

Artificial Subtle Expressions: 人工物のための直感的情報伝達手法の提案

Artificial Subtle Expressions: Intuitive Notification Methodology of Artifacts

小松孝徳¹ 山田誠二² 小林一樹¹ 船越孝太郎³ 中野幹生³

Takanori Komatsu¹, Seiji Yamada², Kazuki Kobayashi¹, Kotaro Funakoshi³, and Mikio Nakano³

¹ 信州大学

¹Shinshu University

² 国立情報学研究所/総合研究大学院大学

² National Institute of Informatics/SOKENDAI

³ ホンダ・リサーチ・インスティテュート・ジャパン

³ Honda Research Institute Japan

Abstract: We describe artificial subtle expressions (ASEs) as artifact-like expressions for creating smooth interaction with users based on our assumption that the artifacts should not produce human-like expressions but artifact-like ones, because it required little budget and manpower. We prepared two types of audio ASEs; one is a flat artificial sound, and the other is a decreasing sound. These two ASEs were played after a robot made a suggestion to the users. We then conducted a simple experiment to observe whether the participants accepted or rejected the robot's suggestion in terms of the two ASEs. The results showed that they accepted the robot's suggestion when the flat ASE was used, while they rejected it when they rejected it when the decreasing ASE was used. Therefore, we found that the ASEs were intuitively interpreted by the participants without their learning its meaning explicitly and with little cognitive load.

はじめに

近年、日常生活における様々なタスクを支援するために、ロボットエージェントや擬人化エージェントなどの様々な種類の人工物が開発されている。これらの人工物は、人間のような外観を持ち、人間のような情報表現を表出するようにデザインされていることが多い。その理由として、一般的に「人間とインタラクションを行う人工物は、人間と同じような外観や情報表現をするべきだ」と考えられているからであり、実際に人工物に対して「人間らしさ」を実現するための研究活動が盛んに行われている[1-3]。例えば、Sugiyama [1]らは、状況を自ら判断することでユーザに対して適切なジェスチャを表出するヒューマノイドロボットを開発し、Kipp and Gebhard [2] はユーザの視線方向に応じて自らの視線方向を調節するアバターエージェントを開発している。

しかしながら、これら「人間らしさ」を実現しようとする研究アプローチには、少なくとも二つの大

きな課題が存在していると筆者らは考えている。一つは、「人間らしさ」を追求するために要する労力とコストの問題である。ロボットエージェントであっても擬人化エージェントであっても、これらの人工物に対して「人間らしさ」実現するためには、多自由度におよぶ関節を使ったスムーズな動作、細かな顔の表情、自然な発話機能などを実装する必要がある。これらの開発には多大な労力とコストが必要であることは想像に難くない。またもう一つの課題は、設計者が追求した「人間らしさ」は、必ずしもユーザにとっても同じように「人間らしく」感じられるとは限らないという問題である。例えば、人工物に「人間らしさ」を実装しようとした場合、その人間らしさがある段階に達すると、ユーザに対して不気味さを与えてしまうという「不気味の谷」[4]問題がその代表的な指摘である。また筆者らのうち、小松・山田 [5]は「適応ギャップ」問題を報告し、その中で、ユーザは人間らしい人工物に対しては過度な期待をしてしまい、その人工物が予想を裏切るような行動を見せてしまうと、それらに対して失望感を抱

いてしまうことを報告している。

このような「人間らしさ」を追求することで生じる課題に対して筆者らは、人工物は「人間らしい」情報表現を表出するべきではなく、「人工物らしい」情報表現を表出するべきだという立場をとる。そのような「人工物らしい」情報は、非常にシンプルな形式の情報だと想定されるため、多大な労力やコストを費やすことなく様々な人工物に対して容易に実装が可能だと考えられる。筆者らは、このような情報表現を ASE (Artificial Subtle Expressions) として提案する。具体的にこの ASE は、以下のような四つの要件を満たす必要があると考えている。

- **補完的**：ASE は人工物とユーザとのインタラクションにおいてあくまでも補完的な役割をするものであり、インタラクションにおける主たるプロトコルを阻害するものではない。つまり、インタラクションの文脈と切り離れた ASE 自体には特定の意味は存在しない。
- **直感的**：ASE の解釈には、明示的な学習は必要なく、直感的に理解される。つまり、ASE の解釈には多大な認知的負荷は必要ない。
- **シンプル**：ASE の実装に要する労力やコストが非常に低い。そのため、ASE は単一のモダリティで構成され、例えばビープ音のような人工音や LED の明滅パターンのようなものが想定される。
- **正確**：特定の意味を正確にユーザに伝えることができる。ASE がユーザに伝える情報としては、人工物の内部状態を想定しており、人間同士のコミュニケーションにおける非言語情報に似た役割を果たす。

ASE が実装されるモダリティとしては、音声情報および視覚情報を想定しているが、本稿では音声情報として実装された ASE を扱うこととする。このような音声情報としての ASE に関する関連研究としては、Blattner らの「earcon」[7] および Gaver の「auditory icon」[6,8] が挙げられよう。これらはコンピュータのデスクトップ上における特定のユーザの操作に対応した人工音を鳴らすことにより、ユーザに直感的な操作感を提供するものである（例。ファイルをごみ箱に移動すると、お皿が割れたような人工音が表出される。）。これらの earcon および auditory icon も「シンプルな人工音でユーザに直感的な理解を促す」という意味では ASE と非常に近いコンセプトを有しているが、これらはインタラクションにおける主たるプロトコルを担っており、インタラクションにおける補完的な役割を担う ASE とは根本的に異なるものであるといえる。

本稿の目的は、ユーザに情報を提示するロボット

に ASE を実装し、ユーザとのインタラクションにおいて ASE が効率的に作用しているのかを実験的に確認するものである。ここで、ASE が明示的な学習や多大な認知的負荷をユーザに与えずにその特定の意味をユーザに直感的に伝達できた場合、ASE の有効性が確認されたとみなすことができる。なお本研究においては、ASE（「人工物らしい」情報表現）と「人間らしい」情報表現との比較は扱わない。なぜなら、ASE が一定の有効性を示せば、その実装に要する労力とコストを鑑みても、「人間らしい」情報表現よりも ASE の応用可能性の方が高いと考えられるからである。

実験

実験設定

人工物とユーザとがインタラクションを行う実験環境として「宝探しゲーム」を使用した（図 2）。このゲームでは、画面の奥に向かって道を進んでいくと、三つの盛り土が現れる。この三つの盛り土のうち一つには金貨が隠されており、その金貨をできるだけ多く集めることが、実験参加者のタスクとして与えられる。三つの盛り土のいずれかに隠された金貨の位置は、ランダムに設定される。三つの盛り土はゴール到達まで合計 20 回現れ、スタートからゴールまでの所要時間はおよび三分程度である。実験参加者には金貨一枚につき 50 円に相当する文具が謝礼として渡されることを事前に説明する。

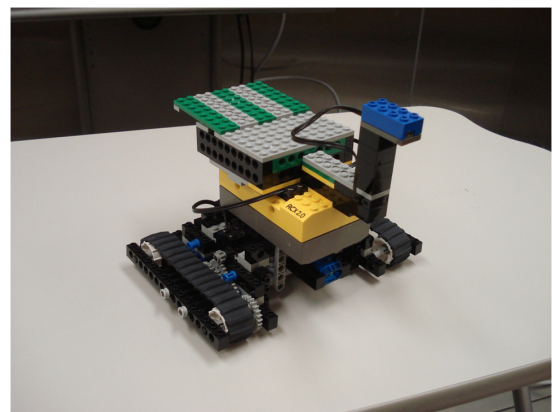


図 1：MindStorms ロボット

ゲーム中に画面が三つの盛り土の前に到着すると、実験参加者の横に置かれた人工物（MindStorms ロボット：LEGO 社製、図 1）から、その人工物が予想した金貨が隠された盛り土の位置が人工音声にて参加者に伝えられる。この場合、参加者はこの人工物の予想通りの盛り土を選んで、予想と違うものを

選んでもよいと教示される。また参加者は、自分が選択した盛り土に金貨が隠されていたのか否かというフィードバックを受けられない設定になっており、最終的な金貨の獲得数を実験後にのみ知ることができる。具体的には、実験中の参加者がある盛り土を選択しても、その盛り土の上にクエスチョンマークが表示されるだけである（図 2）。



図 2：宝探しゲーム

使用した ASE

本実験では、人工物が人工音声を表出した直後に ASE を表出した。人工物は、金貨の隠された盛り土の予想位置を日本語の人工音声で参加者に伝える

(例。「いちばん」「にばん」「さんばん」)。これらの人工音声はクリエイティブシステム社製 Document Talker によって作成された。そして、これらの人工音声の 0.2 秒後に、人工音である ASE を表出した（図 3）。ASE は二種類用意され、一つは提示時間 0.5 秒の三角波で基本周波数が 400 [Hz] で一定のもの（flat ASE）、もう一つは提示時間 0.5 秒の三角波で基本周波数が 400 [Hz] から 250 [Hz] に減少するものである（decreasing ASE, 図 4）。これらの人工音はアドビ社製の Cool Edit 2000 で作成された。このことから、ここで作成した ASE は非常に簡便な作業で作成されるために、**シンプル**であるという ASE の要件は満たされている。筆者らのうち Komatsu and Yamada [9] は、周波数が減少するような人工音はユーザにとってネガティブな印象として解釈されるという知見を実験的に獲得しているために、特に decreasing ASE は人工物のネガティブな内部状態（態度）を伝達すると想定した。本実験での具体的な ASE の解釈としては、人工物が予想した「金貨の隠された盛り土の位置」についての確信度が伝達されると想定している。

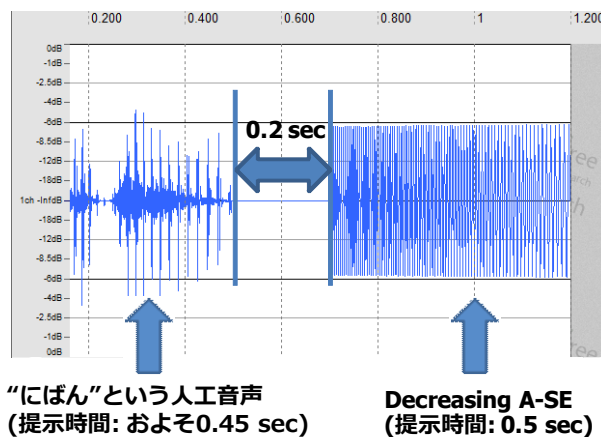


図 3：人工音声と ASE の関係

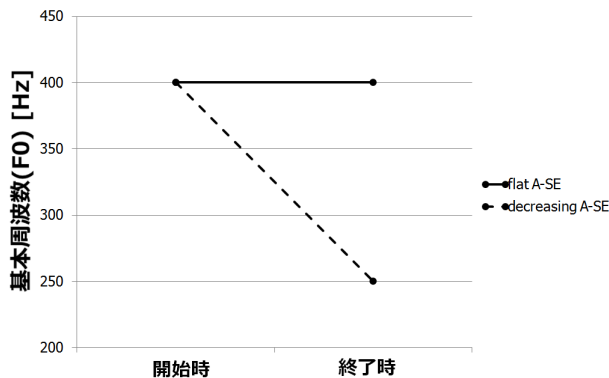


図 4：flat ASE と decreasing ASE

本実験における人工物と実験参加者との間における主たるプロトコルは、人工物の予想した盛り土の位置を参加者に伝えることであり、補完的なプロトコルはその際の人工物の内部状態（確信度）を伝えることである。人工物の人工音声の後に ASE を付与することにより、ASE がメインのプロトコルを阻害することがないために、**補完的**であるという ASE の要件は満たされているといえる。つまり本実験で使用する ASE は四つの要件のうち、シンプル、補完的であるという二つの要件をすでに満たしているために、**直感的**および**正確**であるという残りの二つの要件が満たされているのかを実験的に確認することが必要となる。

実験手順

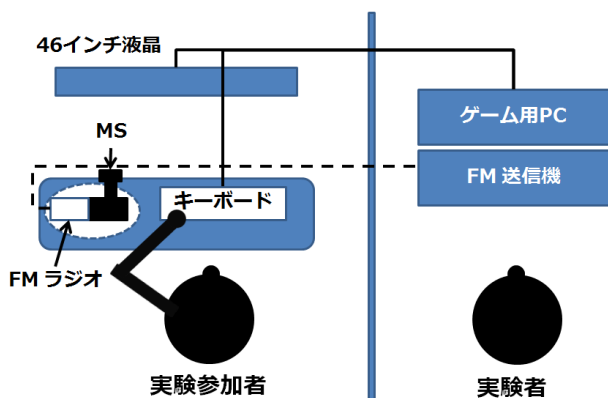


図5：実験設定

実験には19人の大学生が参加した（男性10名、女性9名；22～25才）。宝探しゲームは、実験参加者の前方に配置された46インチ液晶ディスプレイに提示され、金貨の隠された盛り土の位置を予想する人工物は参加者の左斜め前方に配置された（図5および図6）。人工物と参加者の距離はおよそ50cmであり、人工物から表出される音声の音圧は、参加者の頭部付近で50 dB (FAST, A)であった。人工物から表出される人工音声およびASEは、隣室に待機した実験者からWoZ方式にて表出される。

実験開始前に、実験者は実験参加者に対して実験の目的および概要を伝えるが、その際、人工音声の後に表出されるASEに関しては説明も言及も行わなかった。20回の試行のうち、人工物はflat ASEを10回、decreasing ASEを10回表出するが、それらの提示順に関しては、被験者間でカウンタバランスが取られるように配慮された。

人工物は金貨の隠された盛り土の位置を20試行全てに渡って正確に参加者に対して伝えるが、参加

者は実験終了後までその事実を知らされなかった。

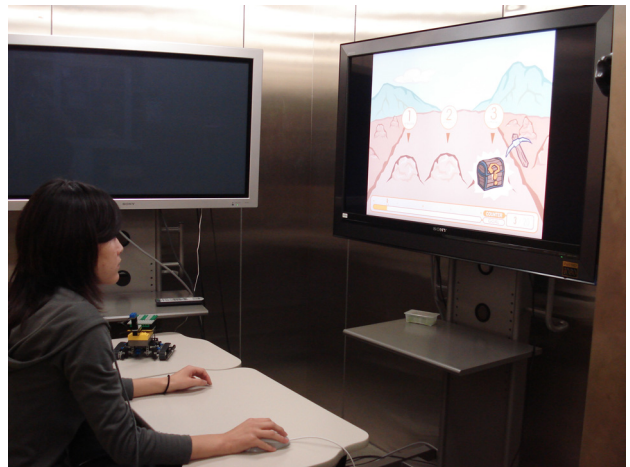


図6：実験風景

本実験の目的は、参加者が人工物の予想を受け入れたのか否かという行動と表出されたASEとの関係を観察することである。また実験終了後に、簡単なインタビューを行い、ASEの存在に気づいていたのか、もしそうならば、それをどのように解釈したのかについての聞き取り調査を行った。図2に示したようなゲームを採用することにより、参加者は自分の選択した盛り土に金貨が隠されていたか否かを知ることができないため、「参加者が人工物の予想に従うのか否か」という現象は、「参加者がASEをどのように解釈しているのか」という点に強く反映されると想定した。

筆者らは以下のような実験結果が得られたならば、ASEの直感的および正確であるという要件を満たすものと考えた：flat ASEが付与された人工音声による人工物の予想は参加者に受け入れられ、decreasing ASEが付与された人工物の予想は参加者に拒絶される。

結果

ASEが実験参加者の行動にどのような影響を与えていたのかを調査するために、実験において参加者が人工物の予想を拒絶した回数（拒絶数）を計測した。この拒絶数は、20試行における10回のflat ASE提示の場合および10回のdecreasing ASE提示の場合ごとに計測された。その結果、flat ASEが付与された予想に対する平均拒絶数は10回中1.73回（SD=1.51）、decreasing ASEに対する平均拒絶数は10回中4.58回（SD=2.43）であった（図7）。一要因二水準（要因：ASEの種類、水準：flat/decreasing、従属変数：拒絶数）の被験者内計画の分散分析の結果、

flat ASE と decreasing ASE が付与された予想に対する拒絶数に統計的有意差が観察された ($F(1,18)=13.38, p<.01 (**)$). つまり, decreasing ASE が付与された人工音声による予想は, 実験参加者によって有意に拒絶される確率が高いことが明らかになった. よって, 本実験においては ASE の役割や存在そのものについて実験者は参加者に対して何一つ言及していないにもかかわらず, ASE の理解が直感的かつ正確に行われていたことを確認することができた. ここから, 前節で予想したような実験結果 (flat ASE が付与された人工音声による人工物の予想は参加者に受け入れられ, decreasing ASE が付与された人工物の予想は参加者に拒絶される.) が確認され, 直感的, 正確であるという ASE の要件が満たされているとみなすことができた.

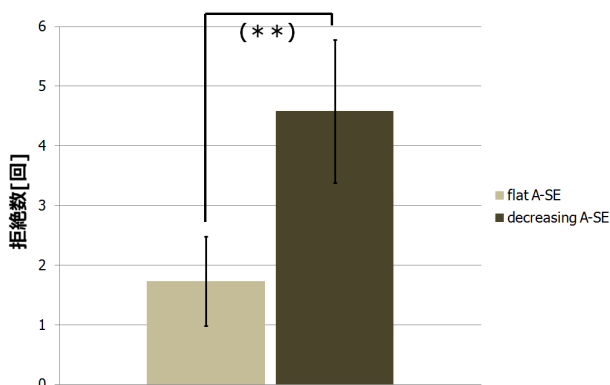


図7: 全参加者 19 人における flat/decreasing ASE に対する拒絶数

実験終了後のインタビュー調査においては, 19 人の実験参加者のうち 5 人が, 実験開始直後に ASE の存在に気づき, その意味を直感的に理解したうえで人工物の予測を受け入れるか否かを決定していたと回答した. その際, 「音程が下がる音 (decreasing ASE) が聞こえると, このロボットは予想に自信がないのかなあと感じた」という感想をこの 5 人から得ることができた. このことから, 参加者にとって ASE の解釈には明示的な学習が必要ではないにもかかわらず, 直感的かつ正確にその意図を理解することが可能なことが確認されたといえる.

その一方, この 5 人以外の残りの 14 人は, 実験終了後においても ASE の存在に気づけなかったと回答していた. ここで, これら 14 人の flat ASE と decreasing ASE に対する拒絶数を比較することで, ASE に対する解釈が無意識的に行われているのか否かを確認できる. この比較においても, decreasing ASE を付与した人工音声の方が, 有意に拒絶数が多かった場合, ASE はユーザの認知的負荷をほとんど必要

せずに解釈されていると強く主張できると考えられる.

これら 14 人の実験参加者において, flat ASE を付与した人工物の予測に対する拒絶数は 2.28 回 ($SD=1.73$), decreasing ASE に対する拒絶数は 3.43 回 ($SD=1.59$)であった (図 8). 一要因二水準 (要因: ASE の種類, 水準: flat/decreasing, 従属変数: 拒絶数) の被験者内計画の分散分析の結果, flat ASE と decreasing ASE の拒絶数に統計的有意差が観察された ($F(1,13)=4.98, p<.05 (*)$). つまり実験中に ASE の存在に気付かなかった参加者においても, decreasing ASE が付与された人工音声による予想は, 参加者によって有意に拒絶される確率が高いことが明らかになった. よってこの実験結果より, ASE の解釈には明示的な説明や多大な認知的負荷は必要ではなく, これらの解釈はむしろ無意識的に行われている可能性が示唆された.

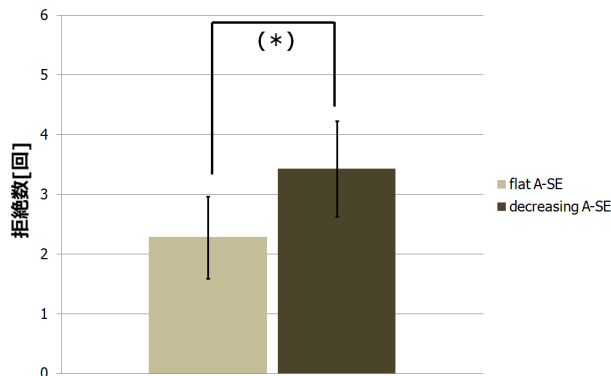


図8: ASE の存在に気付かなかった参加者 15 人における flat/decreasing ASE に対する拒絶数

よって, これらの実験結果より, ASE は実験参加者にとって直感的にかつ正確にその意味が解釈されているために, ASE の要件のうち残されていた二つの要件 (直感的, 正確) を十分に満たしているといえ, ASE の効果が実証されたといえる.

おわりに

本研究では, 「人工物は人工物らしい情報表現をするべきだ」という筆者らの立場に基づいて, 人工物が表出すべき情報表現 ASE を提案し, 実際に ASE がユーザに与える影響を調査する実験を行った. 具体的には, 人工物が表出した情報に対する確信度を示す情報手段として二種類の ASE (flat ASE/decreasing ASE) を設計し, 人工物の表出する人工音声に続いて二種類の ASE を表出した. その際, ASE の設計においては, シンプルかつ補完的である

ように留意した。そして、これらの ASE が表現すべき人工物自信の確信度が、ユーザに対して直感的かつ正確に伝達されていたのかを実験的に分析した。その結果、人工物の予想音声に対して **decreasing ASE** を付与した場合には、参加者はその予想を有意に高い確率で拒絶していることが明らかになった。さらに ASE の存在に気付かなかった参加者であっても、無意識的に ASE に対する正しい解釈が行われている可能性が示唆されたために、ASE は直感的かつ正確に人工物の内部状態を伝える手段として有効だということが確認された。特に、ASE の存在や意味を参加者に全く説明していないにも関わらず、さらには ASE の存在に気づいていないにも関わらず、人工物の内部状態が直感的かつ正確に伝達できていることが確認され、このことは ASE の有効性を非常に明確に示しているものと考えられる。よって、提案した ASE を様々な人工物に実装することで、ユーザとのインタラクションを円滑に構築できる可能性が示唆されたといえよう。

今後はこの人工音による ASE をより現実的なアプリケーションに実装して、その効果を検証することを計画している。具体的には、カーナビゲーションシステムへの導入を検討中である。現在のカーナビゲーションシステムは、その精度が年々向上しているものの、ユーザにとって満足な情報を常に提示できるとは限らないのが現状である。そこで、ナビゲーションシステム自身にとってそれほど確信度が高くない情報を出表する場合には、本実験で利用した **decreasing ASE** のような情報を付与することで、ユーザが直感的かつ認知的負荷を感じずにその意図を理解することで、確信度が高くない情報を出表したシステムに対してストレスを感じることなく操作を続けられるとも考えられる。このように本実験よりもさらに現実的な状況にて ASE の有効性を検証することで、より広範かつ様々な状況における **Human-Agent Interaction** 研究に ASE の応用が可能となると筆者らは期待している。

参考文献

- [1] Sugiyama, O., Kanda, T., Imai, M., Ishiguro, H., Hagita, N. and Anzai, Y. Humanlike conversation with gestures and verbal cues based on a three-layer attention-drawing model. *Connection Science*, 18, 4 (2006), 379-402.
- [2] Kipp, M. and Gebhard, P. IGaze: Studying reactive gaze behavior in semi-immersive human-avatar interactions, In *Proc. the 8th International Conference on Intelligent Virtual Agents (2008)*, 191-199.
- [3] Liu, K. and Picard, W. R. Subtle expressivity in a robotic computer. In *Proc. CHI2003 Workshop on Subtle Expressivity for Characters and Robots (2003)*, 1-5.
- [4] 森政弘. 不気味の谷: 人型ロボットデザインへの注意, 『ロボコンマガジン』, 28 (2003), 49-51.
- [5] 小松孝徳・山田誠二. 適応ギャップがユーザのエージェントに対する印象変化に与える影響, 『人工知能学会論文誌』, 24, 2 (2009), 232-240.
- [6] Gaver, W. W. *Auditory Interfaces. Handbook of Human-Computer Interaction*, Elsevier Science, 1997.
- [7] Blattner, M. M., Sumikawa, D. A. and Greenberg, R. M. Earcons and Icons: Their Structure and Common Design Principles. *SIGCHI Bull.* 21, 1 (1989), 123-124.
- [8] Gaver, W. W. The SonicFinder: An Interface That Uses Auditory Icons. *Human-Computer Interaction* 4, 1 (1989), 67-94.
- [9] Komatsu, T., and Yamada, S. Effect of agent appearance on people's impressions of agents' attitudes *Ext. Abstracts CHI-2008*, ACM Press (2008), 2929-2934.