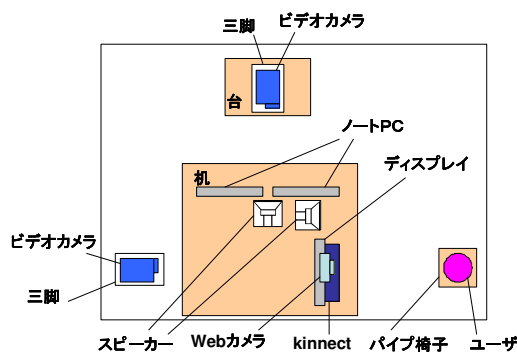


図 2: 聞き手役の被験者がいる部屋

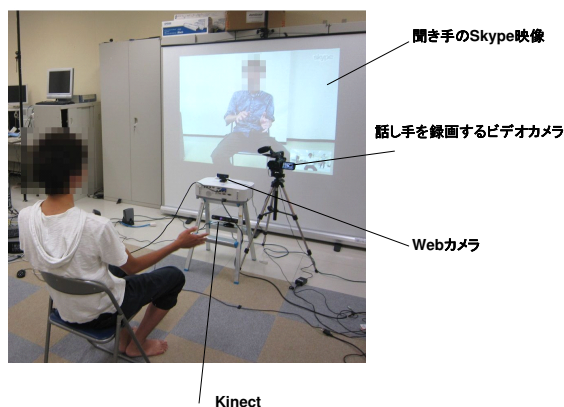


図 3: 話し手役の被験者がいる部屋

ため, 最初の 2 セッションは「家族のいい話」とし, 次の 2 セッションは「家族の悪い話」とした。セッション毎の被験者の役割を表 1 に示す。

3.4 言動の評価

4 セッションの会話終了後, 被験者は互いの態度の印象評価をラベリングした。加えて, 被験者の対話に参加していない第三者も, 後日話し手及び聞き手の態度の印象評価のラベリングを行った。ラベリングには ELAN²を用いた (図 5)。エージェントが動作を検知する際はタイムラグが発生するため, 聞き手の映像は話し手の部屋の Web カメラから取得した Skype の映像を用いた。ラベルの名称及び値の詳細は次項で述べる。

ラベルの開始位置を限定するために, 対話動画の音声

²<http://tla.mpi.nl/tools/tla-tools/elan/>

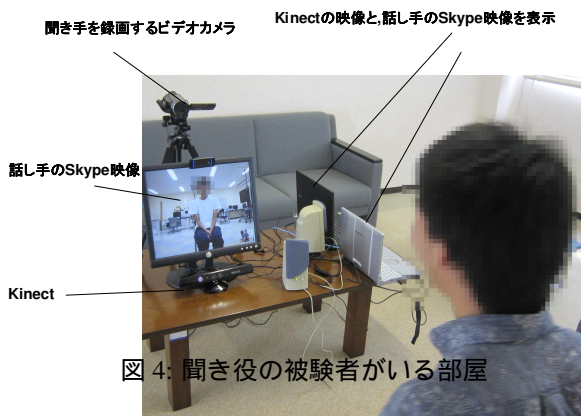


図 4: 聞き役の被験者がいる部屋

表 1: 各セッション (S) における被験者毎の役割

| 話題 | S | 話し手 | 聞き手 |
|--------|---|-------|-------|
| 家族のいい話 | 1 | 被験者 a | 被験者 b |
| | 2 | 被験者 b | 被験者 a |
| 家族の悪い話 | 3 | 被験者 a | 被験者 b |
| | 4 | 被験者 b | 被験者 a |

の発話区間を音声分析ツール Praat³で区切った。ラベリングのルールは以下の 4 つを被験者全員に教示した。

1. ラベルは連続してつける
2. ラベルの開始・終了時刻は Praat の区間にあわせる
3. 複数の発話区間を含んだラベリングも可能
4. 1 つのラベルは最大 10 秒以内

ラベルの名称及び値は、表 2 に準ずる。評価の値は、1(同意しない、全く好ましくない) から 7 まで(とてもそう思う、とても好ましい)の 7 段階である。以下、値の教示について述べる。GoodFeeling は快い、心地よい、親密さを強く感じる、相手が話を注意深く聞いている、相手が共感してくれている、自分の話に夢中になっていると思えるような時には高い値をつけるように教示した。逆に不愉快な気分、落ち込んだ気持ちになった際は低いとした。AttPrtnr は、相手が自分の話についてよくフォローしてくれていると感じる時や、笑顔且つ相槌もしくは質問を投げかけた際に高いとした。相手が自分の話に対して不快感を抱いているように感じた時や、不快感をあらわにした表情で違う意見を述べてきたような時は低いとした。AttMe は、相手の話を適切にフォローしたと感じた時や、笑顔と相槌、質問や打ち解けたような態度をとれば、高い値を与えるように教示した。相手が著しく不快に感じるような発言をしたと感じた場合は、低い値とした。AttSpk と AttLtn は、話し手と聞き手の態度がそれぞれ相手にとって好ましいと思われるかで判断した。笑顔や相槌、質問等を頻繁に行う場合は高く、長時

³<http://www.fon.hum.uva.nl/praat/>



図 5: ELAN 画面

間の沈黙や不愉快さを思わせるような態度を取った場合は、低い値とした。

表 2: ラベルの定義。該当者はそのラベルをつける被験者

| ラベル名 | 説明 | 値 | 該当者 |
|-------------|------------------|---------------|-----|
| Silence | 発話区間の区切り | Sound Silence | なし |
| 10s | 10 秒単位の区切り | なし | なし |
| GoodFeeling | 自分の気分 | 1-7 | 話し手 |
| AttPrtnr | 相手の今の態度が好きか | 1-7 | 話し手 |
| | 自分は今相手にいい態度を取ったか | | 聞き手 |
| AttMe | 自分は今相手にいい態度を取ったか | 1-7 | 話し手 |
| AttSpk | 話し手の態度 | 1-7 | 第三者 |
| AttLtn | 聞き手の態度 | 1-7 | 第三者 |

4 データ分析

4.1 評価間の関係性の分析

聞き手の態度が話し手の気分に影響を及ぼす、という仮説を証明するために、(1) 話し手の気分と、話し手が評価した聞き手の態度 (2) 話し手の態度の自己評価と話し手の気分 (3) 話し手をつけた自分の態度と、第三者がつけた話し手の態度の関係の分析を行った。分析方法は、ラベルの値の相関(ピアソンの積率相関係数)と、ラベルの値の差分の相関についての分析を行った。値については、値を 0.2 秒単位で標準化した 2 本のラベル (1-7) の相関を算出した。差分については、値を 0.2 秒単位で標準化した後、1 本のラベルの値が直前の時間単位から増減があった場合、値の変化分のデータを相関の算出対象とした。被験者同士の言動の 1 つの区切りが、0.6 秒程度であったため、片方のラベルの値が直前の時間単位が

ら増減した場合において、もう片方のラベルの±0.6秒以内の時間単位に値の増減があるかを参照した。双方のラベルで変化が無かった時間単位を除き、相関を算出した。例えば、GoodFeelingとAttPrtnrの相関を算出する場合において、GoodFeelingが会話開始から10~15秒は4,15~18秒までは5の場合、15秒目で4から5に値が増加している。15秒目に1増えた時に、AttPrtnrの14.4~15.6秒で値の増減を参照し、相関の対象を取る値の対象とする。

なお、分析対象のラベルは、4組の被験者(女性1組、男性3組)がラベリングしたものをを用いた。相関の算出に用いたラベルは、GoodFeeling、話し手が評価したAttPrtnr、話し手が評価したAttMe、AttSpkの4つであり、それぞれ約7分の長さである。それぞれの被験者の4セッションの全てのラベルの総数と、相関の算出に用いたラベルの総数を表3にあらわす。

表3: ラベル付けの結果。「Total1」と「Total2」はそれぞれラベルの総数と相関分析に用いたラベルの数を示す

| 被験者ペア | Total1 | Total2 |
|-------|--------|--------|
| A | 1745 | 1207 |
| B | 1536 | 949 |
| C | 1350 | 967 |
| D | 1472 | 1107 |

仮説1:話し手の気分と話し手がつけた聞き手の態度
話し手が評価したGoodFeelingと、話し手が評価したAttPrtnrの2本のラベルを標準化した値と差分について、相関を算出した。

仮説2:話し手の態度の自己評価と話し手の気分
話し手が評価したGoodFeelingと、話し手のAttMeの2本のラベルの値を標準化し、値についての相関を算出した。

仮説3:話し手の態度の自己評価と第三者評価
話し手が評価したAttMeと、第三者が評価したAttSpkの本のラベルの値を標準化し、値についての相関を算出した。

4.2 分析結果

前節の仮説1、仮説2、仮説3を検証するための相関の計算結果を表4にて示す。GoodFeelingと話し手が評価したAttPrtnrの相関及びGoodFeelingと話し手が評価したAttMeの相関は高く、仮説1と仮説2は支持されることが示唆された。話し手が評価したAttMeとAttSpkの相関は高いとはいえず、仮説3は支持されなかった。その要因として、評価の基準値が話し手と第三

者間で若干の差異があることが考えられる。同じ時間単位において、話し手と第三者が互いに4と評価していた場合でも、その値が示す態度の印象は異なる可能性がある。話し手と第三者の値の基準の差異を何らかの形で補正できれば、相関が向上する可能性がある。

表4: 相関分析の結果。「S」はセッション番号を示す。「SM/SL」、「SM/SS」、「SS/3S」はそれぞれGoodFeelingと話し手が評価したAttPrtnrの相関、GoodFeelingと話し手が評価したAttMeの相関、話し手が評価したAttMeとAttSpkの相関を示す

| S | SM/SL | | SM/SS | SS/3S |
|----|-------|------|-------|-------|
| | 値 | 差分 | 値 | 値 |
| A1 | 0.9 | 0.7 | 0.91 | 0.41 |
| A2 | 0.75 | 0.72 | 0.75 | 0.07 |
| A3 | -0.07 | 0.41 | -0.17 | -0.07 |
| A4 | 0.74 | 0.63 | 0.61 | 0.29 |
| B1 | 0.6 | 0.37 | 0.45 | 0.26 |
| B2 | 0.61 | 0.42 | 0.74 | 0.19 |
| B3 | 0.54 | 0.3 | 0.52 | -0.06 |
| B4 | 0.61 | 0.81 | 0.59 | 0.33 |
| C1 | 0.36 | 0.49 | 0.61 | 0.25 |
| C2 | 0.04 | 0.19 | 0.26 | 0.4 |
| C3 | 0.5 | 0.37 | 0.52 | 0.27 |
| C4 | 0.43 | 0.44 | 0.55 | -0.14 |
| D1 | -0.01 | 0.38 | 0.18 | 0.25 |
| D2 | 0.6 | 0.66 | 0.46 | 0.01 |
| D3 | -0.04 | 0.31 | 0.04 | 0.15 |
| D4 | 0.48 | 0.67 | 0.51 | 0.37 |

4.3 事例

話し手の態度の自己評価と、第三者が評価した話し手の態度の相関が低かったが、第三者が話し手の身振りから話し手の態度の変化を検知していた事例があったため、報告する。以下、話し手をS、聞き手をLとする。

事例1:手を叩く

S:お祝いをしてあげることがなかったから

L:うん

S ないと思ってたから

S: (手を叩いて) ほんまうれしかって、それ、めっちゃなんか、(手を叩いて) 感動してん

Sの「ほんまうれしかって」の発話時にSSが6から7に、3Sが4から5に上昇した。手を叩く動作は、気分の高揚を示唆する。

事例2:勢いよく太ももを叩く

S:俺の車っていうか

L:うん

S:(自分に手をむけ)俺と、(画面のほうに手を向け)じいちゃんの(手を左に動かし、勢いよく太ももを叩いて)兼用やねん

L:あー

Sの「兼用やねん」の終了時にSSが4から5へ、3Sが4から5に上昇した。音を鳴らす動作、もしくは勢いよく何かを動かす動作は、気分の変化を示唆する。

これらの結果から、身振り等の非言語情報と話し手の気分の関係をより詳細に分析することで、身振りと気分の関連性を導き出せば、第三者やエージェントでも話し手の気分の変化を検知できる可能性がある。

5 おわりに

本稿では、ラポールを誘発するエージェントの言動を検討するに当たり、Skypeの対話を通じてコーパス収集を行った。また、高齢者の態度の評価(自己、会話相手)を、第三者による評価で代替できないか検証するため、対話時の態度の自己評価、会話相手による評価を行った。データを分析した結果、話し手による聞き手の態度の評価と話し手の気分には相関があることがわかった。また、話し手の気分と話し手の態度にも相関があったため、話し手の気分は態度に何らかの影響を与えている可能性がある。第三者による話し手の態度評価と、話し手による態度の自己評価の相関は低かったが、話し手と第三者の値の基準の差異を何らかの形で補正できれば、相関を向上できる可能性がある。第三者と話し手の態度評価の相関を近づけることが出来れば、第三者が話し手の態度を検知できるようになり、ラポールを誘発する会話エージェントは実現可能と考えられる。

参考文献

- [1] 平成 24 年度版高齢社会白書, 内閣府, 2012
- [2] 加納 政芳, 種田 行男, 清水 太郎, 岸 太一, 井原 一成, 清水 優: Babyloid と高齢者の共生から見えてきたもの, 第 25 回人工知能学会全国大会, 2011
- [3] 小林 優佳, 山本 大介, 田崎 豪, 山地 雄士, 土井 美和子: 高齢者向け対話インタフェース 病院スタッフ・患者間の対話モデルを利用した音声対話ロボット, 第 26 回人工知能学会全国大会, 2012
- [4] 酒井 洋一, 野中 裕子, 安田 清, 中野 有紀子. 認知症患者と語りかけエージェントとのインタラクション分析とシステム構築に向けた提案. ヒューマンインタフェース学会研究報告集. vol. 13, no. 8, 2011
- [5] Tickle-Degnen, L., Rosenthal, R.: The Nature of Rapport and its Nonverbal Correlates. *Psychological Inquiry* 1(4), 285-293 (1990)
- [6] Lixing Huang, Louis-Philippe Morency, Jonathan Gratch.: "Virtual Rapport 2.0" ,In 10th International Conference on Intelligent Virtual Agents, Reykjavik, Iceland, pp.68-79(2011)
- [7] Timothy Bickmore and Laila Bukhari and Laura Pfeifer Vardoulakis and Michael Paasche-Orlow and Christopher Shanahan.: "Hospital Buddy: A Persistent Emotional Support Companion Agent for Hospital Patients", in 12th Conference on Intelligent Virtual Agents, Santa Cruz, CA, USA, pp.12-14(2012)