

感情表現手段としての非言語音に関する実験報告

和氣早苗* 池邊亮志** 河野泉+
上窪真一+ 福住伸一**
s-wake@bp.jp.nec.com

NEC 情報通信メディア研究本部*, NEC コンシューマ PC 事業部**,
NEC インターネットソフトウェア開発本部+,
〒630-0261 奈良県生駒市高山町 8916-47

近年、擬人化エージェントやペットロボットが情報機器のインタフェースとして、またコミュニケーションのパートナーとして注目されている。本論文では、擬人化エージェントの感情を表現するメディアとしての非言語音の有効性を評価した。音声対話型の情報検索エージェントシステムである EDS (Emotional Dialog System) を対象として擬人化エージェントの感情を表す 11 種類の感情表現音を制作し実装した。その上で、音の有無、音の数、発音位置の違いによって、感情の感じられやすさ、ユーザの好みがどのように変化するかを実験的に検証した。その結果、非言語音は感情表現のメディアとして有効となり得るものの、使い方によってはユーザに不快感を与えたり、他メディアの表現を阻害するものにもなりかねないため注意が必要であることがわかった。

A study of using nonverbal sound for expressing emotional information

Sanae Wake*, Ryouji Ikebe**, Izumi Kohno+,
Shin'ichi Uwakubo+, Shin'ichi Fukuzumi**

Computer & Communication Media Research*, Consumer PC Division**,
Internet Software Development Division+,
NEC Corporation

Recently, human-like characters and entertainment robots are paid attention as a new type of computer interface and human's partner. In this paper, we study the effectiveness of nonverbal sound for expressing emotional information. Using EDS (Emotional Dialog System) as a platform, we designed eleven nonverbal sounds to express emotion of the human-like character. Then we made an experiment to clear whether users can feel emotion through the sounds, and whether users like it. As a result, nonverbal sound can be effective for expressing emotional information, but it have to be paid attention to be use because sounds tends to be too persistent and obstruct another media.

1. コミュニケーションにおける非言語音

警告音, 操作反応音, 通知音などは「報知音」と総称され, これらは「論理情報」を直感的かつ明示的に聴取者に伝達する非言語音 (nonverbal sound) である. 我々はこれまでにこの報知音を対象とし, 設計手法や有効性に関する研究を行ってきた^[1]. 一方, 実世界において非言語音は論理情報の伝達以外にも様々に用いられる. その一つに「感性情報」の伝達があり, 鳴き声, 感嘆詞などの非言語音が感情の表現に利用される. 本研究ではこの「感情表現」手段としての非言語音に關しての考察を行う.

生物間のコミュニケーションでは, 感情を交換する「感情コミュニケーション」が重要である^[2]. 感情コミュニケーションを電子機器において実現しようとする研究も増えつつあり^[3,4], 特にペットロボットや擬人化キャラクタを利用したインタフェースは昨今注目を集めている^[5]. ここではロボットやキャラクタとのインタラクションにおいてユーザに感情を意識させることが目的の一つともなっている.

現実の感情コミュニケーションは動作, 表情, 発話の抑揚などで実現される. 音という観点からこの感情コミュニケーションを捉えると, 生物は感情を鳴き声などの音で表現する. 人間も, 感嘆詞や笑い声, 泣き声など言語以外の音声を利用する. このような, 感情を表現する非言語音を本論文では「感情表現音」と呼ぶ.

映画やアニメなどでは, 演出技法の一つとして感情表現音が積極的に利用されている. 言葉を話さないキャラクタが音で登場人物とコミュニケーションするという手法はもはや一般的である. 今回このような非言語音による感情コミュニケーションを仮想エージェントとユーザの間に実装し, 非言語音の感情表現手段としての有効性, 効果的な利用方法について実験を通して考察する.

実験の遂行にあたっては, 音声対話型の情報検索エージェントシステムである EDS (Emotional Dialog System) を利用した. EDS では擬人化エージェントの感情がユーザとの対話に応じて変化し, この感情に準じて感情表現音が発音される.

2. EDS - Emotional Dialog System -

研究のプラットフォームとなる EDS を説明する.

2. 1 EDS の目的と概要

EDS は車への搭載を前提とした音声対話型の情報検索エージェントである. 車内空間をより楽しく豊かにすることを目的とし, ユーザの嗜好や入力に応じて提案や反応を行う「有能かつ人間味のある秘書エージェント」の実現を目指す. 図 1 に EDS の画面例を, 図 2 に音声対話の例を示す. ユーザは, キャラクタとして表現される擬人化エージェントとの音声対話を通して BGM 選択を行うが, ここでユーザの感情認識およびエージェントの感情表現を利用して, 感情を含んだインタラクションの実現する点が EDS の特徴である.

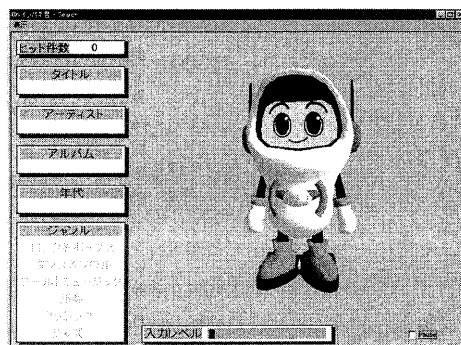


図 1. EDS インタフェース画面

また HCI (Human-Computer Interaction) の観点から, 感情的なインタラクションがタスクやインタラクションそのものにいかに影響するかを明らかにすることも EDS 研究の目的である. EDS の詳細は同研究会に報告の「感情表現を用いた対話システム EDS の開発(1)および(2)」を参照されたい.

A: 「…邦楽はどうですか? (提案文)」
U: 「いいねいいね (ユーザ応答)」
A: 「♪気に入ってもらえてなによりです(反応文).それならオ フコースのYesNoがお勧めです.どうですか? (提案文)」
U: 「ダメダメ, 全然ダメ (ユーザ応答)」
A: 「♪ばっちりだと思ったのに. (反応文). それなら…」

(♪マークは感情表現音の発音位置) A=Agent, U=User

図2. EDS対話例

2. 2 擬人化エージェントの感情

2. 2. 1 感情モデル

EDSの擬人化エージェントは11種類の反応感情を持つ。反応感情の生成モデルを表1に示す。音声対話は図2に示すように、“提案文(エージェント) → ユーザ応答・指示 → 反応文(エージェント)”の繰り返して進行するが、ここで擬人化エージェントの反応感情は、提案感情(提案時のエージェントの自信度)と、その提案に対する利用者感情(肯定, 否定とその程度)に応じて決定される。例えば、エージェントは自信満々で提案した内容を否定されると慌てる、というようなことである。

表1. 反応感情の生成モデル

利用者感情 提案感情	否定 興奮	否定 落ち着いた	中立	肯定 落ち着いた	肯定 興奮
自信度 大	慌てる	不平	平穩	満足	誇り
自信度 中	落胆	落胆	平穩	喜び	喜び
自信度 小	悲しい	諦め	平穩	安堵	驚き

2. 2. 2 感情表現のメディア

擬人化エージェントの反応感情は以下の4種類で表現される。

(1) 表情 (視覚メディア)

各感情につき1種類の表情。

(2) 動作 (視覚メディア)

各感情につき1種類の動作。

(3) せりふ (聴覚メディア)

反応感情の生成モデル(表1)の各セルに割り当てられる2種類以上のせりふよりランダムに出力される。せりふには録音音声を利用した。

例: 誇り・まかせといてよ, そうでしょうとも!

落胆・考えたのに…, こんなはずじゃなかったのに。

(4) 感情表現音 (聴覚メディア)

各感情につき1種類の感情表現音。せりふの直前に発音される(例: 「ピロツ(感情表現音). かなり考えたのに. ではジャズはいかがですか?」)。感情表現音には、自然音や楽器音の録音ではなく、波形編集を行った加工合成音を利用した。加工合成音はより創作の自由度が高く、感情を誇張して表現することが可能だからである。音質的に自然音よりも合成音の方が、キャラクタの容姿に合うと考えたことも一因である。

3. 感情表現音の制作

擬人化エージェントの感情を表す感情表現音をサウンドデザイナーとの共同作業で制作した。その制作プロセスを以下に示す(図3)。

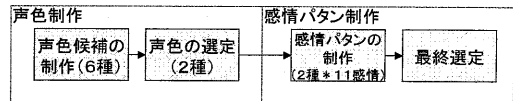


図3. 感情表現音デザインのプロセス

3. 1 声色制作

3. 1. 1 声色候補の制作

以下に示す声色制作上の注意項目に準じて、まず6種類の声色が制作された。

- 長くとも2秒までの音。
- 機械的イメージの声色と、生物的イメージの声色を候補として制作。
- 比較のため、各声色について「普通」「喜ぶ」「落胆」の3感情の感情表現音制作。
- 日本人しかわからない表現は用いない。

3. 1. 2 声色の選択 (2種類)

6種類の声色の中から2種類(機械的イメージ1種類, 生物的イメージ1種類)の声色を, 以下の観点から選択した.

- 音が感情を表すと感じられること
- 容姿との違和感が少ないこと
- 制作した, 普通, 喜ぶ, 落胆の感情の差がわかりやすいこと.

3. 2 感情パタン制作

3. 2. 1 感情パタン制作

選出した2種類の声色について, 表1に示す11種類の感情表現音を制作した. 制作にあたっての各感情の説明項目と例(項目4以外)を表2に示す.

表2. 感情の説明

	(1) 発音の条件 (2) 言葉で表現したときの例 (3) 注意事項 (他感情との比較など) (4) 表情および動作データ
喜ぶ	(1) 誉められたとき, すんなりと推薦が受け入れられたとき (2) 「気に入ってもらえてなによりです」 (3) 軽い感じの喜び. 「満足」「誇り」の方がプラス方向に強い感情.
落胆	(1) 提案した内容 (音楽, 商品情報など) が否定されたとき (2) 「いいとおもったのになあ. . .」

3. 2. 2 最終選定

制作した2セットの感情表現音を, 表情, 動作, せりふと組み合わせて出力し, いずれの声色がよいかを主観的に評価した. 評価の観点は, 前記の声色選定時と同じとした. その結果, 感情がよりわかりやすく, キャラクタの容姿とも違和感のない生物的イメージの音を EDS の感情表現音と選定した. 擬音で例を示すと, ピュルツピュルツ, ピロツという類の音である.

4. 実験

制作した感情表現音を利用し, 感情表現における非言語音の有効性について実験を行った.

4. 1 実験の目的と評価観点

実験の目的は, 感情表現音の有効性を評価することにある. EDS のインタラクションにおいて, 感情表現音の有無, 利用する感情表現音の種類数, 発音位置の違いによって, 感情の表れやすさ(ユーザにとってのわかりやすさ), ユーザの好みがどのように変化するかを検証する【実験1】. また感情表現における非言語音の評価が, 視覚メディアの有無で変化するかどうかを調べる実験も行った【実験2】. 評価の観点は, 実験1, 2共に, 感情の感じられやすさと, ユーザの嗜好である.

4. 2 実験1: 視覚メディア利用時の感情表現音評価

擬人化キャラクタの感情を, 聴覚メディア(感情表現音, せりふ)と視覚メディア(表情, 動作)を使って表現する状況で, 感情表現音の使い方を4パターンに変化させ, 感情の感じられやすさ, 使い心地がどのように異なるかを比較評価した.

4. 2. 1 被験者

被験者は7名(男性5名, 女性2名). 学生6名と研究員1名で, 年齢はいずれも20代であった. 全て, 初めてEDSシステムを利用する者である.

4. 2. 2 比較実験用システム

「せりふ」「表情」「動作」は共通で, 「感情表現音」を表3に従って変化させた4種類のシステムを準備した.

表3. 感情表現音の使い方

感情表現音	有無	種類	出力位置*
A (11種類)	有	11種類	前
B (1種類)	有	1種類※	前
C (音なし)	無	—	—
D (位置違い)	有	11種類	後

※1種類の感情表現音は感情「普通」音を利用.

*出力位置とはせりふと感情表現音の出力順序

【前】は「感情表現音+反応文+提案文」の順.

【後】は「反応文+感情表現音+提案文」の順を表す.

4. 2. 3 実験方法

実験の手順を説明する。

(1)説明： EDS の目的と実験の概要を説明。但し、比較実験用の4つのシステムの違いは被験者には伝えない。アンケート用紙は予め読んでもらう。

(2)システムの使用： 20 分間、4 種類のシステムを自由に使ってもらふ。時間内であれば、同じシステムをどのような順序で何度使ってもよい。

(3)アンケート回答： 4 種類のシステムを比較して、以下に示すアンケート項目に回答頂く。

- Q 1. キャラクタの感情が感じられた順を記す
- Q 2. 使いたい (好きな) システムの順を記す
- Q 3. 各システムの違いがわかったか
- Q 4. 自由コメント

4. 2. 4 結果

実験 1 の結果を示す。まず、Q 1 「感情の表れ」および Q 2 「嗜好」についての結果の集計データを図 4、図 5 に示す。

【集計方法】個々の評価結果について、評価の高いシステム順に、8、6、4、2 点を付与し、システム毎の合計値を算出した。評価が同位のシステムについては、それぞれに中間の得点を与えた (A=B>C>D の場合、A、B には 7 点を付与)。

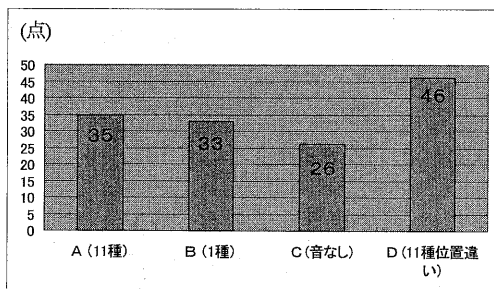


図 4. 実験 1：感情の感じられやすさ評価結果

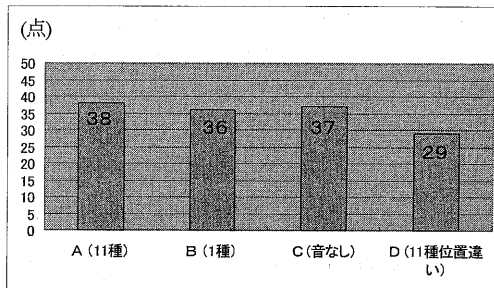


図 5. 実験 1：使いたいシステム評価結果

Q 3 の 4 種類のシステムの違いに関する結果を示す。

- ・ 4 種類の違い全てを正しく認識 … 0 名
- ・ D の音の使い方が他と違うことを認識 … 4 名
- ・ C に音が無いことを認識 … 2 名
- ・ 音の有無や違いに全く気づかない … 2 名

A=11 種類, B=1 種類, C=音なし, D=位置違い

Q 4 で得られた代表的なコメントを示す。

- ・ A[11 種類]が最も感情的だと感じた。
- ・ A[11 種類]は次にキャラクタがどうなるかわかる気がした。
- ・ C[音なし]は仕草だけで、いやらしさが無い。
- ・ D[位置違い]は自己主張が強い。それほど強い主張は必要ない。
- ・ D[位置違い]は効果音がきつく耳障り。
- ・ D[位置違い]は感情があからさますぎる。
- ・ 各システム感情が違うと感じた。
- ・ ABC は大差ない。特に BC は余り印象がない

4. 2. 5 結果のまとめ

実験 1 の結果より以下の傾向が読み取れる。

【感情の感じられやすさ】

- ・ 音は (たとえ 1 種類でも) ないよりもある方が感情は感じられやすい。
- ・ D[位置違い] が最も感情が感じられやすい。

【ユーザの好み】

- ・ D[位置違い]はわざとらしく受け入れられ難い。

各実験システム毎に結果をまとめる。

●Aシステム [11種類の感情表現音]

全体的によい評価であった。“感情の感じられやすさ”では評価が2番目，“ユーザの好み”も有意差はないものの数値上は最もよい評価が得られた。コメントも「最も感情的」等があり、少なくとも否定的な結果や意見はない。11種類の感情表現音利用を明確に認識した被験者はいないが、無意識にその多様性を感じたのではないかと思われる。

●Bシステム [1種類の感情表現音]

最も特徴のないシステムと評価された。評価は“感情の感じられやすさ”“ユーザの好み”共に3番目。生データを参照しても、最もわかりやすいと評価した者も、最も使ってみたいと答えた者もない。またコメントもBに関するものはない。

●Cシステム [感情表現音なし]

“ユーザの好み”の評価は悪くないが，“感情の感じられやすさ”評価は最も低い。音がないことへの否定的意見はないものの、肯定的意見もない。結果からは、感情表現という観点からは、音はあの方がよいといえる。

●Dシステム [11種類の感情表現音がせりふの後に発音]

最も感情が感じられると評価されたが，“ユーザの好み”は最も評価が低い。音が耳につき耳障りというコメントも多く、せりふを聞くことに集中した直後に音が聞こえるため良くも悪くも音の印象が強くなったと考察される。

4.3 実験2:聴覚メディアのみ利用時の感情表現音評価

擬人化キャラクターの感情を、ディスプレイを利用せず聴覚メディア（感情表現音、せりふ）のみを使って表現する状況で、感情表現音の使い方を2パターンに変化させ、感情の感じられやすさ、好みがどう異なるかを比較評価する。

4.3.1 被験者

被験者は5名(男性4名、女性1名)。研究員4名、

事務員1名で、年齢はいずれも20代であった。被験者は実験1の者とは異なり、すべて初めてEDSシステムを利用する者である。

4.3.2 比較実験用システム

実験1で利用したA[11種類の感情表現音]、C[感情表現音なし]の2種類のシステムを比較刺激とした。

4.3.3 方法

AとCの2種類のシステムについて比較実験を行った。実験手順は4.2.3項に示すものと同じである(但しシステムの使用時間は10分間)。アンケート項目も実験1と同じである。

4.3.4 結果

Q1「感情の表れ」およびQ2「嗜好」の結果に関しての集計データを図6、図7に示す。

【集計方法】各被験者の比較評価結果について、評価の高いシステムに2点、低いシステムに0点、評価が同位の場合にはそれぞれに1点を付与し、各システムの合計値を算出した。

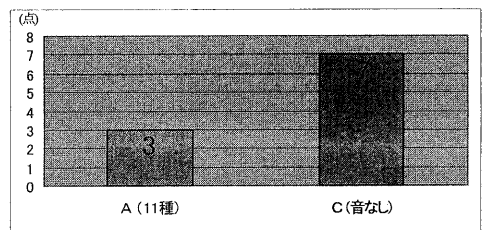


図6. 実験2:感情の感じられやすさ評価結果

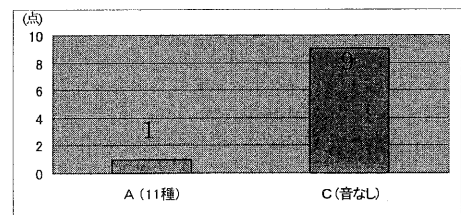


図7. 実験2:使いたいシステム評価結果

Q3のシステムの違いに関する結果を示す。

- ・音の有無を認識 … 3名(5人中)
- ・音の違いに気づかない … 2名

A=11種類, C=音なし

Q4における代表的なコメントを示す。

【A [11種類] に関して】

- ・Aは自分勝手な印象
- ・音はコンピュータっぽくていや。ない方がいい
- ・感情がわかりやすいが、反面反応が単純過ぎる
- ・感情の起伏が激しくわかりすぎて好みでない

【C [音なし] に関して】

- ・音声認識されたかどうか間が空くので寂しい

4. 3. 5 結果のまとめ

集計結果とヒアリングより以下の結果を得た。

【感情の感じられやすさ】

- ・感情はC [音なし]の方が感じられやすい。
- ・ただし中にA[11種類]の方に強く感情を感じるものもいることがコメントからわかる。

【ユーザの好み】

- ・C [音なし]の方が受け入れられる。

音の有無を認識した被験者の割合は実験1より多かった。注意が聴覚メディアのみに集中するからだと考えられる。視覚メディアがなくなること、実験1の結果(音がある方が感情が感じられやすい)がより強まること予想したが、予想に反して逆の結果(音はない方が感情がわかりやすく、また受け入れられやすい)との結果となった。

4. 4 考察 - 感情表現音の効果と利用メディア

実験1-視聴覚メディアで感情表現を行う実験-によると、感情表現音を利用する方が感情が感じられやすいとの結果が得られた。一方、実験2-聴覚メディア(せりふと感情表現音)のみで感情表現を行う実験-では、感情表現音はない方が感情がわかりやすく、また音のない方を利用

したいとの結果が得られた。これは、特に感情の表れやすさに関して、実験1の結果と矛盾している。この理由を以下に考察する。

一つには、聴覚メディアのみの実験環境では、感情表現音がエージェントの発する音だと認識されにくかったことが理由にあると類推できる。実験2では擬人化エージェントの容姿は提示されず、キャラクターは被験者の想像に任される状況であった。その上で、実験で使われた「せりふ」には録音音声が発用されることで、被験者の多くは人間に近いキャラクターを想像し、人間の音声とは音質の異なる感情表現音をキャラクターが発する音だと直感的に認識することが困難であったことが考えられる。その結果、感情表現音は対話と関係のないノイズ、あるいは対話とは異質なシステム音として捉えられ、音はない方がよいという評価になったのではないかと推測できる。

これより得られる知見は、同一キャラクターについて複数種類の音(せりふと感情表現音など)を利用する場合には、その音色の親和性に充分配慮する必要がある、ということである。動作や口の動きという視覚メディアの仲立ちがある場合は、ユーザは音とキャラクターの関係を理解しやすい。しかし、仲立ちがない場合に音質が異なる音が利用されると、それらはキャラクターと無関係の音と認識される可能性がある。動作との関連も含め、せりふの声色との相性を考慮して、感情表現音をデザインする必要がある。

二つめの原因としては、聴覚メディアのみの環境ではユーザの注意が録音音声である「せりふ」に集中することが原因と考えられる。

先に述べたように、実験で使われた「せりふ」は感情のこもった録音音声であった。EDSのインタラクションは音声対話によって進行するためせりふを聞かずにのインタラクションはあり得ないが、せりふが抑揚を含むことでユーザは

「せりふ」に表れる感情を聞取することに注意を払ってしまったと推測できる。つまり、視覚メディアがないことでユーザの注意力の分散が少なくなり、その分ユーザは、情報量が多く聞き慣れた人間の音声に全ての情報を読み取ろうと、意識を集中させたのではないだろうか。その結果、感情表現音はその作業を阻害するものと認識されてしまった可能性がある。言い換えると、注意深く音声を聞こうとする被験者にとっては、感情表現音なしの方が音声に意識を集中できるので、感情が感じられやすいと評価されたと考えられる。

それでは、感情表現には人間の音声のみを利用するがよく、感情表現音や他のメディアは使わない方がよいのであろうか。しかし、実験1によると、音は使わないよりも使った方が感情はわかりやすいという傾向が出ている。また、実験2についても、感情表現音を聞くことでより感情を捉えた被験者もいる。これらの結果からは感情表現音の感情表出に対する有効性を見ることができる。ただ、他メディアとの組み合わせ方、利用する音そのもののデザインによっては、感情表現が有効に機能しないばかりか、他の表現の効果をも阻害しかねない、ということが言える⁹⁾。

本実験結果から結論づけられることは、非言語音は感情表現のメディアとして有効となり得るものの、その使い方によっては、他メディアによる感情表現を阻害したり、使用感を損ねたりするものにもなりかねない、ということである。使い方としては、感情表現音は必要以上に強い印象をユーザに与えることのないせりふ前に発音するのがよいという示唆が得られた。ただ、これは感情表現音の音質、音色（音声との相性）や各音のデザインなどとの関連も考えられ、どのような音が適しているかは今後評価検討すべき課題である。

5. まとめ

本論文では、感情表現における非言語音利用の観点から、擬人化エージェントの感情を表現する感情表現音を制作し、その有効性を考察した。実験の結果、非言語音は感情表現のメディアとして有効であるが、使い方によってはユーザに不快感を与えたり、他メディアの表現を阻害する危険があることがわかった。

今後は、聴覚上の心地よさの問題は音質や音色に大きく影響されるため、音そのもののデザインについて評価検討を行う必要がある。

【謝辞】

本論文中には、アルパイン株式会社殿と NEC との共同研究成果が含まれている。討議に参加していただいたアルパイン株式会社の斉藤文夫、大石哲也、斎藤望、穴戸博他の諸氏に感謝する。

【参考文献】

- [1] 和氣他, "音響インタフェース設計手法「報知音多次元設計手法」確立への一考察", 電子情報通信学会論文誌 D-II Vol.J82-D-II No.10 pp.1721-1728, 1999 Oct.
- [2] 下原, "感性と理性を結ぶ「人工生命」", 感性・人間・コンピュータ, 大沢他共著, 第1章 pp.10-46, 富士通ブックス
- [3] 土佐, "芸術とテクノロジー—ファジィ技術を用いたインタラクティブアート『ミック&ミュージ』—。解説 シリーズ・感性について考える2-" 日本ファジィ学会誌, 9, 5, 648-655, 1997
- [4] 牛田, 平山他, "自立的行動決定モデルに基づくインタフェースエージェント", 電子情報通信学会論文誌 D-II Vol.J82-D-II No.10 pp.1655-1665, 1999 Oct.
- [5] 河野他, "擬人化キャラクタによる情報伝達", 第13回人工知能学会全国大会論文集 S6-05, pp139-142, 1999
- [6] 岩宮, "音と映像の感性的調和", 第4回 日立中研研究会予稿集, 39-44, 1996.6. (<http://animal.kyushu-id.ac.jp/~iwamiya/hitati.html>)