

## 多人数エージェントによる会話の雰囲気生成 —文字ばかり読んでないで空気読め—

木村 幸士<sup>†</sup> 湯浅 将英<sup>†</sup> 武川直樹<sup>†</sup>

<sup>†</sup> 東京電機大学 情報環境学部 〒270-1382 千葉県印西市武西学園台 2-1200

E-mail: DEN03471@nifty.com, {yuasa, mukawa }@sie.dendai.ac.jp

### あらまし

本研究は、擬人化エージェントによる発話交替シミュレータを用いて様々な発話交替のパターンを生成し、評価実験により会話の雰囲気との関係を調べる。複数の人が会話をしているとき、人はたとえその会話に参加していなくても、会話の雰囲気やその人たちの間柄などを推測できる。他者の会話中の発話開始タイミングや表情が会話の雰囲気を作り出しているという仮説のもとで、発話交替シミュレータによって擬似的に再現された発話の衝突あるいは沈黙などの会話パターンを評価し、発話交替パターンと人が感じる会話の雰囲気との関係を明らかにした。本研究の結果は、会話の雰囲気や空気の読めるエージェントやロボットの実現に寄与できる。

キーワード 擬人化エージェント, 発話交替, 多人数会話, 会話の雰囲気, KY

## Generating Conversational Atmosphere using Turn-taking Agents —Don't Read Books, Read the Atmosphere!—

Koji KIMURA<sup>†</sup> Masahide YUASA<sup>†</sup> and Naoki MUKAWA<sup>†</sup>

<sup>†</sup> School of Information Environment, Tokyo Denki University, 2-1200 Muzai Gakuendai, Inzai, Chiba, Japan

E-mail: DEN03471@nifty.com, {yuasa, mukawa }@sie.dendai.ac.jp

**Abstract** This paper investigates conversational atmosphere in conversation scenes generated by multi-party turn-taking agents. When we overhear others' conversation, we can infer relationship among the participants and understand conversational atmosphere. We assume that the conversational atmosphere is generated by the turn-taking rules of participants and their facial expressions. We conducted experiments to evaluate the atmosphere by using various kind of conversational scenes, such like the utterance collision where more than two persons start to speak and the silence where no one intend to speak. Experimental results show that the turn-taking patterns and facial expressions influence on the conversational atmosphere. The findings in the experiment will contribute to development of life-like agents and robots that have the function to comprehend and maintain the conversational atmosphere like humans have.

**Keyword** Life-like agent, turn-taking, multi-party conversation, conversational atmosphere, KY

### 1. はじめに

人は、自分が関わる会話だけでなく、他者同士の会話からも「会話が盛り上がっている」「仲が良さそうに話している」あるいは「何か暗い話題を話しているようだ」などの雰囲気を容易に読み取ることができる。うまく雰囲気が読み取れずに間違った言動をすると「空気が読めない (KY)」などと言って揶揄されてしまう [1]。しかし、会話の場の雰囲気を具体的に定義し、雰囲気を感じ方を言葉にして説明することは大変難しく、さらにその雰囲気を感じるメカニズムも明らかとなっていない。

本研究では、他者同士の会話を見て感じる雰囲気として特に発話交替の役割に注目する。発話交替は、話

者と聞き手の役割が次々と交替することを指すが、同時に話せるのが一人であること、全員が沈黙するのが避けられるなど、会話を制御する交差点としての役割を持つ [2-5]。この交差点で、誰がどのように話を始め、誰が話をやめるかという駆け引きが、会話の面白味や緊張感、満足感などを生み出し、それが「会話の雰囲気」として表現される。

本研究では、発話交替が作る「会話の雰囲気」の生成メカニズムを明らかにすることを目的として、様々な発話交替のパターンを生成し観察できる発話交替シミュレータ ARABAHIKA を作成した [6]。ARABAHIKA では、多人数からなる擬人化エージェントの発話開始、表情などが調整できる。本稿では、ARABAHIKA による発話交替のパターンと会話の雰囲気との関係性を評価実

験により調べた結果を報告する。この研究の成果は、「会話の雰囲気とは何か」、「会話の雰囲気はどのように作られ、理解されるか」という会話コミュニケーションの基本問題の解明と、場の雰囲気を理解して人と気持ちを通じ合わせる能力を持つ擬人化エージェントやロボットの実現に寄与する。

## 2. 従来研究

心理学や言語学では、会話中の発話交替と視線の振る舞いとの関係が分析され、話者によって見られた聞き手が、次に話し出す傾向があることが明らかにされている。この知見を利用して、ユーザの視線の動きを検出し[7-10]、発話の交替を実現するエージェントやロボットが開発されている[11-14]。しかし、従来までのエージェントやロボットの発話交替は、「聞き手となるロボットは、話者を見なければならぬ」「話者が、発話を終わるときロボットを見たら、ロボットは話し出さなければならぬ」など単純なルールに従っているだけであった。このような単純なルールを用いた場合には、人は話したいときに思いのまま話させてもらえず、話したくないときにも話さざるを得ず、活発な楽しい会話の雰囲気を作ることができない。

社会心理学や認知心理学では、他者の気持ちを読み取る「マインドリーディング」の研究が多数ある[15-19]。徳永らは、発話交替時における聞き手の話したい／話したくないという「発話マインド」を調べ、会話で人は発話マインドをお互いに表出して解釈していることを述べている[5]。また、Poggiらは、話者の声の音量や韻律などの非言語情報の表現と発話欲求の関連を分類している[16]。しかし、それは個人に着目した非言語情報表現と発話欲求に限定され、複数人に関わる会話全体の雰囲気までを考慮したものではなかった。会話の雰囲気について、伊藤らは擬人化エージェントを用いた研究により、ユーザが指定した雰囲気を実現するために視線や表情などのノンバーバル情報を調節する方法について述べている[20]。しかし、実現した雰囲気についての評価は不十分であり、エージェントの行動と雰囲気との関係を明確にする課題が残った。

本研究では、複数の擬人化エージェントを用い、様々な発話開始と表情などのパターンを擬似的に再現させ、発話交替にかかわる会話の雰囲気の評価を試みる。特に、3人会話において、話者の発話終了時に2人の聞き手が同時に話す、話者がそっぽを向いていても誰かが話すなどの様々な発話交替と視線の関係を用いて、会話の雰囲気の評価する。そして、複数人の非言語情報によって生成される雰囲気の評価し、多数の非言語情報と多人数コミュニケーションの関係

を探る。

## 3. 擬人化エージェントを用いた発話交替の再現

本研究では、擬人化エージェントを用いた発話交替のシミュレーションシステムを開発し、様々な発話交替を再現させ、そのときの会話の雰囲気を評価実験により評価する。

人の体を模したCGキャラクターである擬人化エージェントは、人の認知的負担を軽減できるインタフェースとして注目されている。これまでに、擬人化エージェントシステムを簡易に構築可能なGalateaやTVMLなどのツールが開発されている[21]。本研究ではTVMLを基にした発話交替シミュレーションシステムARABAHIKAを使用した[6]。ARABAHIKAによる会話の例を図1に示す。図1は、3体のエージェントが向き合い、人間の言語を話しているように聞こえる無意味音声(ARABAHIKAやUKUJARAH等)を発話する。無意味語を用いた理由は、会話に意味内容を持たせると、その内容が発話交替の評価に影響し、発話交替のみの影響評価が難しいと考えたためである。



(a) 笑顔の例



(b) 怒り顔の例

図1 ARABAHIKAの画像例

無意味語であってもエージェント会話を見る人からは3体のエージェントが話し合いをしているように見える。また、このシステムでは、擬人化エージェントの発話交替におけるパラメータを自在に調整できる。たとえば、「50%の確率で話し出す」、「話者に見られた人が75%の確率で話し出す」ことを人同士の会話で再現することは困難であるが、エージェントであれば、それらを再現することが可能である。

なお、予備実験において、人に近いリアルなCGのキャラクターを用いたところ、「表情がわかりにくい」との指摘があったため、ここでは目、眉の動きが明確で、表情の動きがわかりやすいアニメ的なキャラクターを使用した。

## 4. 実験

### 4.1. 実験刺激パターン

評価のため3体のエージェントによる12の会話パターンを作成した(表1)。これらの会話は、発話交替パターン(GT, BT, NT)、表情、話者の話し終わりから次話者の初めの時間間隔、によって変化する。

#### (1) 発話交替パターン

##### I. 話者視線による次話者選択パターン

###### (GT: Gaze-based turn)

話者に見られた聞き手が話すパターン。たとえば、エージェントAがエージェントBを見て話し終わると、Bが次に話し始める、BがエージェントCを見て話し終わるとCが話し出す。

##### II. 二人同時に話し出すパターン

###### (BT: Both Participants' Turn)

話者に見られた聞き手、見られない聞き手の両者が話し出し、発話が衝突する頻度が高いパターン。発話の衝突とは、たとえば、エージェントAがエージェントBを見て話し終わると、BとエージェントCの両方が話し始めるものである。発話衝突後、どちらが話し続けるかは、システムがランダムに決定する。実験では、75%の確率で衝突を発生させた。

表1 実験刺激パターン(以降ではこの略称を用いる)

	会話交替	表情	間
1	GT	笑	早
2	GT	笑	遅
3	GT	怒	早
4	GT	怒	遅
5	BT	笑	早
6	BT	笑	遅
7	BT	怒	早
8	BT	怒	遅
9	NT	笑	早
10	NT	笑	遅
11	NT	怒	早
12	NT	怒	遅

### III. ターンなし

#### (NT: No turn)

話が終わっても、誰も話し出さない頻度が高いパターン。たとえば、エージェントAがBを見て話し終わっても、BもCも話し出さない。この場合、数秒間の沈黙が起こった後に、再びAが話し始める。本実験のこのパターンでは、25%の確率で、この沈黙が起こるとした。

#### (2) 表情

(1)で述べた発話交替パターンに、笑顔(図1a)と怒り顔(図1b)の2種を付加する。

#### (3) 話者交替間隔(間)

発話交替に際し、1体のエージェントが話を終えて、別のエージェントが話し出すまでの時間を、二段階で変化させた。これらは予備実験により、話し始めてから次話者に話に移るまでを3.3秒とした「間が早い」、4.2秒とした「間が遅い」を用いた。

### 4.2. 評価方法

前節で説明した実験刺激パターンを実験協力者に提示し、5段階の形容詞対によるSD法により評価させた。事前の予備実験から「雰囲気」が「親和性」、「覚醒度」の2因子から構成されると仮定してこれらの評価項目とした。これら2つの因子は顔や表情、対人関係の印象評価でよく使われ、「仲がよさそうな雰囲気(=親和性が高い)」、「テンションが高い雰囲気(=覚醒度が高い)」に対応する。用いた形容詞対を下記に示す。

・親和性に関するもの:「協調 - 対立, 友好 - 敵対, 仲がいい - 仲が悪い, 親密 - 疎遠」

・覚醒度に関するもの:「激しい - 穏やか, 高揚 - 消沈, 張りつめた - 緩んだ, 興奮 - 冷静」

実験刺激には、エージェント同士の会話が途切れることなく次々と発話が交替する会話シーンを提示した。なお、表1の実験刺激のパターンの順番については、実験協力者ごとにランダムに並び替えて提示した。

### 4.3. 仮説

検証すべき仮説を以下に示す。

#### 仮説1: 表情における親和性

「笑顔」の会話は、「怒り顔」の会話よりも「親和性」の評価が高い。

#### 仮説2: 表情における覚醒度

「怒り顔」の会話は「笑顔」の会話よりも「覚醒度」の評価が高い。

#### 仮説3: 話者交替における覚醒度

「覚醒度」の評価はBT > GT > NTの順になる。

#### 仮説4: 間における覚醒度

「間が早い会話」は「間が遅い会話」よりも「覚醒度」の評価が高くなる。

仮説1と仮説2は、先行研究に基づけば、笑顔のとき親和性が高く、怒り顔ならば覚醒度が高いとされているために今回のエージェント実験でも同様の結果となると考えた。仮説3は、発話を複数の人が開始する、一人だけ開始する、誰も開始しない、の順に発話への意欲が下がり、それに伴い覚醒度が低くなると考えた。

仮説4については、間が早いと高揚し興奮した雰囲気、間が遅いと消沈し緩んだ雰囲気に感じられると予想したためである。

#### 4.4. 結果

理系大学生10名（男性7名、女性3名）の実験協力者に対して実験を行った。親和度、覚醒度ごとに形容詞対の評価平均値を算出した結果を図2に示す。図は縦軸が覚醒度、横軸が親和性を表す。t検定を実施し、パターン毎の有意差を調べた。

##### 仮説1、仮説2の検証

表情以外のパラメータが同一の条件間（たとえば、「笑・GT・遅」と「怒・GT・遅」条件）で比較した

結果、「笑」の会話は、「怒」の会話よりも親和性評価が高いことが分った。また、「怒」の会話は「笑」の会話よりも覚醒度の評価が高いことがわかった。t検定の結果、表情が「笑」か「怒」の場合には、親和性、覚醒度のどちらも有意な差があった。「BT・遅」では覚醒度  $p < 0.05$ 、他の場合には  $p < 0.01$  で有意であった。したがって仮説1および仮説2は支持された。

##### 仮説3の検証

話者交替について、表2に示す。覚醒度を評価した結果、 $BT > GT > NT$ の順になった。GTとNTでの「笑・遅」「怒・遅」、BTとGTでの「笑・早」「怒・早」以外の覚醒度においては、 $p < 0.05$ で有意差が認められた。特にNTとBTについては覚醒度に明確な差があり、仮説3は支持された。また、GTに関しても、NTとの「笑・早」「怒・早」での比較、BTとの「笑・遅」「怒・遅」で有意差が認められた。

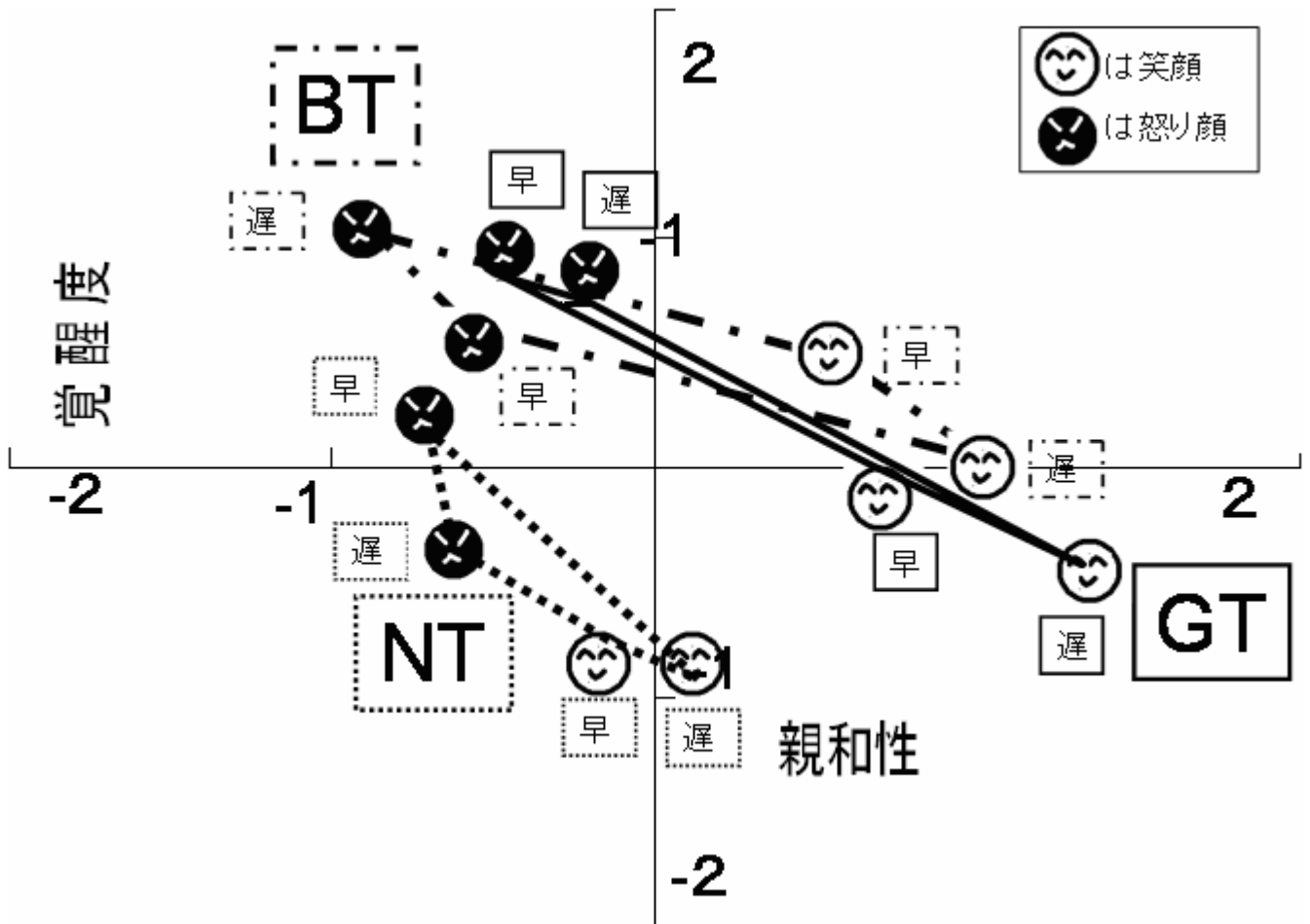


図2 会話の雰囲気と発話交替、表情、間の関係。縦軸は覚醒度、横軸は親和性、「BT（両方とも話す）」、「GT（視線での交替）」、「NT（沈黙あり）」は発話交替「早」「遅」は話者交替間隔

表2 発話交替について t 検定結果

	親和性	覚醒度
<b>GT &gt; NT</b>		
笑・早	—	*
笑・遅	**	—
怒・早	—	**
怒・遅	—	—
<b>BT &gt; GT</b>		
笑・早	—	—
笑・遅	—	*
怒・早	—	—
怒・遅	—	*
<b>BT &gt; NT</b>		
笑・早	**	*
笑・遅	**	**
怒・早	—	**
怒・遅	—	**

(注)\*は  $p < 0.05$ , \*\*は  $p < 0.01$  を示す。

一方、親和性については、t 検定の結果、有意であったものは「NT・笑・早」と「BT・笑・早」の間、「GT・笑・遅」と「NT・笑・遅」の間、「GT・笑・遅」と「BT・笑・遅」の間であり、他の比較においては有意差が認められなかった。発話交替パターンの違いの親和性への影響は部分的であった。

#### 仮説4の検証

「間」を変えることによる覚醒度は、「間が早い会話」は「間が遅い会話」よりも若干「覚醒度」の評価が高くなったが、「間」以外のパラメータが同一の条件間である6組において  $p < 0.05$  のものはなく、仮説は棄却された。

## 5. 考察

発話交替に関して、親和性より覚醒度のほうが明確な差が見られた。覚醒度については発話交替の各パターンで「テンションが高い雰囲気」あるいは「テンションが低い雰囲気」であることは判断しやすかったと考えられる。特に、NTとBTの覚醒度の有意差がどの表情でも見られることから、「誰も話し出さない」と「二人同時に話し出す」の違いは、「テンションの低い会話」か「テンションの高い会話」を判断するために重要な情報であったと言える。

一方、親和性については発話交替パターンの変化だけでは、「仲がいい雰囲気」あるいは「仲が悪い雰囲気」ということは判断しづらいことが分かった。たとえば、同じ笑顔であれば、黙っていても、話が盛り上がっていても、仲が良く見えてしまうことがあると考えられる。

今回の実験において、表情の会話の雰囲気への寄与

が非常に大きいという結果が得られた。これは顔表情がコミュニケーションに影響を与えるという先行研究の知見と一致する。複数人の会話の雰囲気についても表情の要素が強く影響することがわかった。

また、間の変化は会話の雰囲気に全く影響を与えなかった。これは間の設定が適切でなく、「遅い」と「早い」の時間の差が実験協力者に知覚されにくかったためと考えられる。今後、追加実験により知覚しやすい時間を設定したい。

## 6. まとめ

発話交替が作る会話の雰囲気の生成メカニズムを明らかにするため、エージェントにより擬似的に再現させた発話交替パターンについて、評価実験を実施した。実験の結果、発話交替パターンの違いが会話の雰囲気の生成に影響していることが分かった。特に、二人が同時にしゃべりだす会話交替パターンと、沈黙がある会話交替パターンでは、会話の覚醒度の評価が大きく異なることが得られた。しかし、発話交替パターンと会話の親和性との関係は有意ではなかった。一方、会話参加者の表情は親和性に影響していることがわかった。感情コミュニケーションにおいて重要な役割を持つ顔表情が、複数人の会話の雰囲気の生成にも強く関連することが確認できた。

今後の課題として、「間」に関して今回の設定では雰囲気に影響を与えなかったために、パラメータ設定の見直しをする。表情や発話交替パターンを時間毎に変化させた場合、会話の雰囲気にどのように影響を与えるかを調べる予定である。

## 謝辞

本研究の一部は、文部科学省科学研究費補助金 基盤研究(B)16300032, 2006, 同若手研究(B) 19700119, 東京電機大学総合研究所研究(課題番 Q04J-10,Q06J-14), および東京電機大学先端工学研究所重点課題による。

## 文献

- [1]KY:<http://dic.yahoo.co.jp/newword?index=2007000362&ref=1>
- [2]榎本美香, 伝康晴: 3人会話における参与役割の交替に関わる非言語行動の分析, 人工知能学会研究会資料, SIG-SLUD-A301, pp.25-30 (2003)
- [3]石崎雅人, 伝康晴: 談話と対話, 東京大学出版会 (2001).
- [4]榎本美香, 伝康晴: 3人会話における参与役割の交替に関わる非言語行動の分析, 人工知能学会研究会資料, SIG-SLUD-A301, pp.25-30 (2003).
- [5]徳永弘子, 湯浅将英, 武川直樹: 3人会話の発話交替における視線行動と発話マインド分析-聞き手の立場から見た発話・非発話の戦略-, 電子情報通信学会技術研究報告書 HCS2007-40, pp.7-12 (2007)

- [6]湯浅将英, 徳永弘子, 武川直樹: 発話交替シミュレーションシステム ARABAHIKA -人間観察に基づく自律的発話交替エージェントの提案 -, 情報処理学会インタラクシオン 2007 論文集 pp.87-88(2007)
- [7]Kendon, A: Some functions of gaze direction in social interaction, *Acta Psychologica*, Vol.32, pp.1-25 (1967).
- [8]Argyle, M. and Cook, M.: *Gaze and Mutual Gaze*, Cambridge University Press (1976).
- [9]Sacks, H., Schegloff, E. and Jefferson, G.A.: simplest systematics for the organisation of turn-taking in conversation, *Vol.50, No.4*, pp.696-735 (1974).
- [10]Vertegaal, R. and Ding, Y.: Explaining effects of eye gaze on mediated group conversations: Amount or synchronization-, In *Proc.2002 ACM conference on computer supported cooperative work* (2002).
- [11]松坂要佐, 小林哲則: ROBITA: グループ会話ロボット, 人工知能学会研究会資料 SIG-Challenge-0113, pp.1-8 (2001).
- [12]Scassellati, B.: Theory of mind for a humanoid robot, *Autonomous Robots*, Vol.12, No.1, pp.13-24 (2002).
- [13]Kojima, H. and Yano, H.: A robot that learns to communicate with human caregivers, In *Proc. of the First International Workshop on Epigenetic Robotics*, pp.47-52 (2001).
- [14]Eichner, T., Prendinger, H., Andre, E. and Ishizuka, M.: Attentive presentation agents, In *Proc. IVA-07*, pp.283-295 (2007).
- [15]Baron-Cohen, S.: *Mindblindness*, The MIT Press (1995).
- [16]Poggi, I., Pelachaud, C. and Caldognetto, E.M.: Gestural mind markers in ECAs, In *Proc AAMAS 2003*, pp.1098-1099 (2003).
- [17]Malle, B.F. and Hodges, S.D.: *Other Minds*, The Guilford Press (2005).
- [18]Sperber, D. and Wilson, D.: *Relevance: Communication and Cognition*, Blackwell Pub(1995).
- [19]Ono, T. and Imai, M.: Reading a Robot's Mind: A Model of Utterance Understanding based on the Theory of Mind Mechanism, In *Proc. AAAI-2000*, pp.142-148 (2000).
- [20]伊藤 淳子, 角所考, 美濃導彦: 親密性平衡モデルに基づく CG キャラクタ同士の会話映像における雰囲気制御, 電子情報通信学会技術研究報告書, HIP2002-69 pp.37-42, (2003)
- [21]TVML: <http://www.nhk.or.jp/str1/tvml/>