

# 効果音のレトリックと物語生成における利用

## The Rhetoric of Sound Effect and its Use in Narrative Generation

平松 雅也\*<sup>1</sup>  
Masaya Hiramatsu

小方 孝\*<sup>2</sup>  
Takashi Ogata

岩手県立大学大学院\*<sup>1</sup>  
Graduate School of Iwate Prefectural University

岩手県立大学\*<sup>2</sup>  
Iwate Prefectural University

In this research, we pay attention to “sound effect” in narrative representation. Our goal is to develop a mechanism in our narrative generation system that automatically attaches sound effects into visual narrative using animated movie and other types’ narrative representations. First, we propose a framework of system that is composed of sound effects attaching processing, sound effects controlling processing, sound effects database and sound effect rhetoric knowledge base. Second, we collect sound effects from a movie and extract parameters to define basic characteristics of each sound effect, and moreover attempt a preparation to more complex rhetoric in the unit of a scene. Third, we implement a system that controls the use of sound effects based on above framework and analysis.

### 1. はじめに

筆者らが進めている「物語生成システム」[小方 2003ab]における物語表現媒体には、主に言語、映像、音の三つがあるが、本研究ではこのうち音に注目する。物語における音の表現には、音楽、効果音、台詞の音声、環境音等があるが、ここでは「効果音」を取り上げる。[久保 2003]によれば、効果音とは、映画・演劇・アニメ・ゲーム等において、演出の一環として付け加えられる音を言う。本研究の目標は、物語生成システムによって生成された映像情報に対して、効果音を自動的に添付するシステムの構築である。但し効果音そのものの音響的合成ではなく、何らかの形で用意された効果音の断片を物語の内容等を参照して意図的に添付する機構を目指す。本稿では、まず次の2節でこのようなシステムの構想と試作、3節と4節ではこのシステムにおいて重要な役割を果たす「効果音のレトリック」に向けた検討を行う。ここで言う効果音のレトリックとは、物語の状況や映像表現に対応して、何らかの効果音を使用するための様々なルールを意味する。3節では、一つの映画を素材として、効果音の基礎的要素を整理し、4節では効果音のレトリックのための最も基礎的な機構を2節のシステムの枠組みの上に実装した例を示す。この試作は、NHKが開発したCGアニメーション構築用ツールであるTVML(TV program Making Language) [林 2004]で記述されたスクリプト言語を入力として、効果音の再生を行うコマンドを自動的に書き加えるものである。情報処理の観点から効果音を扱う研究として、例えば[有田 2004]や[金 2007]のように受け手による効果音の受容を分析した研究があるが、本研究のように生成の観点から効果音処理の自動化を目指す研究は少ない。但し、効果音そのものの生成ではなく、その「編集」がここでの目標である。

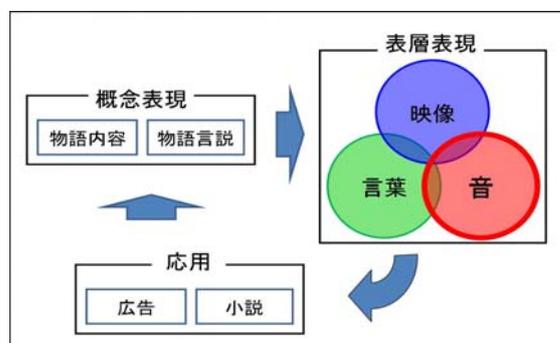


図1:物語生成システムの概要

### 2. 効果音自動添付システムの枠組み

本研究は、効果音の分析・整理とシステム実装・検証の往還によって進めることを予定している。そこでまず「効果音自動添付システム」と呼ぶシステムを構想した(図2)。利用する映像ツールとしては前述のTVMLを想定して議論を進める。システムは効果音添付処理、効果音データベース、効果音制御、レトリック知識ベース4つの機構からなる。このような枠組みをまず用意し、この中に分析結果により得られた具体的な知識やレトリックを入れて行くという手順で研究を進めて行く。

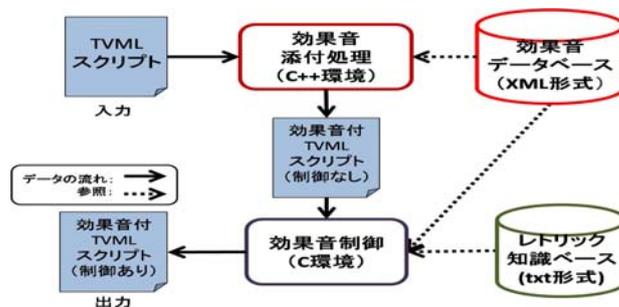


図2:効果音自動添付システム

処理の流れであるが、初めに効果音が添付されていないTVMLスクリプトが効果音添付処理機構に入力される。TVMLスクリプトは図3で示すようにセットアップ記述部と動作記述部から構成されており、効果音添付の対象は後者である。効果音

添付処理は、表 1 に示すような効果音データベースに格納されている効果音のキー(後述)を動作記述部から検索し、発見された場合その一行下に効果音を添付する。下行に添付することにより音の対応物との同期が取れる。効果音のキーとしては現在、表 1 に示すように、〈場所〉、〈動き〉、〈無生物〉及び〈生物〉を用意している。表 1 の効果音の種類は、ビクターエンタテインメント株式会社から発売されている効果音 CD『効果音大全集』(2005)等から収集したものである。例えば「砂浜の波」と命名された効果音に対して、〈場所〉として「砂浜」、〈動き〉として「波打つ」、〈無生物〉として「波」というキーを用意する。そのキーによって検索された効果音のファイルが全て効果音添付の対象となる。

```
//カメラのセットアップ
camera: assign( name=CamA )
camera: movement( name=CamA, x=0.0, y=1.0, z=3.0, tilt=5.0, vangle=70 )
camera: switch( name=CamA )

//照明のセットアップ
light: assign( name=light1 )
light: model( name=light1, type=flat, x=1.0, y=1.0, z=1.0, r=1.0, g=1.0, b=1.0 )
light: assign( name=light2 )
light: model( name=light2, type=flat, x=-1.0, y=1.0, z=1.0, r=1.0, g=1.0, b=1.0 )

//スタジオのセットアップ
set: assign( name=location )
set: openmodel( name=location, filename="desert%desert.obj" )
set: change( name=location )

//キャラクターのセットアップ
character: casting( name=agent1 )
character: openmodel( modelname=abeno, filename="abeno%abeno.bm" )
character: bindmodel( name=agent1, modelname=abeno )
character: setvoice( name=agent1, voicetype="male01" )
character: position( name=agent1, x=-1, y=0, z=0, d=90, posture=standing )
narration: narratorvoice( voicetype="man01" )

skipscript(switch=off)

character: sequencepose(name=cagent1, sequencepose=abc,wait=no)
//ナレーションの実行
narration: talk( who=narrator, text="むれにおわれ", caption="群れに追われ",
character: shake(name=agent1, state=on, level=3)
wait(time=1)
character: walk(name=agent1, x=-4, d=270, pitch=2)
```

セットアップ部  
動作記述部

図 3:TVML スクリプトの構造

表 1:効果音の分類表の一部

分類	効果音名	場所	動き	無生物	生物
背景	砂浜の波	砂浜	波打つ	波	
背景	太平洋の波	太平洋	波打つ	波	
背景	磯の波	磯	波打つ	波	
背景	峡谷の急流	峡谷	流れる	川	
背景	せせらぎ		流れる	川	
背景	滝	滝	落ちる	滝	
背景	水車	川	回る	水車	
背景	ししおどし	池	鳴る	ししおどし	
背景	夕立		降る	雨	
背景	雨垂れ		降る	雨	
背景	近雷		鳴る	雷	
背景	遠雷		鳴る	雷	
背景	弱い風		弱く吹く	風	
背景	強い風		強く吹く	風	
背景	木枯し		吹く	風	
背景	台風		吹く・降る	風・雨	
背景	吹雪		吹く・降る	風・雪	
背景	風鈴(南部風鈴)		揺れる	風鈴	
背景	警笛	駅	鳴る	新幹線の警笛	
動作	アスファルト道	アスファルト	歩く		人
動作	砂利道	砂利道	歩く		人
動作	交差点	交差点	歩く		人

以上のように、システムは、TVML スクリプトの記述にキーと一致する文字を発見した場合、発見した位置の下行に効果音ファイルを可能な限り添付する。

### 3. 効果音の分類・整理—効果音レトリックに向けて—

システムの枠組みでは、可能な効果音がすべて添付された TVML スクリプトを対象に、「効果音制御」が行われる。効果音制御とは、TVML スクリプトの一単位において、どんな効果音を、どのように使用するか全般を指す言葉とし、ここではこの制御をレトリックという観点から検討する。システムにおいては、制御のための具体的な諸知識・ルールが効果音レトリックに格納されるものとする。レトリックとは一般に、言語表現を初めとした様々な表現の技法や技法の組み立てを言う[内海 2007]。ここでは、個々の効果音の調節、複数の効果音の組合せ的使用、何らかの意図や戦略に基づく効果音あるいは効果音群の組合せ的使用等の問題が含まれる。

このような意味での効果音のレトリックへの手掛りを得るために、北野武の映画『菊次郎の夏』(1999)の調査を試みた。音楽と台詞の発話を除いたものを効果音と見なし、映画の各シーンごとにすべての効果音を抽出した。

その結果について考察し、まず効果音のタイプを次のように分類した—(1)対応する事物がある(推定可能な)効果音:①可視(映像として映っており目に見える)もの、②不可視(映像として映っておらず、目に見えない)もの。(2)対応する事物がない(推定不可能な)効果音。

この映画は全部で 186 シーンあり、映画全体の効果音の使用回数は述べ 352 回(但し数を算定するのが困難なものもあるので概算)、効果音の種類(上の「対応する事物があるもの・ないもの」に適当な名前を付け、その名前が共通のもの)は 169 種類あった。以下に、可視の効果音、不可視の効果音、対応する事物が推定不可能な効果音の一覧を掲げる(カッコ内はそれが出現するシーンの数)。

**可視の効果音:** 足音(53)/風鈴(2)/車(11)/叩く(6)/ドア(9)/サッカーボール(1)/自転車(3)/転ぶ(1)/花火(1)/水(3)/皿(2)/ビニール袋(1)/車のドア(7)/ノート(2)/リュック(1)/チャック(1)/タップダンス(2)/グラス(2)/人のギャギャ(2)/人を引きずる(1)/鐘(1)/グラスに飲み物を注ぐ(1)/焼き鳥を焼く(1)/ホウキ(1)/倒れる(2)/車の発車音(1)/ハザード(1)/車の車内音(2)/「ギーッ」という音(1)/サングラスを置く(1)/水に飛び込む(2)/泳ぐ(2)/テーブルを叩く(1)/肩たたき(1)/自動ドア(1)/車を叩く(2)/木を投げる(1)/オレンジをキャッチ(2)/オレンジを落とす(2)/ロボットダンス(1)/紙袋(1)/おにぎりを落とす(1)/溝に落ちる(1)/雨(2)/杖で叩く(1)/車のブレーキ(1)/タイヤパンク(1)/スリップ(1)/車が崖に落ちる(1)/タイヤをける(1)/波(3)/天使の鈴(5)/射的(1)/机を叩く(1)/神社の鈴(1)/シャッターを叩く(1)/風車(1)/カエル(1)/カエルが池に潜る(1)/とうもろこしを取る(1)/バイク(3)/水をバシャバシャする(1)/石を置く(1)/いす(1)/弓を射る(1)/太鼓(1)/「コツツ」という音(1)/縄がきしむ(1)/縄がほどける(1)/「ピュン」という音(1)/「ピュン」という音(1)/ドラム缶を倒す(1)

**不可視の効果音:** 天使の鈴(7)/車のクラクション(3)/街の騒音(3)/着地(1)/セミ(3)/車(27)/バイクのエンジン音(1)/鳥(9)/風鈴(4)/虫(13)/風(36)/水(2)/車のエンジン音(2)/引き出し(1)/ドア(2)タップダンス(1)/劇場のベル(1)/自転車(1)/人のギャギャ(6)/歓声(3)/スタートのピストル(1)/鐘(2)/アナウンス(1)/電話のベル(1)/カラス(3)/プールに飛び込む

(1)/車のガラスを割る(1)/小便(1)/屁(1)/雨(2)/車のブレーキ(1)/車にひかれる(1)/足音(2)/波(2)/笛(3)/太鼓(4)/石をぶつける(1)/景色が落ちる(1)/シャッターを叩く(1)/とうもろこしを取る(1)/バイク(1)/リールを巻く(1)/矢が刺さる(1)/いびき(1)

対応する事物がない効果音:「ドッド」<sup>①</sup>という音(2)/違う世界にいるような音(1)/カミナリのような音(1)

これらは、人物の動きに依存した効果音、生物(人物以外)の動きに依存した効果音、無生物の動きに依存した効果音、場所に依存した効果音、何にも依存しないいわば純粋な効果音(上のどれにも属さない効果音)に分けられるが、これは 2 節のシステムにおける検索キーの種類にほぼ対応している。

ここで対応する事物がない効果音とは、映像からも音からもそれが何の音であるか推測するのが困難であった音であり、上の一覧の「ドッド<sup>①</sup>という音」は、主人公(正男)の夢の中で使用されているが、このシーンでは、音に対応するものが映し出されない(映されているものの何かがこの音に対応しているとは考えにくく)。この映画ではこの種の効果音が 3 種類使用されているが、いずれも現実の世界と夢の世界とが混合するシーンにおいて使用されており、受け手の立場から見れば、両者の境界・区別の認識が促進されると思われる。

ところで可視の効果音において、対応している事物と効果音の使用されるタイミングが同期している場合と同期していない場合がある。例えば、歩くという動作で足が地面に着くと同時に足音が出る場合やカメレオンが横を向くと同時に効果音が鳴る場合は同期していると考えられる。この映画における効果音は大部分対応事物と同期しているが、他方で対応事物と効果音が同期していない場合(非同期の場合)もある。例えば、素材の映画では、「天使の鈴」が振られ少し遅れてから鈴の音が流れる、といった非同期の効果音使用が四回見られた。また、効果音の音量にも工夫が凝らされており、例えば、波と車の音は可視の方が不可視より音量が大きく、逆に足音は可視より不可視の方が音量が大きかった。このように、同期/非同期の区別や音量の調節が効果音の性質を規定する一因となっている。

これらは個々の効果音の性質を規定する要素であり、レトリックという観点からは最も基礎的なものに当る。これら基礎的なレトリックをシーンあるいはそれ以上の単位を対象に、何らかの方策に従って組み合わせる使用することによって、より高度なレトリックが構成される。この整理は別稿に譲るが、映画におけるシーンレベルでの効果音使用がどのように行われているかを示す例を以下に 2 つ掲げる(図 4, 5)。ここで、□で囲っているものは使用されている効果音であり、そこから延びる矢印はその持続範囲を示している。図 4 は、トラックの運転手と主人公の菊次郎がケンカをするシーンであり、画面の奥で菊次郎とトラックの運転手が言い争いをしているところから始まり、その後殴り合いのケンカが始まり、運転手が画面から消え菊次郎だけが現れる。ここでは、シーン開始から虫とカラスの鳴き声が流されているが、運転手と菊次郎が殴り合いを始めた瞬間に今まで流れていた鳴き声が止んで無音になる。そしてケンカが終るとまた虫とカラスの鳴き声が流れ始める。つまり、物語中の重要な行為の部分で効果音を故意に消去することで、効果音と沈黙の対比を強調し、効果音のない行為の強調している、というのが普通の解釈である(別の解釈もあり得る。例えば、重要な出来事は、前後であるかも知れない)。図 5 のシーンでは、登場人物の動きに伴って効果音が鳴り、その背後で通奏低音的に別の効果音が鳴り続ける。

シーンの時間	t →			
出来事と効果音	トラック停止 菊次郎と運転手が前に立っている	菊次郎と運転手がトラックの後ろに移動しケンカ	ケンカが殴りあいになる	菊次郎のみトラックの前に現れる
可視のものとの効果音	トラック 菊次郎 運転手 野原			
不可視のものとの効果音				

図 4: 菊次郎(主人公)と運転手がケンカするシーン

シーンの時間	t →			
出来事と効果音	お兄さんとお姉さんのところに正男が行く	お兄さんがロボットダンス <sup>①</sup> を披露	お兄さんがロボットダンス <sup>②</sup> を披露	お兄さんがロボットダンス <sup>③</sup> を披露
可視のものとの効果音	正男 お兄さん お姉さん 野原			
不可視のものとの効果音				

図 5: 主人公のお兄さんがロボットダンスをするシーン

#### 4. 基礎的な効果音レトリックの実装

ここで実装するのは、レトリックと言うより、レトリックを構成するための基礎的な効果音の調節、レトリックとしては最も原始的な水準のものであるが、2 節に述べた枠組みに沿って、効果音の取捨選択、キャラクタ動作の効果音のタイミング制御、音量の制御の 3 つを行うシステムを作成した。効果音の取捨選択とは、効果音添付処理において無差別に添付された効果音のうち何れを残し何れを削除するかを決定することである。具体的には、全てのキーによって検索、添付された効果音全てに対して、特定のキーに対応する効果音のみを残す。図 6 を使用し、説明する。

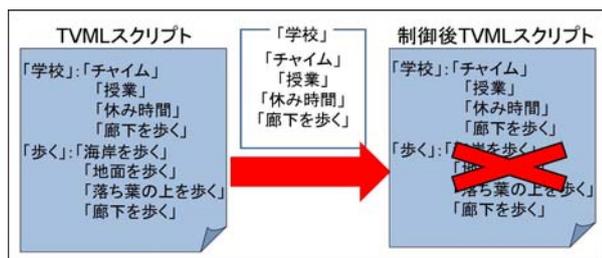


図 6:効果音の取捨選択

TVML スクリプト中で、「学校」という言葉に「チャイム」、「授業」、「休み時間」、「廊下を歩く」の効果音、「歩く」という言葉に「海岸を歩く」、「地面を歩く」、「落ち葉の上を歩く」、「廊下を歩く」の効果音が添付されていた場合、それに対して「学校」に関連する効果音のみを残すと、「学校」から検索された「チャイム」、「授業」、「休み時間」、「廊下を歩く」及び「歩く」から検索された「廊下を歩く」の 5 つとなる。

次に、キャラクタ動作のタイミング制御とは、TVML のコマンドとして用意されている walk(歩く), jump(跳ぶ)のようなキャラクタの動作に対して、動きと効果音のタイミングをもっともらしいタイミングに設定する制御を行う。制御可能なキャラクタの動作コマンドと対応している効果音、タイミングを表 3 に示す。なお、この制御の使用の有無はユーザが決める。

表 3:キャラクタ動作コマンドのタイミング

コマンド	効果音	タイミング
walk	足音	walk, step のコマンドが記述されている上行に書き換える
step		
jump	ジャンプ	jump コマンドが記述されている上行に書き換える
	着地音	jump コマンドが記述されている下行に wait コマンドを挿入する
tumble	倒れる音	tumble コマンドが記述されている下行に wait コマンドを挿入する

最後に、こうして決定された効果音に対して、音量の調節を行う。なお今回は、音量自体は単に大, 中, 小の 3 種類を予め用意した。その方法として、「背景音強調」、「動作音強調」、「指定音強調」の 3 つを用意した。背景音強調は表 2 において「背景」に分類されている効果音の音量を大とし、「動作」に分類されている効果音の音量は小とする。動作音強調はその逆である。どちらを選ぶかはユーザが決める。さらに指定音強調とは、音量をユーザが直接指定できるモードである。

実行例として、図 7 の TVML スクリプトを入力し、「学校」キーに基づく効果音を残し、「背景音強調」を適用した出力例を図 8 に示す。所期の機構はおおよそ達成されたが、評価は今後とする。

```
//ナレーションの実行
narration: talk( who=narrator, text="たみちゃんのかよう学校は、むらのいちばんたかいところ
narration: talk( who=narrator, text="うらは磯だった", caaption="裏は磯だった。")
narration: talk( who=narrator, text="こうていのすみにはさるごやがあって", caaption="校庭")
narration: talk( who=narrator, text="ととして磯からおいだされた。ほすぎるのこんたがず)

.....

narration: talk( who=narrator, text="むれにおわれ", caaption="群れに追われ。", wait=no)
character: shake(name=agent1, state=on, level=3)
wait(time=1)
character: walk(name=agent1, x=-4, d=270, pitch=2)
```

図 7:入力 TVML(一部)

```
.....
//ナレーションの実行
narration: talk( who=narrator, text="たみちゃんのかよう学校は、むらのいちばんたかいところ
sound: playfile( filename=チャイム1.mp3 )
sound: playfile( filename=授業1.mp3 )
sound: playfile( filename=休み時間1.mp3 )
sound: playfile( filename=廊下を歩く2.mp3 )
narration: talk( who=narrator, text="うらは磯だった", caaption="裏は磯だった。")
narration: talk( who=narrator, text="こうていのすみにはさるごやがあって", caaption="校庭")
narration: talk( who=narrator, text="ととして磯からおいだされた。ほすぎるのこんたがず)

.....

narration: talk( who=narrator, text="むれにおわれ", caaption="群れに追われ。", wait=no)
character: shake(name=agent1, state=on, level=3)
wait(time=1)
character: walk(name=agent1, x=-4, d=270, pitch=2)
sound: playfile( filename=廊下を歩く2.mp3 )
```

図 8:出力 TVML(一部)

## 5. おわりに

以上、筆者らの物語生成システムの研究の文脈において、TVML で記述された映像情報に対して効果音を自動添付する方法として、効果音自動添付システムの構想を示し、この枠組みを下に、効果音のレトリックの要素や方法を考察するために映画の調査分析を行い、個々の効果音の特徴を規定する要素を幾つか定め、それに基づいて上記システム枠組みを使用し簡単な試作を実装した。また、映画において個々の効果音がシーンという単位の中でどのように使用されているか、つまり本格的な効果音レトリックの問題にアプローチする端緒として、実例による考察も行った。

しかしながら、そもそも効果音は物語における如何なる情報との関わりにおいて存在するのか、という根本問題の考察に遡って、今後研究を再編成して行く必要があることを認識している。効果音は自然の中の音とは全く水準を異にする音の演出であり、その意味ではそのすべてがレトリックであると言える。今後は 3 節で示したような実際の映像等の分析を本格的に継続し、その知見から物語生成システムにおける効果音のシステムの再構成を行って行く予定である。

## 参考文献

[林 2004]林正樹:めざせ! テレビ番組クリエイター, 芸術評論社, 2004.

[有田 2004]有田元則, 岩宮眞一郎:映像転換時の切り替えパターンと効果音の音列パターンの構造特徴が音と映像に与える影響, 情報処理学会研究報告, 71-76, 2004.

[久保 2003]久保光男:音声素材が効果音に変わる時, 映像情報メディア学会誌 Vol.57, No.8, 956-958, 2003.

[金 2007]金基弘, 岩崎敬吾, 岩宮眞一郎:テロップ・プレゼンテーションにおけることばと効果音の印象の意味的調和の効果, 日本音響学会誌 63 巻 3 号, 121-129, 2007.

[小方 2003a]小方孝:物語の多重性と拡張文学理論の概念—システムナトロジーに向けて I—, In 吉田雅明(編), 複雑系社会理論の新地平, 127-181, 専修大学出版局, 2003.

[小方 2003b]小方孝:拡張性文学理論の試み—システムナトロジーに向けて II—, In 吉田雅明(編), 複雑系社会理論の新地平, 309-356, 専修大学出版局, 2003.

[内海 2007]内海彰, 金井明人:認知修辞学の構想と射程, 認知科学, Vol.14, No.3, 236-252, 2007.