

卒業論文

効果音の探索的検索支援に関する一検討；
効果音の音響特徴に基づく類似性の検証

平成 27 年 1 月

関西大学 総合情報学部

岡本 香帆里

目 次

1	序論	1
1.1	本研究の背景	1
1.2	効果音の検索	1
1.3	効果音検索の問題点	2
1.4	本研究の目的	3
2	関連研究	4
2.1	人がイメージした効果音を表現する方法	4
2.2	効果音検索システムに関する研究	4
2.3	Exploratory Search	5
3	デザイン指針	8
3.1	対象となるユーザ	8
3.2	効果音検索のモデルケースの考察	8
3.3	デザイン指針に基づいた効果音検索システムの概要	10
4	効果音に対応付くオノマトペの収集	12
4.1	実験目的	12
4.2	実験概要	12
4.3	実験準備	12
4.4	実験手続き	13
4.5	収集したオノマトペの統制	14
5	音響特徴量の抽出とクラスタリング	21
5.1	音響特徴量の抽出	21
5.2	音響特徴を用いたクラスタリング	22
6	分析	24
6.1	オノマトペの形態的特徴	24
6.2	音を聴取した際にもつ印象	27
6.3	まとめ	30
7	議論	33
7.1	効果音と対応付くオノマトペの収集方法	33
7.2	抽出した音響特徴量数	33
7.3	デンドログラムを分割する高さ	34
7.4	デンドログラムの分析の観点	34
8	結論	36

1 序論

本章では、本研究の実施に至った背景を説明し、対象とする課題を明確にする。

1.1 本研究の背景

近年、User-Generated Contents¹ (UGC) が整ったことで誰でも簡単にコンテンツの制作・配信が可能になった。UGC 環境で制作・配信されるコンテンツとしては、映像やゲームなどが挙げられる。それらのコンテンツの多くは、視覚情報と聴覚情報で構成されている。視覚情報とは、動画や画像など視覚から獲得する情報のこと、聴覚情報とは、音楽や登場人物のセリフなど耳から獲得する音の情報のことである。聴覚情報の中でも音は、コンテンツの演出の役割を担うため、コンテンツをより際立たせるために重要である。

UGC における音には、歌謡曲、BGM、効果音の 3 種類がある。歌謡曲は、映像のオープニングやエンディング、劇中歌や挿入歌として使用される。BGM は、コンテンツ制作者が演出すべきであると判断した場面の背景に流す音楽として使用される。効果音は、演出に表情をつけることやコンテンツの視聴者に気付きを与えることを目的として使用される。

3 種類の音のうち、特に効果音はコンテンツの演出をする上で重要な役割を担っている。例えばファンタジー映画では、魔法のように現実世界に存在しないものが登場する。魔法が登場する場面では効果音を利用することで、視覚情報との相乗効果により、現実世界に魔法が存在しているかのような感覚や臨場感を与えている。また、アクション映画でよく見られる爆発の場面では、大きな爆発音によって迫力が増している。このように臨場感や迫力を増加させる役割の他にも、効果音はコンテンツの視聴者が視覚からは獲得できない情報を付加する役割も担う。例えば、登場人物が公園を歩く場面に噴水が存在していなくても、「噴水」の効果音を付与することで噴水の近くを歩いているかのような演出をすることが可能である。コンテンツを効果的に演出するためには、その場面に適切な効果音を選ぶことが必要である。例えば、ゆるやかに川が流れている映像に滝の音を使用すれば、映像と音にギャップが生じ、視聴者は違和感を感じるであろう。

効果音は、特定のコンテンツのために作成されることが多い歌謡曲や BGM と違い、1 つの効果音が特定の作品や場面でのみ利用されるわけではなく、同一の効果音が異なる場面や作品で利用されることも多い。例えば、映像制作時に利用した「雨」の効果音を別の映像に利用することや、映像以外のコンテンツを制作する際に利用することも可能である。このように再利用が可能な性質から、効果音はあらかじめ大量に収集され、データベース化されている。データベースを作成しておけば、いつでもその中にある効果音を検索して再利用することができる。

1.2 効果音の検索

コンテンツ制作者（以下、効果音検索者）が効果音を検索する際は、付与すべき効果音のイメージを言語化し、効果音の名称や説明文を頼りに効果音データベースの中から適切な効果音を検索することが一般的である。

効果音データベースとして利用されているものには、例えば、効果音 CD や、効果音検索サ

¹ Web サイトの利用者によって制作・生成されたコンテンツのこと

イトなどが存在する。効果音 CD とは、効果音ばかりが記録されている CD のことであり、記録されている効果音の名称や説明文の一覧を参照することによって探せる。効果音検索サイトとは、音の発生源をクエリとして入力したり、サイト運営者が設定したカテゴリ (e.g., 生活音、人の動作) を選択したりすることで、それらの内容に一致するタグが付与された効果音をまとめて確認できるものである。効果音検索者は、これらのデータベースからいくつか選出した効果音の候補を聴取・比較し、場面に付与すべき適切な効果音を決定する。

データベースから効果音を選出する際、効果音 CD を利用する場合は、効果音の名称や説明文の一覧を参照する。その一覧から、言語化された効果音イメージに一致。または類似した効果音をいくつか選出する。

効果音検索サイトを利用して効果音を検索する場合は、効果音イメージを言語化したものを作りとして入力する方法と、カテゴリを選択する方法の 2 種類の方法が存在する。効果音イメージを言語化したものをクエリとして入力する方法は、入力するクエリとして、音の発生源の他にオノマトペを利用するサイトも存在する²。オノマトペとは、擬音語 (e.g., ザーザー、カチャカチャ) や擬態語 (e.g., ピカピカ、ヒラヒラ) の総称のことである。入力したクエリに一致したタグが付与された効果音が一覧で表示されるため、その一覧から効果音を選出する。カテゴリを選択する方法は、カテゴリを選択すると、そのカテゴリに属する効果音の名称や説明文が一覧で表示される。そのため、効果音検索者は、付与すべき効果音が存在すると推測されるカテゴリを選択し、その一覧から効果音を選出する。

このようにデータベースから各々の方法で選出した効果音を聴取・比較することで、付与すべき効果音を決定する。

1.3 効果音検索の問題点

コンテンツ制作において、1.2 節で述べた方法で効果音を検索することが一般的であるが、効果音 CD や効果音検索サイトでの検索方法では、3 つの問題点があると考える。

(1) 効果音の用途の多様性

効果音は、あらかじめ付与された名称や説明文に記述された用途以外にも使用可能である。例えば、「木の扉をノックする」効果音は「まな板で野菜を切る」効果音として、「スノーノイズ」は「大雨」の効果音として、それぞれ応用可能である。しかし、効果音に付与された名称や説明文からは音の特徴を把握することは難しく、あらかじめ検索候補から除外してしまうおそれがある。

(2) 音情報のテキストによる表現の限界

効果音に付与される名称や説明文に記述されたテキスト情報のみによって音情報を表現することは困難である。例えば、説明文の参照からは効果音の発生源を知ることはできても、その効果音の「響き」や「強さ」といったニュアンスを捉えることはできない。そのため、僅かにニュアンスが異なる効果音を探すためであっても、本来探している効果音と無関係の効果音まで聴取しなければならない。

²<http://taira-komori.jpn.org/freesound.html> (2015 年 1 月 11 日確認) .

(3) 効果音のイメージの明確化における困難性

効果音を検索する際、付与すべき効果音のイメージを明確化することが難しい場合も存在する。例えば、「UFO」や「魔法」などの実在しないものに効果音を付与する場合、効果音検索者は付与すべき効果音イメージのみを頼りに効果音を検索しなければならない。また、効果音検索では検索の過程で効果音のイメージが明確化されていくことも多い。これは、すでに聴取した効果音を手がかりに様々な効果音を探索する過程で、新たな効果音の発見や効果音同士の類似性から新たな効果音の連想が期待されるためである。しかしながら、既存の検索手法では、新たに発見・連想した効果音のためにクエリを再生成し、検索結果として得られた効果音を全て聴取する必要があるため、円滑に探索的検索が行えるとはいえない。

これらの問題は、効果音を選定・比較・決定するまでに聴取する効果音数の総量の増加に通底する。そのため、多大な聴取時間を要することになる。

1.4 本研究の目的

この問題を解決するために、本研究ではオノマトペによって効果音を表現することで効果音同士の類似性の視覚的な提示を可能にし、探索的な効果音検索の容易化を図るシステムの実現を目指す。提案システムでは、効果音同士の類似性を可視化することで、聴取しなければ分からなかつた効果音同士の関係性を視覚的に確認できるようになるため、1つずつ効果音を聴取しなければならない既存の検索に比べて、短時間で効率的な検索の実現が期待される。加えて、効果音を何度も聴取する中で欲しい音が変化したときも、クエリを再生成することなく類似した音を辿ることで効果音を探索的に検索できることが期待される。

本研究で提案するシステムでは、効果音の名称や説明文 (e.g., 利用シーン) の文脈、波形情報の特徴 (以下、音響特徴) の観点に加えて、オノマトペの表象の類似性にも着目した。この3要素により、効果音検索者は文脈やオノマトペの表象の類似性から欲しい効果音を大まかに探し、その後、音響特徴の類似によって効果音の細かいニュアンス (e.g., もっと軽い音、もっと響いている音) を調整するように探していくことが可能になる。これらの作業を、可視化された類似性を辿りながら行うことを可能にすることで効果音の探索的検索を実現する。

本稿では提案システム実現の端緒として、音響特徴量を用いた効果音の階層的クラスタリングと、クラスタリングされた効果音集合に見られるオノマトペの表象情報との関係性について考察する。

2 関連研究

本章では、関連研究を述べる。

2.1 人がイメージした効果音を表現する方法

効果音を検索する際は、付与すべき効果音を言語化して表現し、それをクエリとしてシステムに入力するため、人が音をどのように表現するのかについて知る必要がある。

和気らは、人がどのように音を表現するのかについて、聴取した音を他人に説明するという課題を用いて明らかにしている [12, 3]。サンプル音には、音楽、音声、環境音の3種類の音が使用された。その結果、得られた音の説明は(1)波形情報に関する説明、(2)音の発生源、(3)主観表現を伴う説明、の3種類に分類されている。波形情報に関する説明では、音の物理量の説明に加え、オノマトペも用いられる。音の発生源の説明は、その音は、何の音なのか、どうしたときに鳴るのか、いつ発せられるのか、どんな状況で発せられるのか、に加えて比較表現も用いられる。主観表現での説明は、その音を聴取したときの主観的な感想が述べられ、その感想には主に形容詞が用いられる。オノマトペは、音の種類に関係なく、比較的多く用いられる。

この知見は、効果音検索者がどのように探したい効果音を表現するのかを考える上で有用である。

2.2 効果音検索システムに関する研究

効果音検索者は、付与すべき効果音を言語化したものをクエリとしてシステムに入力・参照し、効果音を検索する。その検索部分に関する研究がいくつか存在する。

Wake らは、オノマトペ、音の発生源、主観表現の3種類をクエリとした効果音検索システムを提案している [3]。効果音検索者は、この3種類のうち、最低1種類以上クエリを入力する。データベース内の効果音に付与されたラベルと入力したクエリとの類似度に応じて検索得点が算出され、得点順に効果音が表示される。検索得点を算出することにより、効果音に付与されたラベルと入力したクエリが一致していないても効果音が提示されることに加え、類似した効果音も提示される。音の発生源により分類されたリストを用いた検索と、和気らのシステムを用いた検索との比較実験により、前者のシステムの方が検索にかかる時間は短いが、後者のシステムの方が発想が広がることが確認されている。

また、青木らは入力したクエリと一致した効果音以外の音も提示するために、類義語と連想語を考慮したシステムを提案している [13]。例えば「爆発」と「爆発音」「ダイナマイト」と「爆発」が、それぞれ類義語と連想語として対応付けられている。これらを対応付けることで、入力したクエリが効果音に付与されたラベルに登録されていない場合でも、効果音が提示される。

清水らは、音響特徴を用いて効果音を分類したシステムを提案している [10]。このシステムは水の音に関する効果音を対象にして、各効果音から音響特徴を33項目取得し、それに基づいて群平均法で分類している。効果音検索者には、その分類を木構造で表示することで効果音が提示される(図2.1)。効果音には、音響特徴により自動でオノマトペが生成されたものが付与される。木構造には、波の音同士でクラスタが形成されたり、激しく水が流れる音同士でクラスタが形成されるなど、効果音の類似性が反映されている。

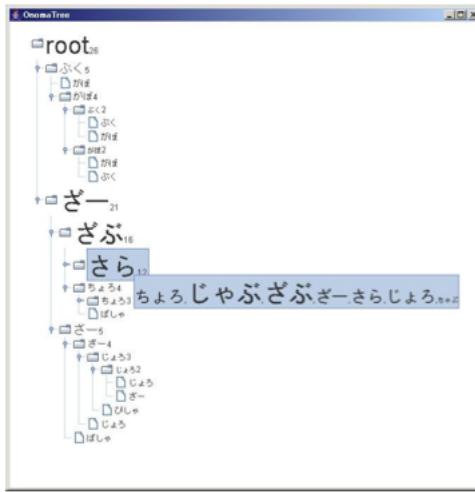


図 2.1: 木構造で分類されている効果音 [10]

2.3 Exploratory Search

効果音検索者は、システムによって提示された効果音同士を比較しながら、付与すべき効果音を決定する。その比較の中では、多数の効果音を聴取するため、付与すべき効果音が変化することもある。そのため、様々な効果音を渡り歩くように聴取できることが望ましい。そのような検索を行うためには、Exploratory Search のモデルが適していると考える。

Exploratory Search [5] とは、検索活動の中で得られる情報や知識を元にして繰り返し検索を行い、不明確な自らの情報要求を明確にしていく情報検索モデルのことである。検索活動を行う中で新たな知識を獲得したり、人が持つ知識を分析、統合、評価、要求することを繰り返したりすることで検索者の知識は増大する。情報を得る度に検索者の情報要求が変化していくことにより、不明確な要求を明確にできる。Exploratory Search は、検索者の要求が不明確であることが重要であり、それを明確にしていくことが課題のひとつである。要求の不明確さは、Taylor により以下の 4 種類に分類されている [2]。

(1) 直感的な要求 (the visceral need)

要求を意識していることも意識していないこともあります、漠然と不満を感じても言葉で表せない状態。この要求は変化する。

(2) 意識されている要求 (the conscious need)

要求を意識しているが明確に表せない状態。この状態の検索者は現状を他人に説明でき、その対話の中で要求を明確にできることもある。

(3) 形式化された要求 (the formalized need)

要求を具体的に説明できる状態。解決手段を考えていたり、考えていなかったりする。他人に説明しなくとも、自らで解決手段を知る可能性がある。

(4) 解決に至る要求 (the compromised need)

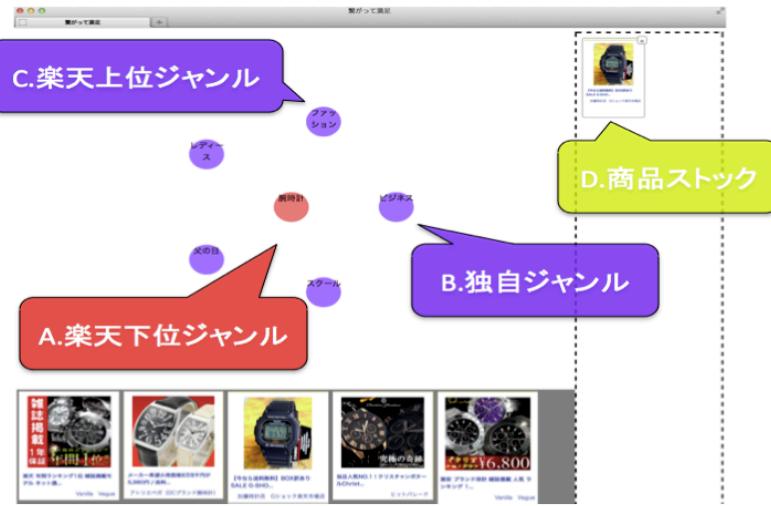


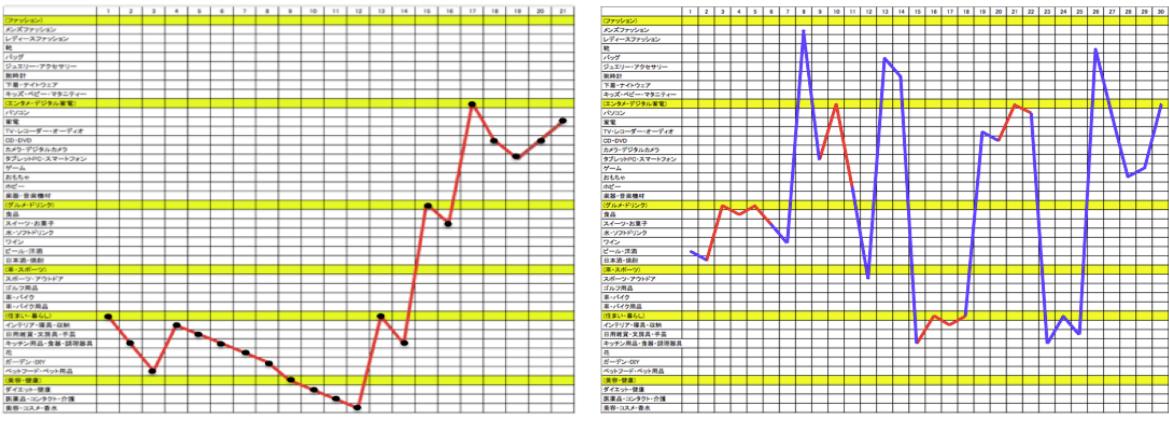
図 2.2: 腕時計を選択した時のシステム画面 (文献 [11] より図引用)

要求を具体的に説明でき、解決手段がわかっている状態。解決するために、何をどう探せばいいのかがわかつっていたり、探し方を変えたりできる。

効果音を検索する際は、1.2 節で述べたように、付与すべき効果音のイメージを言語化しなければならない。しかし、言語化できる要求は、形式化された要求や解決に至る要求のレベルまで明確になっている要求である。1.3 節で述べた (3) の問題点のように、要求を明確化することが困難である場合は、直感的な要求や意識されている要求に分類される。具体的には、効果音が欲しいがどのような効果音が欲しいのかがわからなかったり、欲しい効果音を言葉で表現できなかったりするような要求である。このような要求をもっている検索者は、検索活動を繰り返すことによって、自らの要求を明確にできるであろう。

曖昧なものを明確化していく必要があるものとして、プレゼント選択がある。プレゼントは、相手の趣味や嗜好を考慮し、喜んでもらえるものを選択する必要がある。しかし、相手の趣味や趣向が明確でない場合は、それらを推測し、明確にしなければならない。

盆子原らは、EC サイトにおける商品探索プロセスに着目したプレゼント探索に Exploratory Search モデルを取り入れたシステムを提案している [11]。EC サイトは、商品にジャンルを設定することで整理されている。そのジャンルは、具体的な商品を表わす下位ジャンル (e.g., 腕時計) が、意味や内容を受け継ぐ上位ジャンル (e.g., ファッション) に繋がっており、それを次の階層に繋げていくようなツリー構造が成されている。しかし、このツリー構造で商品を検索する際は、上位ジャンルから下位ジャンルの商品を閲覧し、下位ジャンルに要求を満たすものが存在しなかった場合、再び上位ジャンルへ遷移して他のジャンルを閲覧しなければならないとされている。これは、商品をいくつも閲覧したり、探索したりするプレゼント選択においては不向きであるため、提案システムでは、EC サイト (楽天市場) の上位ジャンル、下位ジャンルに加え、使用用途や使用場面を想起させる内容の独自の上位ジャンル (e.g., ビジネス、父の日) が新たに設定されている (図 2.2)。この独自の上位ジャンルにより、多様な観点から商品を探せるようになっている。さらに、商品ストックを設けることで、興味を持った商品を



(A) 従来型のシステム

(B) 盆子原らのシステム

図 2.3: 実験結果の例 (文献 [11] より図引用)

ストックし、見比べながら商品を選択できるようになっている。

ツリー構造が用いられた従来型のシステムと盆子原らのシステムとを用いて、「身近な人のためにプレゼントを購入する」という課題が課された比較実験が行われた。この実験結果から、従来型のシステムよりも盆子原らのシステムを用いる方が、幅広いジャンルを遷移しつつ商品探索が行われることが示されている(図 2.3)。このことから、Exploratory Search モデルを取り入れると、検索結果を多岐に渡って遷移しながら効果音を検索できることが推測される。

3 デザイン指針

本章では、効果音検索システムのデザイン指針について述べる。

3.1 対象となるユーザ

提案システムのデザイン指針を立てるにあたって、探したい効果音を曖昧な表現でしか表出できない検索者（以下、検索者）をシステムの対象ユーザとする。以下、3.2 節にて効果音検索のモデルケースの考察からデザイン指針を定め、3.3 節にてデザイン指針に基づいた効果音検索システムの概要を述べる。

3.2 効果音検索のモデルケースの考察

図 3.1 に、検索者が効果音を検索する様子の一例を示す。この例では、Aさんは「足音」の中でも「歩く」という効果音に焦点を当て（Phase 1）、聴取することで「走る」という効果音の検索に移行し、「もっと軽い音」のように細かいニュアンスの調整を図った（Phase 2）。その後、付与したい効果音が見つけられなかつたため、データベース内全体を俯瞰し、要求をオノマトペで表現し、そのオノマトペに類似した効果音を選択するという段階を踏んでいる（Phase 3 から Phase 4）。この例では、最終的に検索者のイメージに合致した効果音は、検索開始時に想定していた効果音とは異なるものであった。また、効果音を聴取する過程で検索者自身により付与する効果音のイメージが明確化されている。そのため、効果音を最終的に決定するまでの過程で、「歩く」「走る」「足音・コミカル」という 3 つのカテゴリ、合計 6 つの効果音を聴取しなければならなかつた。検索者のイメージに合致した効果音を円滑に検索するためには、聴取前に効果音のニュアンスや効果音同士の類似性を把握可能であることが望ましい。

以上の考察から、本稿では効果音検索システムに必要となるデザイン指針を以下の 5 点に集約した。

- データベース内から効果音を絞り込み可能であること
- 効果音の細かいニュアンスで検索可能であること
- 聽取前に、効果音のニュアンスが把握可能であること
- 類似した効果音を示唆可能であること
- オノマトペによる検索要求に対応可能であること

本研究では、これらのデザイン指針を満たす効果音検索システムの実現を目指す。これらのデザイン指針のうち、特筆すべき点として、オノマトペを用いた効果音の表現 [12] が挙げられる。オノマトペとは擬音語・擬態語の総称であり、感覚的である一方で繊細かつ微妙な描写を可能にする言葉もある [7]。また、音響特徴とオノマトペの特徴および音の聴取によって感じる印象の 3 者には対応が見られ、音に起因する印象や音響的特徴を捉える上でオノマトペが有効な表現手段となり得ることが示されている [14]。これらのことから、効果音検索者はオノマトペを用いることで、検索目的の効果音を感覚的かつ具体的に検索可能となり、効果音の聴取前におおよその効果音のニュアンスを把握できるようになると期待される。

Aさんは新しいアニメーションを制作している。今回制作しているアニメーションでは、主人公に架空の小さい生き物が同行している。

—Phase 1—

Aさんは、その架空の生き物の足音の効果音を探すために、効果音検索サイトを開いた。まずAさんは「人の動作」と書かれたカテゴリから、「歩く」という音に注目した。森の中を歩いている場面に付与するため、「砂利の上を歩く」という音を聴いた。その音を聴いたAさんは、架空の生き物は小さいため、その音では歩幅に合わないと考えた。

—Phase 2—

そこで「走る」という音から探すことにした。Aさんは「玉砂利を走る」という音を聴いたが、「もっと軽い音が良いし、こんなにジャリジャリして欲しくない」と思い、「土、砂を走る」という音を聴いた。しかし、音の軽さが同じだった。

—Phase 3—

Aさんは「架空の生き物だから、実際に出る足音でなくても良いかもしない」と考え、他にどのような足音があるのかを見てみることにした。Aさんは効果音の一覧を眺める中で、「足音・コミカル 01」という音を見つけた。しかし、「ブニブニして気持ち悪い。ポンポン弾んでいる音が良い」と思ったため、「足音・コミカル 02」という音を聴いたが、その音からは重たい生き物が歩いているという印象を受け、探している音とはかけ離れた音だった。

—Phase 4—

次に、「足音・コミカル 03」という音を聴いたところ、「ポンポン弾んでいる音ではないけど、こっちのポコポコしている音の方が合うかもしれないし、これなら歩く速さにも合うから使える」と確信し、その音に決めた。

図 3.1: 探したい効果音が不明確な検索者が効果音を検索する様子の一例

表 3.1: 効果音の表現に用いる観点の詳細説明とその活用目的

観点	詳細説明	効果音検索での活用目的
文脈	効果音のタイトルや説明文。説明文には、効果音の発生源や使用場面が記載される。	音の発生源に関する因果を捉える。効果音の大まかな絞り込みに用いる。
音響	効果音の音響特徴。音響波形から抽出される特徴量。	細かなニュアンスを捉える。効果音のニュアンスを微調整するような検索に用いる。
オノマトペの表象	オノマトペの情報。具体的に、破裂音・濁音などの音の種類や、ABAB 型などの語形特徴。	効果音を視覚的に捉える。効果音の大まかな絞り込みや効果音同士の類似性を瞬時に把握するために用いる。

3.3 デザイン指針に基づいた効果音検索システムの概要

これまで、効果音は「文脈」や「音響」のみによって表現されることが多かった。本研究では、前節で定めたデザイン指針に則り、オノマトペによって効果音を表現することで、効果音検索に対して新たに「オノマトペの表象」という概念を取り入れる。これにより効果音同士の類似性は、「文脈」「音響」「オノマトペの表象」という異なる 3 種類の観点で表現可能になる。提案システムにおいて、効果音の表現に用いる「文脈」「音響」「オノマトペの表象」それぞれの詳細説明と効果音検索における活用目的をまとめたものを表 3.1 に示す。このように、活用目的に応じて 3 種類の観点を使い分けることで、直感的かつ円滑な効果音の探索的検索の実現を図る。提案システムは、図 3.2 のように、効果音同士の関係性が 3 種類の観点それぞれによって構成されたデータベースを基に構築される。図 3.2 では、中心の青いノードに対して、3 種類の観点それぞれの類似性によって効果音同士の関係性が構築されている。同図から「ザーザー雨(1)」は、「オノマトペの表象」の観点で見ると「サーサー雨(1)」と類似していること、「文脈」の観点で見ると「バリバリ雷」と類似していること、「音響」の観点で見ると「ザッパーん海」と類似していることが見てとれる。提案システムは、効果音同士の類似性を視覚的に確認しながら探索的に検索することを可能にする。そのため、検索者は単に効果音を検索するだけでなく、1) 「この効果音はこのクエリを入力すれば検索できる」という効果音の検索方法の学習、2) 「P 音が使用できるならば、P 音に類似した Q 音も使えるのかもしれない」という連想、3) 「このような音もあるのか」という新しい効果音の発見、といった二次的な経験も得ることができる。また、図 3.2 のようなリンク構造を持つデータベースを用いることで、効果音を検索する中で探したい効果音が変化した場合であっても、効果音の類似性を辿ることでクエリを再生成することなく効果音の検索が可能になると考える。

提案システムを構築するためには、効果音とオノマトペを対応付けた上で「文脈」「音響」「オノマトペの表象」の 3 種類の観点それぞれでの類似性の提示技術が必要となる。類似性の提示技術には、表 3.2 の情報を使用する。

本稿ではその端緒として、音響特徴によって階層的にクラスタリングされた効果音に付与さ

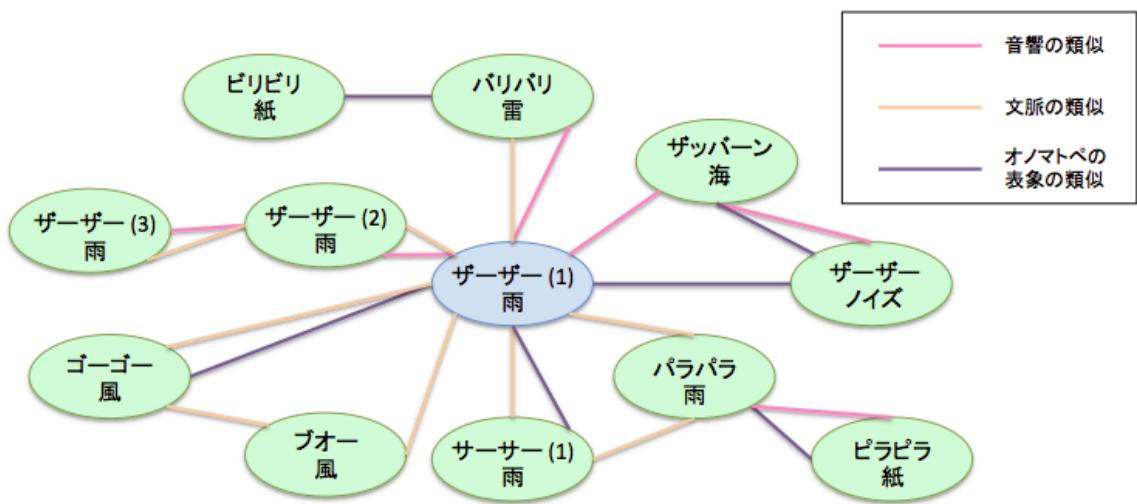


図 3.2: 効果音同士の関係性の模式図 . 効果音同士の類似性が「文脈」「音響」「オノマトペの表象」という異なる 3 つの観点でそれぞれ表現されている .

表 3.2: 類似性提示に使用する情報

観点	類似性提示に使用する情報
文脈	効果音の名前や音の発生源の共起情報
オノマトペの表象	オノマトペの音の種類や語形特徴
音響	効果音の音響特徴

れたオノマトペを参照し，音響の類似性とオノマトペの表象との関係性を考察する．以下，オノマトペと効果音の対応付けについては 4 節，音響特徴に基づく効果音のクラスタリングについては 5 節にてそれぞれ詳細を述べる．

4 効果音に対応付くオノマトペの収集

本章では、分析で扱うオノマトペの収集方法について述べる。

4.1 実験目的

文献 [14]において、音の印象や音響特徴を把握する上でオノマトペによる表現が有効であることが示されている。オノマトペはシグナル情報である音響特徴を、ヒトが理解し易いシンボル情報に置き換えたものであると捉えることができる。効果音にオノマトペを対応づけることで、temporal な聴覚情報である音を spatial な視覚情報に変換可能となる。視覚情報に変換することで、聴取するだけではわからなかった効果音の類似性を字面で確認可能にすることを狙う。

そこで、本稿で扱う効果音にオノマトペを対応づけるため、聴取した効果音に対応するオノマトペを自由記述で回答させる主観評価実験を行った。

4.2 実験概要

本実験では多数の協力者から回答を得るために、作成した実験プログラムをインターネットで公開して行った。実験プログラムの詳細については、4.3.2 節で述べる。実施期間は 2014 年 10 月 26 日～2014 年 11 月 25 日の 31 日間とし、114 人からの回答が得られた。

実験では、まず、実験協力者はオノマトペと実験手順について理解をした後、効果音の聞き逃がしを防止するためにイヤホンの装着を求められた。次に、実験協力者に効果音を提示し、「他人がその効果音を聴いた場合、どのようなオノマトペで表現するのか」を回答させた。回答は自由記述形式とし、1 つの効果音に対して複数のオノマトペを回答可能とした。他人の表現方法を実験協力者に回答させたのは、個人によるオノマトペの表記揺れを回答段階において、ある程度統制するためである。

実験協力者の負担を考慮し、一人の実験協力者には選定した効果音 100 データを 5 グループに分割した中の 1 グループ（20 個）に回答してもらうこととした。効果音の選定については、4.3.1 節で述べる。

4.3 実験準備

4.3.1 効果音の選定

本実験で使用する効果音データベースの作成にあたり、インターネット上の効果音配布サイト³から 100 データを選択した。選択にあたっては、各タイトル・各音源につき 1 データずつ効果音を選定した。例えば、「ドアをノックする」という効果音であれば「鉄製のドア」と「木製のドア」から 1 データずつ効果音を選定した。「鉄製のドア」が 3 種類存在するなど、同一音源の効果音が複数存在する場合は、その内の 1 データをランダムに選定した。この方法で効果音 70 データを用意し、残りの 30 データはそれまでに選定されていない効果音の中からタイトル・音源の重複を考慮せずランダムに選定した。選定した効果音リストを、表 4.1 に示す。

³<http://taira-komori.jpn.org/freesound.html> (2015 年 1 月 11 日確認)。

4.3.2 実験プログラムの作成

本プログラムは、効果音に対応付くオノマトペを収集するために作成された。実装は、HTML、CSS、JavaScript、PHP を用いて行った。図 4.6 の画面が表示された際に、0 ~ 4 までの数字が割り当てられた効果音グループ 5 つの中からランダムに 1 つが選出され、そのグループに属する効果音 20 データがランダムな順序に並び替えられた（図 4.1）。並び替えられた効果音は、先頭から順に実験参加者に提示された。

収集データの管理を容易にするために、実験協力者ごとに ID を振り当てた。ID には、選出されたグループ番号と、図 4.6 の「実験を始める」ボタンが押下された時刻を使用した。例えば、選出されたグループ番号が「4」で、時刻が「11 月 15 日 18 時 29 分 30 秒 195 ミリ秒」であった場合、振り当てられる ID は「4_11_15_18_29_30_195」となる。

実験協力者が「実験を始める」ボタンを押下した際、「振り当てられた ID.txt」という名前のファイルが作成される。その txt ファイルには、「再生された効果音名：実験協力者が回答したオノマトペ：図 4.6 の「次へ」ボタンが押下された時刻（時：分：秒）」が、同ボタンが押下される度に書き込まれる（図 4.2）。「次へ」ボタンを押下した時刻を取得したのは、実験 1 回当たりのおおよその所要時間を把握するためである。

4.4 実験手続き

実際に使用した実験プログラムの画面と併せて、実験協力者が行った手続きを述べる。

(1) 実験に同意する（図 4.3）。

同意書を読み、「同意する」にチェックをした後「OK」ボタンを押下する。

(2) オノマトペについて理解する（図 4.4）。

オノマトペについての説明を読み、理解できれば「次へ」ボタンを押下する。

(3) 実験手順について理解する（図 4.5）。

実験協力者自身が行う手順について理解する。理解できれば「次へ」ボタンを押下する。

(4) 実験準備をする（図 4.6）。

実験協力者は、手順に沿って実験準備をする。まず、実験協力者はイヤホンを装着する。次に、再生ボタンを押して音量の確認を行う。音量調節は、パソコンの本体側で行う。準備ができれば「実験を始める」ボタンを押下する。

(5) オノマトペを回答する（図 4.7）。

実験画面が表示されると、自動的に効果音が再生される。実験協力者は、再生された効果音を聴取し、その音の一般イメージを a のテキストボックスに入力する。入力が完了すれば「次へ」ボタンを押下する。ボタンを押下すると、次の音源が再生される。テキストボックスに何も入力が無ければ、次の音源に遷移しない。b の再生ボタンを押下することにより、何度も効果音を再生できる。実験協力者は、この手順を試行終了まで繰り返す。現在の試行回数は、c に表示される。d のヘルプにより、A) オノマトペ、B) 実験概要、C) 操作方法、をそれぞれのタブをクリックして切り替えることで、いつでもその内容を確認できる。ヘルプには、それぞれ以下の内容が記載されていた。

- A) この効果音を他の人が聞いた時、他の人はどのようなオノマトペでこの効果音を表現するのかを予想し、それをなるべくたくさん回答してください。効果音は何度聞き直しても構いません。
- B) オノマトペとは、例えば「カエルが“ぴょん”と飛んだ」の“ぴょん”という擬態語や「ライオンが“ガオー”と鳴いた」の“ガオー”という擬音語の総称のことです。
- C) 回答をフォームに入力した後、「次へ」ボタンを押すと次の効果音へ移ります。何も記入がなければ、ボタンを押しても次の音源へ移りません。

(6) 実験の終了(図 4.8) .

全試行が終了すると、実験が終了したことを伝えられる。

4.5 収集したオノマトペの統制

この実験で収集されたオノマトペは、表記揺れを統制するために表 4.2 の規則に従って前処理が施された。前処理は、クリーニング、正規化、標準化の順で行われた。本実験の目的は効果音に対応付くオノマトペラベルの収集であったため、クリーニングでは、オノマトペ以外の回答が削除された。正規化では、重複したオノマトペ末尾の促音の除去と、重複した母音の 2 つ目以降を長音符に置換を行った。標準化では、個人ごとに音の長さを表現するために使用する長音符の数が異なると考え、長音符を 1 つとした。

その後、ある効果音に対して最も共通して付与されたオノマトペを、その効果音を表すオノマトペと仮定した。その際、ひらがな、カタカナの表記の違いは考慮しなかった。

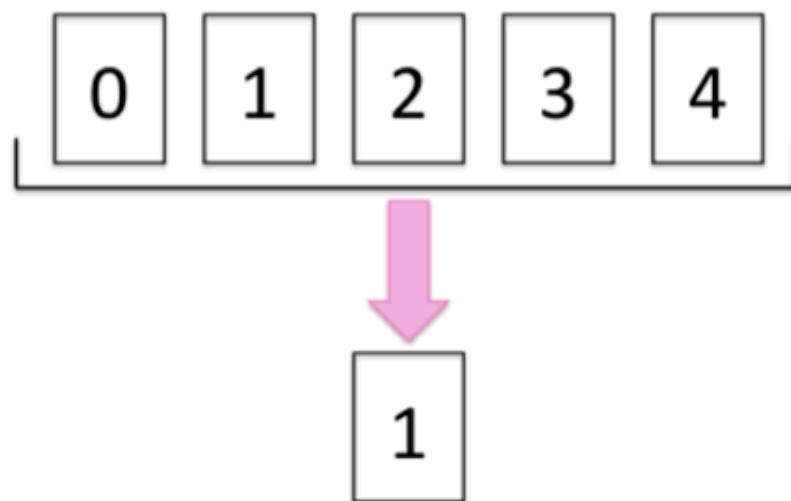
表 4.1: 選出された効果音 100 データ

音源名			
answer_01	door_chime_01	noise_01	rain_01
attack_02	door_chime_02	opening_01	reciever_01
autoshaver_01	door_key_01	opening_02	roll_02
bell_02	door_knock_01	opening_04	roll_04
bell_03	door_knock_02	opening_05	scissors_01
bell_04	door_knock_06	opening_06	sello_tape_01
boiling_01	drinking_01	opening_08	shaking_01
boiling_02	eating_01	paper_01	spray_01
broadcast_stopping_01	eating_02	paper_03	spray_02
broom_01	eating_03	paper_07	staple_01
brushing_01	electric_01	phoning_01	striking_01
bubble_01	explosion_03	phoning_03	swish_01
button_01	fabric_01	phoning_04	switch_02
camera_01	falling_01	polishing_01	tea_with_a_spoon_01
clapping_01	fire_01	pot_01	tissue_01
click_01	ice_water_01	pouring_01	toaster_01
clinking_01	key_01	power_point_01	underwater_01
clinking_02	keyboard_01	pulling_01	vibration_01
clinking_03	keyboard_02	punch_01	water_01
clock_01	kick_01	putting_03	water_07
coin_01	lighter_01	putting_04	water_11
cork_01	magic_tape_01	putting_06	wave_01
crash_01	match_04	putting_09	zenmai_01
cutter_01	money_box_01	putting_11	zipper_01
cutting_01	motor_01	putting_15	zipper_02

表 4.2: 収集されたオノマトペの表記揺れの統制ルール

処理	統制前	統制後
クリーニング	雷	削除
クリーニング	ホッキスをぱちんととめる	ぱちん
クリーニング	? や。などの記号	削除
正規化	カチツツ	カチツ
正規化	ざあああああん	ざああーーーん
正規化	グオウアアアん	グオウアーネン
標準化	ざああーーーん	ざああーん

① 5つの効果音グループからランダムに1グループが選出される



② 効果音グループ内の効果音がランダムに並び替えられる

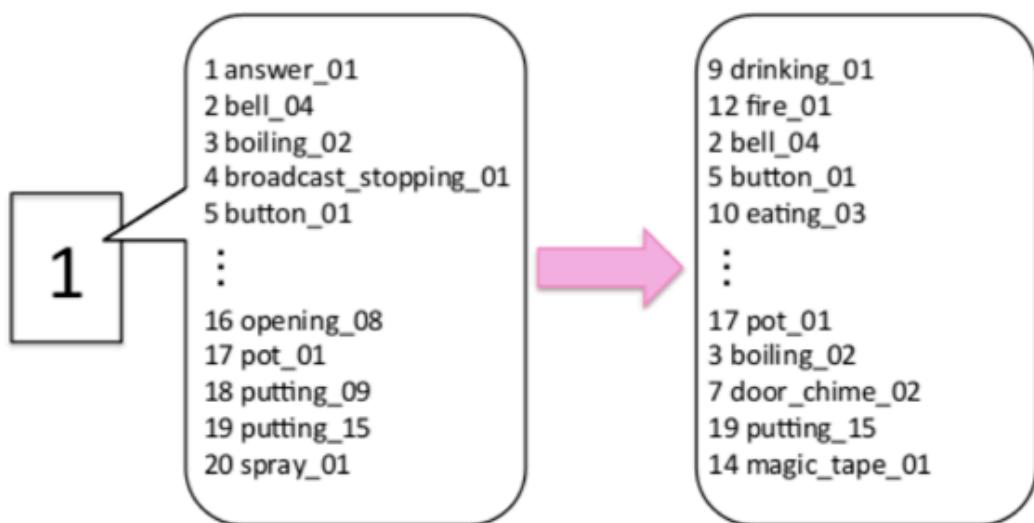
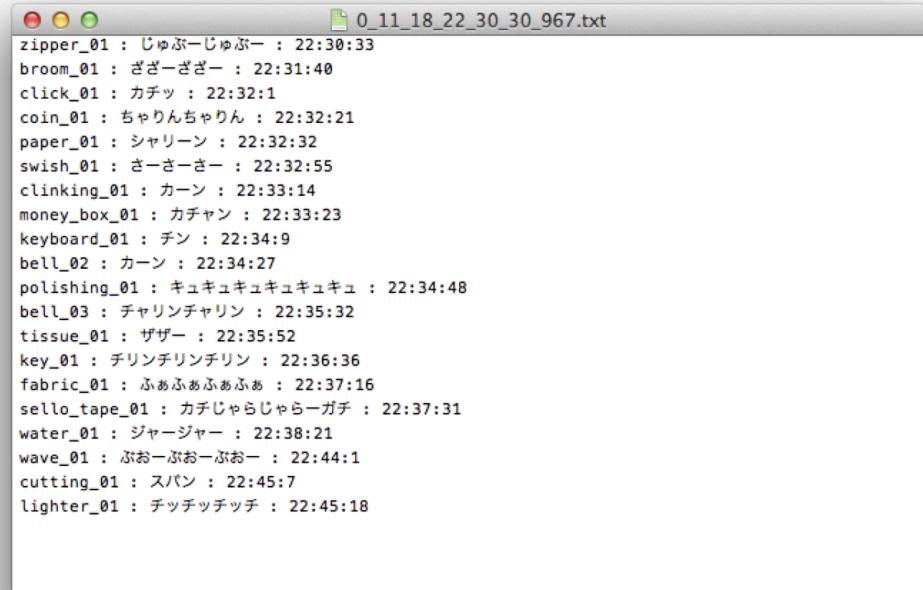


図 4.1: 「実験を始める」ボタンが押下されてから実験協力者に最初の効果音が提示されるまでの処理



0_11_18_22_30_30_967.txt

```
zipper_01 : ジュブー・ジュブー : 22:30:33
broom_01 : ざざーざざー : 22:31:40
click_01 : カチッ : 22:32:1
coin_01 : ちゅりんちゅりん : 22:32:21
paper_01 : シャリーン : 22:32:32
swish_01 : さーさーさー : 22:32:55
clinking_01 : カーン : 22:33:14
money_box_01 : カチャン : 22:33:23
keyboard_01 : デン : 22:34:9
bell_02 : カーン : 22:34:27
polishing_01 : キュキュキュキュキュキュ : 22:34:48
bell_03 : チヤリンチャリン : 22:35:32
tissue_01 : ザザー : 22:35:52
key_01 : チリンチリンチリン : 22:36:36
fabric_01 : ふあふあふあふあ : 22:37:16
sellotape_01 : カチジャラジャラーガチ : 22:37:31
water_01 : ジャージャー : 22:38:21
wave_01 : ぶおーぶおーぶおー : 22:44:1
cutting_01 : スパン : 22:45:7
lighter_01 : チッチッチッチ : 22:45:18
```

図 4.2: 実験協力者ごとに作成された txt ファイル例



図 4.3: 同意書画面



図 4.4: オノマトペの説明画面



誰かに相談せずに回答は一人で行ってください。

次へ

図 4.5: 実験手順説明画面



1. イヤホンの準備

イヤホンを装着してください

2. 音量チェック

本体側で音量を調節してください

再生ボタンを押すと音を再生します

※音量注意※

3. 準備完了

準備が完了したら下のボタンを押してください

実験を始める



図 4.6: 実験準備画面

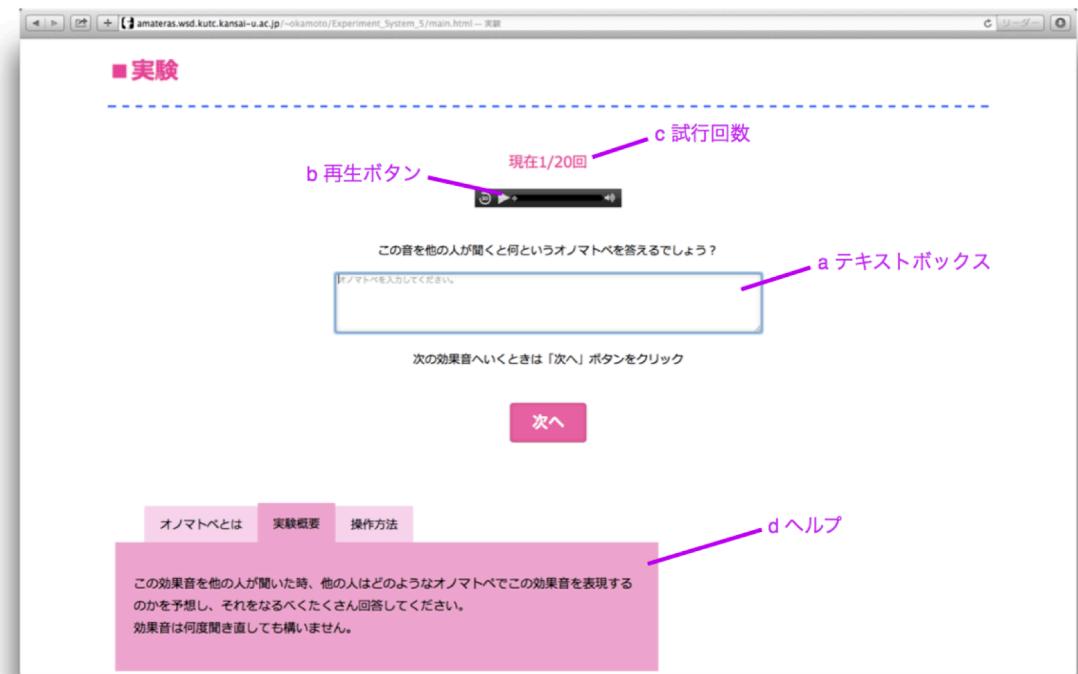


図 4.7: 実験画面

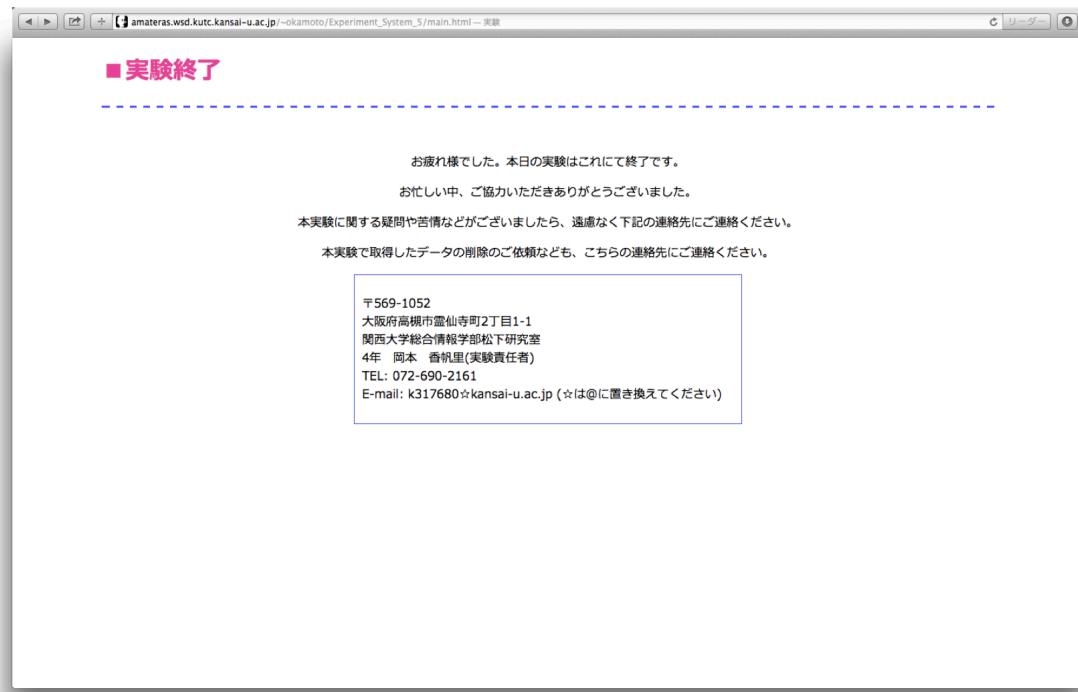


図 4.8: 実験終了画面

5 音響特徴量の抽出とクラスタリング

本章では、効果音の音響特徴量の抽出とクラスタリング手法について述べる。

5.1 音響特徴量の抽出

音楽情報処理の研究分野では、音響特徴量を用いることで、ヒトが認知する類似性に則した音楽の分類が可能であることが知られている [9]。

音響的特徴量でクラスタリングされた効果音集合について、4節にて対応づけられたオノマトペを考察することで、音響特徴に基づいたクラスタとオノマトペの形態的特徴・音象徴との関係性を検証する。

音響特徴量の抽出には、音楽情報処理ツールである MIRtoolbox⁴[1] を用いた。MIRtoolbox とは、数値解析ソフトの MATLAB⁵ 上で動作するツールであり、音声ファイルから調性やリズム、及び構造のような音楽的特徴を抽出することができる。本稿では、印象的側面 [14] という音色の特徴の一面に基づき、特徴量を抽出した。音色の印象的側面とは、形容詞で音色の特徴を表現できる性質のことである。印象的側面は、3ないし4の独立した音色因子に集約できると考えられており、代表的な因子は美的因子、金属性因子、迫力因子といわれている [14]。以下では、3つの因子と抽出した音響特徴について説明する。

- 美的因子

美的因子は、“澄んだ—濁った”や“綺麗—汚い”といった表現語対が対応する。一般的に、不協和音が少ないと澄んだ音、不協和音が多いと濁った音になるといわれているため、美的因子の“澄んだ—濁った”は不協和音の多さと言い換えることができる。不協和音の多さは、和音の中心の音であるルート音に従わない音の量から、おおよその量を測れるため、特徴量 Inharmonicity を効果音から抽出した。

表 5.1: 抽出した音響特徴

番号 i	特徴名	説明
1	RMS energy	音量
2	Low energy	弱音の割合
3	Tempo	テンポ
4	Zero cross	波形が 0 値をとる回数
5	Roll off	85%を占める低音域の割合
6	Brightness	1500Hz 以上の音域の割合
7	Inharmonicity	ルート音に従っていない音の量
8	Mode	major と minor の音量の差

⁴<http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
(2015年1月23日確認)。

⁵<http://www.mathworks.com/> (2015年1月23日確認)。

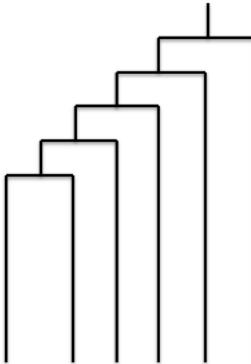


図 5.1: 鎖効果の例

- 金属性因子

金属製因子は，“鋭い—鈍い”や“固い—柔らかい”といった表現語対が対応する。一般的に、低音が少なければ軽い・やわらかい音と表現され、低音が多くなると重い音と表現されることが多い。そのため、低音の割合を示す特徴量 Roll off、高音域の割合を示す特徴量 Brightness を効果音から抽出した。

- 迫力因子

迫力因子は，“弱々しい—力強い”といった表現語対が対応する。迫力因子は音の大きさで測ることができると考えたため、音量を示す特徴量 RMS energy と、弱音の割合を示す特徴量 Low energy を効果音から抽出した。

音色の印象的側面に加え、文献 [6] を参考にして、効果音 1 データにつき表 5.1 に示した 8 つの音響特徴量 f_i をそれぞれ抽出した。ただし、周期性がない効果音についてはテンポを示す特徴量 f_3 を抽出できないため、 $f_3 = 0$ とした。

抽出した音響特徴同士はそれぞれ値域が異なる。そこで、効果音 s について観測された各特徴量 $f_i(s)$ を、式 (5.1) により 0 ~ 1 の範囲に正規化した $f'_i(s)$ を特徴量として扱う。

$$f'_i(s) = \frac{f_i(s) - f_i^{\min}}{f_i^{\max} - f_i^{\min}} \quad (5.1)$$

ここで、 f_i^{\min} と f_i^{\max} は、実験に用いた全ての効果音から得られた特徴量 f_i の最小値と最大値を示す。

5.2 音響特徴を用いたクラスタリング

効果音の音響特徴を用いてクラスタリングすることで、音響特徴が類似した効果音同士の関係性を確認する。階層クラスタリングでは、類似度ごとに分類された結果がデンドログラム(樹形図)として出力されるため、視覚的に関係性を確認することができる。また、デンドログラムを任意の高さで分割することで、音響特徴の類似度に応じて形成されたクラスタ群に効果音を分類できる。デンドログラムのラベルを 4 節で得たオノマトペにすることで、聴覚情報であ

る効果音を視覚情報とした確認が可能になる。それにより、効果音を聴取した際に生じる個人による印象の違いを考慮することなく、一意に定まったオノマトペの字面により効果音の類似性を確認できる。

本稿では、抽出した音響特徴を用いて、Ward 法 [4] で効果音を階層的にクラスタリングした（式 5.2, 5.3）。

$$d(C_1, C_2) = E(C_1 \cup C_2) - E(C_1) - E(C_2) \quad (5.2)$$

ただし、クラスタ C_i の中心を \bar{x}_i として

$$E(C_i) = \sum_{x \in C_i} (x - \bar{x}_i)^2 \quad (5.3)$$

Ward 法とは、クラスタ分析手法の 1 つである。この手法では、クラスタ内の平方和が最小となるようにクラスタが形成されていく。図 5.1 のような、あるクラスタに分類対象が 1 つずつ吸収されてクラスタが形成されていく鎖効果という現象が起こりにくいことも特徴である。鎖効果が起こると、どの距離でクラスタを分割しても、どのクラスタ群にも分類されない単独の効果音が存在してしまう。これを避けるために、本稿では、Ward 法を採用した。

6 分析

本章では、分析について述べる。

本研究が企図するシステムでは、「文脈」「オノマトペの表象」「音響」の3種類の観点での探索を提供する。探したい効果音が不明確な検索者は、各観点の類似性を渡りながら効果音を探索することで、漠然とした効果音のイメージを徐々に明確にしていくことが可能になる。

このような探索のためには、構成される効果音のクラスタは観点ごとに最適化されているべきである。その一方で、「音響」と「オノマトペの表象」については、擬音語が音自体を言語化するものである以上、一定の相関が期待される。そのため、効果音探索におけるアクセシビリティの向上には、音響特徴によって得られたクラスタを言語化した場合、そのクラスタ内の効果音が共通点を持ちながらも多様な形態のオノマトペで構成されていることが望ましいと考えている。

この点について検証するため、5節によって得られたデンドログラムを2つの観点で分析する。本稿の分析では、「単一の効果音のみで構成されるクラスタが存在しない」という条件を満たすようにクラスタ間の高さを調整して、デンドログラムを分類した。その結果、100個の効果音は、図6.1のように9群に分類された。この群ごとにそれぞれの特徴を考察していく。

6.1 オノマトペの形態的特徴

それぞれの群がどのような形態のオノマトペから構成されているのかを知るために、オノマトペの形態的特徴から比較を行った。表記揺れを除くための前処理として、オノマトペを共通の記号に置き換えた。具体的には、第1音をA、第2音をBと表現し、それが交互に現れる場合を“繰り返し型”的オノマトペとした。ここで、流音(ら行)はうなり音など早い振幅変動を表現するために多用されるため[14]、ら行については、Rと分けて表現することとした。また、促音・撥音・長音符はそのまま用いることとした。前処理を行った後のデンドログラムを図6.2に示す。これらを表6.1のように評価していく。項目の要素を含んでいるオノマトペが群内の50%を超えた項目を表6.2に示す。

この表より、まず、どの群についても50%を超える項目が存在することから、各群にdominantな形態的特徴が存在していることが示唆された。ただし、促音に関してはどの群においても50%

表 6.1: オノマトペの評価例

オノマトペ例	記号変換例	分類項目
ザー	A -	長音符
コツ	A つ	促音
シャキン	AB ん	撥音
ブルルル	ARRR	R
カチャカチャ	ABAB	繰り返し
チリン	AR ん	R, 撥音
チャリンチャリン	AR ん AR ん	R, 撥音, 繰り返し

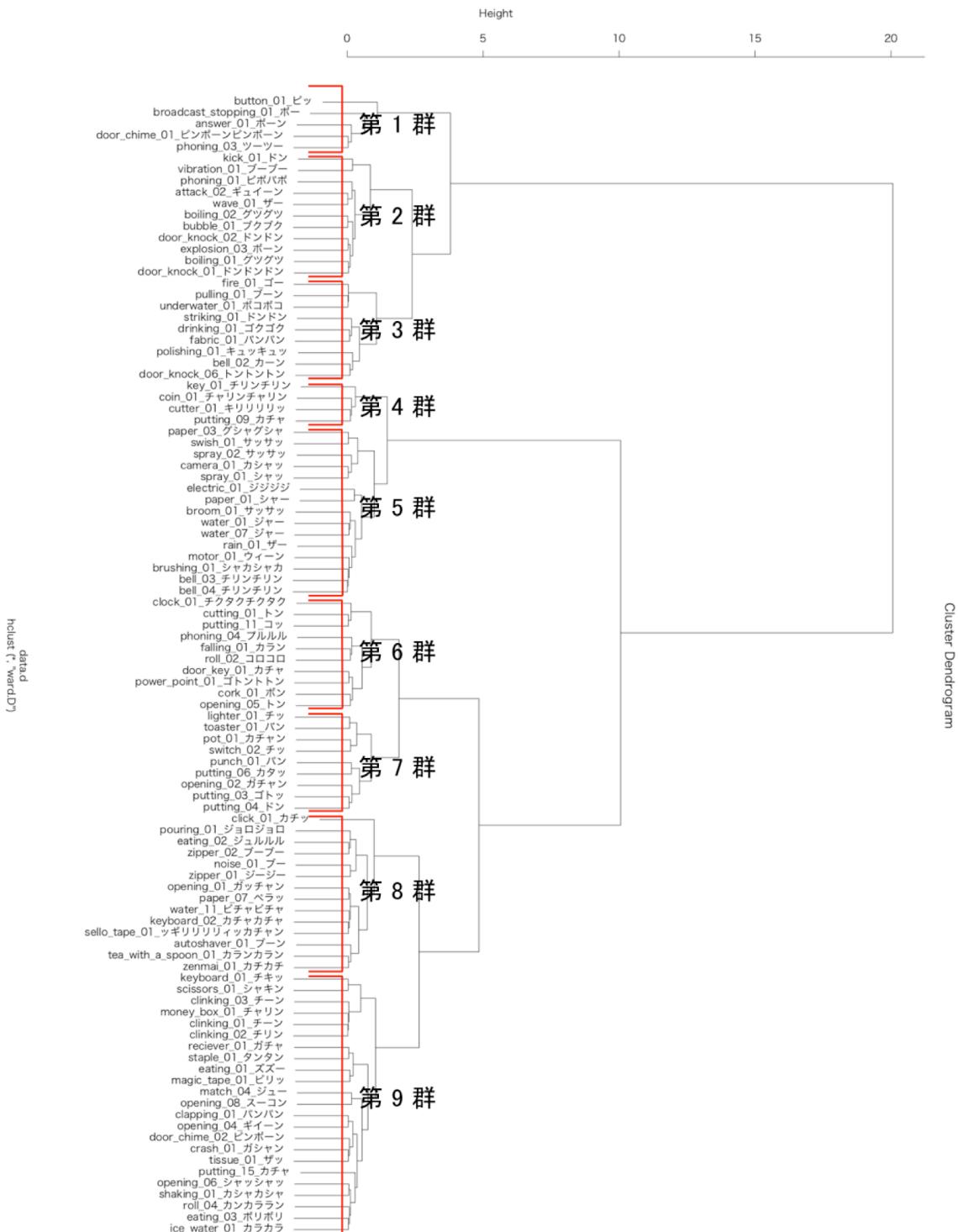


図 6.1: 全 9 群に分類されたクラスタリング結果

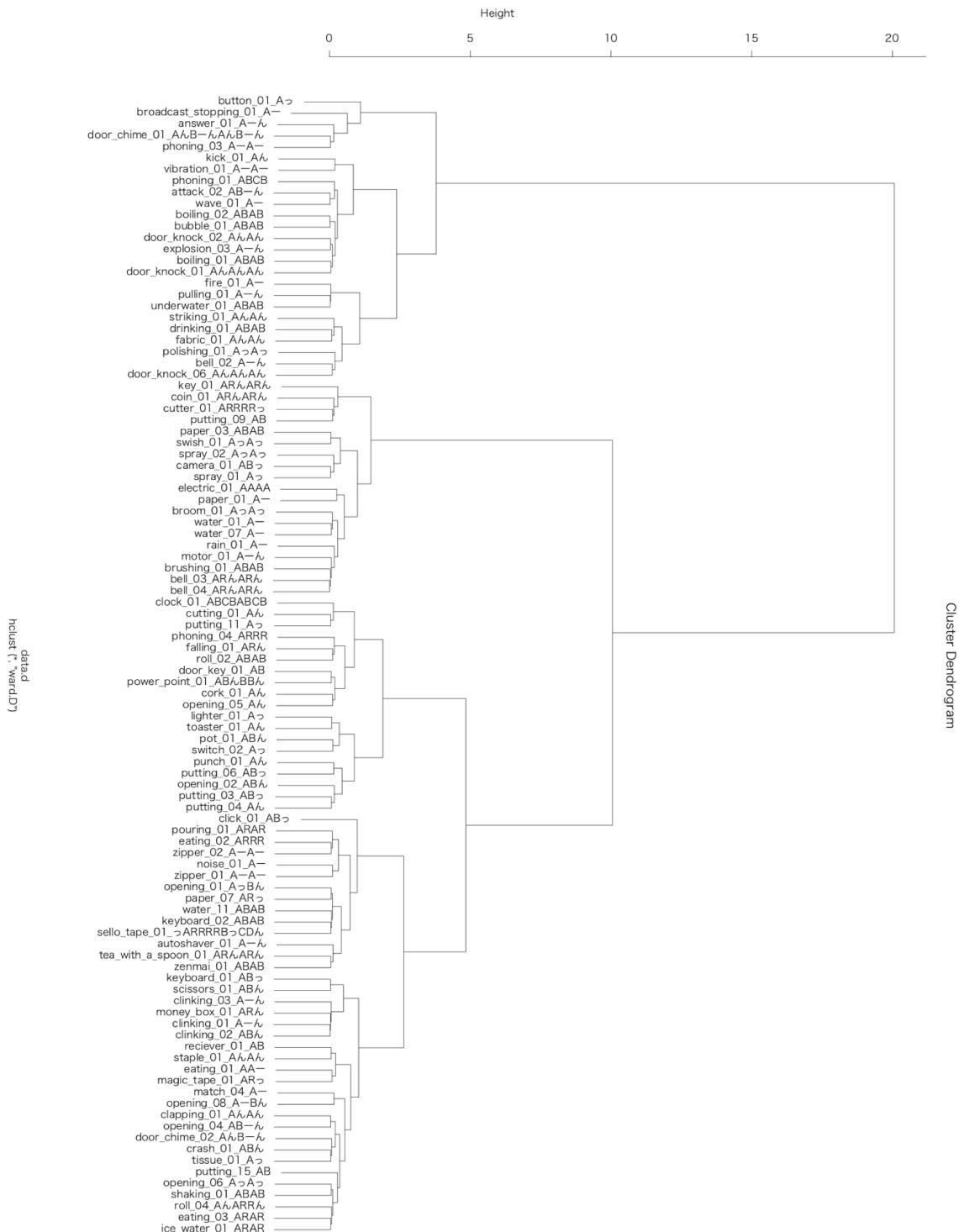


図 6.2: 共通の記号に置き換えられたオノマトペラベル

表 6.2: 各群における 50% を超えた項目の有無

群	長音符	促音	撥音	R	繰り返し
第 1 群	-	-	-	-	-
第 2 群	-	-	-	-	-
第 3 群	-	-	-	-	-
第 4 群	-	-	-	-	-
第 5 群	-	-	-	-	-
第 6 群	-	-	-	-	-
第 7 群	-	-	-	-	-
第 8 群	-	-	-	-	-
第 9 群	-	-	-	-	-

を下回っており、促音を持つオノマトペが dominant となる群は存在しないことが伺える。

第 2 群と第 8 群のペア、および第 6, 7, 9 群のトリプルについては、クラスタを構成しているオノマトペの形態的特徴が類似していることが伺える。例えば、第 6 群は「Aっ」型や「Aん」型のような短いオノマトペが 50% 以上を占めており、第 7 群も同様に短いオノマトペで形成されたクラスタであった。デンドログラム上でこれらの 2 群は近い距離にあり、形態的特徴によって分類されていることが伺える。

第 5 群は特徴的なクラスタであり、デンドログラムの高さでさらに 2 つに分離できる。一方は促音をもつオノマトペが 80% で、他方は長音符をもつオノマトペが 50% を占めていた。また、第 9 群は、群内に存在する ARAR 型同士が近い距離にあった。

6.2 音を聴取した際にもつ印象

音を聴取した際に持つ印象を、オノマトペの音象徴的意味の観点から分析する。オノマトペの音象徴的意味では、効果音を聴取した際に連想される印象を把握することができる。音象徴的意味から分析するために、オノマトペを音喻の要素ごとに分類する。前処理として、オノマトペをローマ字に置き換え、母音と子音ごとにデンドログラムのラベルを用意したものを、図 6.3 (母音)、及び図 6.4 (子音) に示す。

田守らの分類 [8] を参考に、音韻形態とオノマトペがイメージさせる音象徴的意味をまとめたものを、表 6.3 (母音) 及び表 6.4 (子音) に示す。本分析を行うにあたり、項目ごとに型名をつけた。この型に分類されるオノマトペを計数し、クラスタの特徴をまとめたものを表 6.5 に示す。

まず、クラスタに分類されていた母音を分析する。

表 6.5 より、どの群も複数の種類の母音で形成されていることがわかる。*v* 型をもつオノマトペは、4 節で収集されたすべてのオノマトペのうち 1 種類しか存在しなかったため、第 8 群にしか出現していなかった。第 2 群は、4 つの型で構成されており、*i* 型と *ii* 型、*iii* 型と *iv* 型がそれぞれ同じ割合で存在していた。第 5 群の *ii* 型のように、特定の母音が群内の dominant となっている場合を除き、群内の母音の種類に差は見られなかった。音の周波数によって擬音

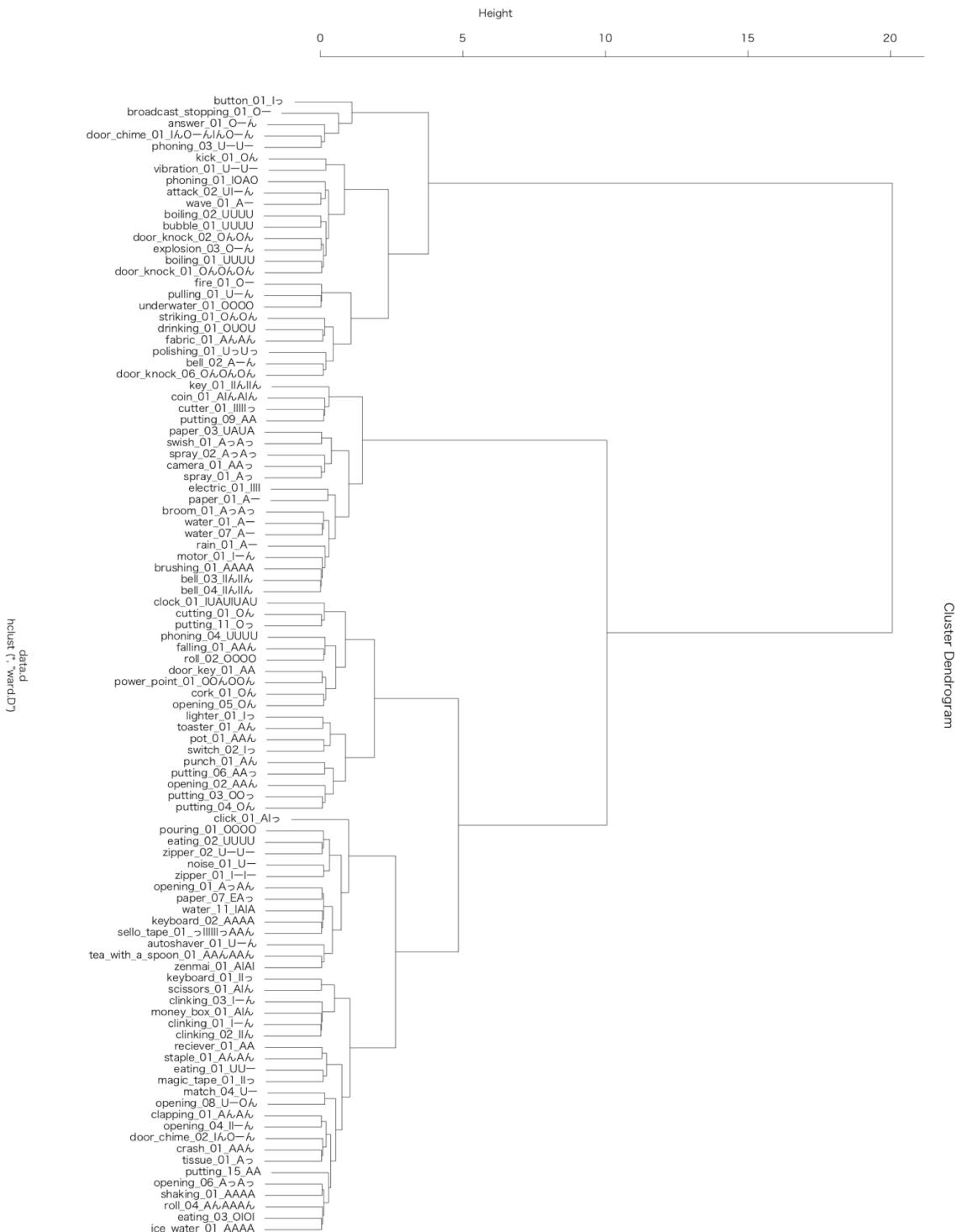


図 6.3: 母音に置き換えられたオノマトペラベル

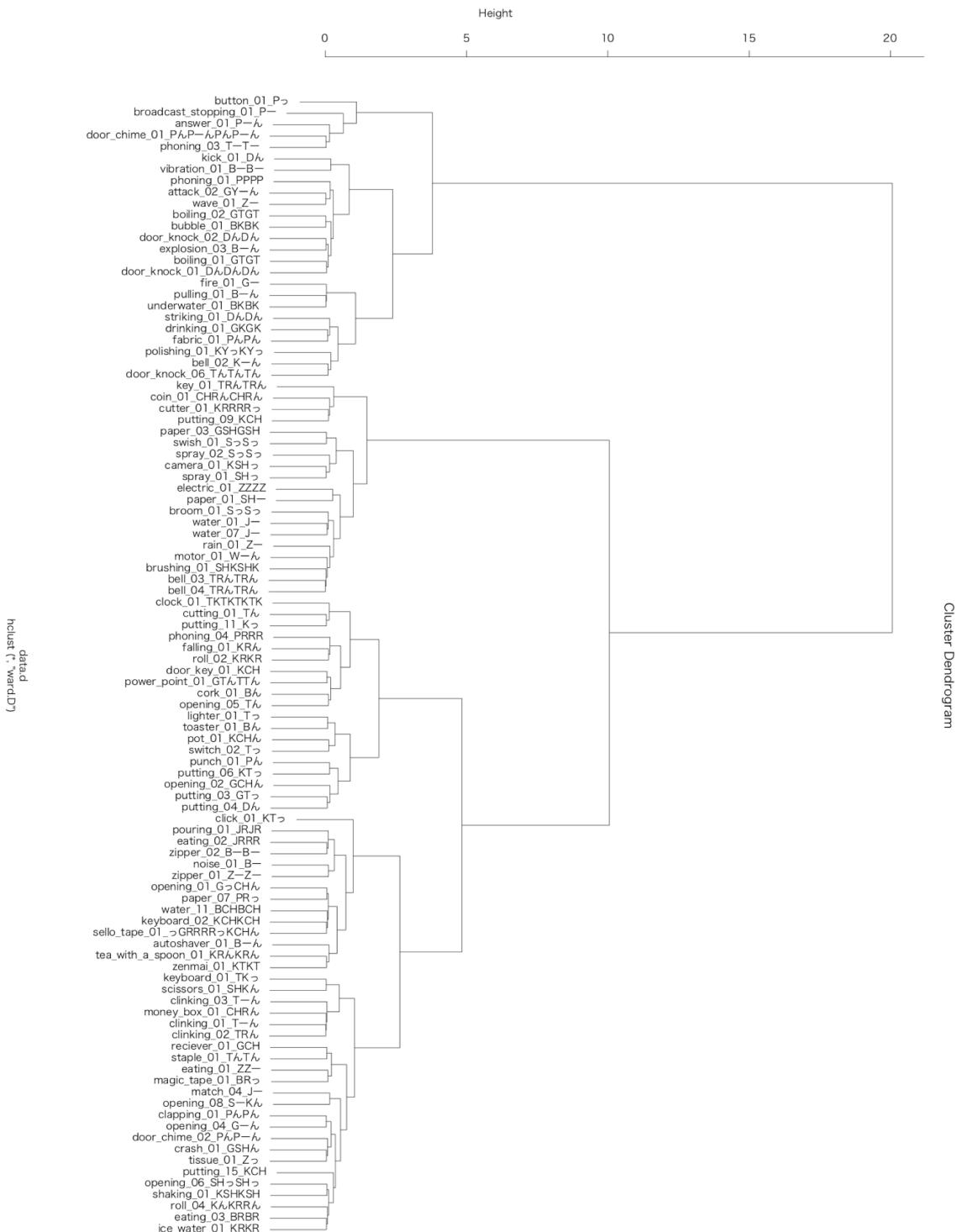


図 6.4: 子音に置き換えられたオノマトペラベル

語表現に用いる母音が変化するため，分類された効果音の高さにはらつきがあることが一因と考えられる。

つぎに，クラスタに分類されていた子音を分析する。

表 6.5 より，子音も母音と同様に，どの群も複数の型で構成されていることがわかる。V 型，VI 型，VII 型が見られないのは，今回収集したオノマトペ内に，その子音が含まれているオノマトペが存在していなかったためである。

第 5 群のように母音の特徴が顕著に表れた群も存在したが，第 2 群の I 型と III 型のように型に分類された母音の割合が同じものもあった。

分類されたオノマトペが一番多い型をその群を象徴する特徴とし，子音と母音を併せて群ごとの音象徴的意味を分析する。

母音と子音の音喻を鑑みると，第 1 群は，丸く，活発さが感じられる音，第 2 群は，柔らかく控えめな打撃音，第 3 群は，控えめな金属音，第 4 群は，甲高いゆったりした音，第 5 群は，広がりのある軽い接触音，第 6 群は，控えめな金属音，第 7 群は，派手な打撃音，第 8 群は，金属のような固い表面を打撃した音，第 9 群は，金属のような硬い表面と接触した派手な音，という分類になっていると言えよう。

ある効果音と，その効果音と同一のクラスタに属する効果音を被験者 3 名に聴取させたところ，クラスタ内の効果音の内，50% の効果音が類似している音であることが確認された。

6.3 まとめ

形態的特徴と音象徴的意味の両方の観点を合わせると，各々の群が異なる特徴を持つクラスタとして構成されている様子が伺える。また，類似したオノマトペのみによって構成されているクラスタがなかったことから，異なる形態的特徴で表現される効果音同士が同一のクラスタに含まれており，探索的な検索を行う際の“気づき”的の一助となることが期待できる。

以上のことから，本研究で提案するシステムの機能のうち，音響が類似した効果音の提示方法については実現可能であることが示唆された。今後，被験者実験を行い，それらがどのように機能し得るかを検証する。

表 6.3: 母音の音象徵

型	音喻の要素	音象徵的意味
i	/i/	線，一直線に延びたもの，甲高い音
ii	/a/	平べったさ，大きな表面に影響を与える出来事，関わっているものの全面的関与，広がり，華やかな，派手な，目立った出来事
iii	/o/	丸いもの，小さい出来事，小さい部分にしか影響を与えない出来事，目立たない・控えめな出来事
iv	/u/	口や鼻のような小さい丸い穴と関係のある出来事，柔らかくて控えめな音
v	/e/	動作の不適切さないし下品さ

表 6.4: 子音の音象徵

型	音喻の要素	音象徵的意味
I	/p, b/	物体に打ち当たる，破裂するといった急で爆発的な動作や出来事，ぴんと張った状態，突然性，力強さ。/p/：活発さや活動的な動作と関連
II	/t, d/	打撃
III	/k, g/	金属のような固い表面との接触
IV	/s, z/	1 モーラの語基：滑らかさ，障害の欠如，動作が急でないこと。2 モーラの語基：軽い接触，摩擦，小粒の動き，表面の張りがない，流動する液体，障害の不在，ゆったりした動き，静けさ。穏やかさ，壮快さ。人の性格：こぎれいさ，スマートさ，冷静さ
V	/h/	息，息の吐き出し。2 モーラの語基：不確定，頼りなさ，弱さ，繊細な優雅さ
VI	/m/	はっきりしない状態，落ち着きの無さ，理性の無さ。1 モーラの語基：抑制や不明瞭さ
VII	/w/	1 モーラの語基：動物や人の発する音

表 6.5: 群ごとの分類

群	母音	子音
第 1 群 (5)	i 型 (2), iii 型 (3), iv 型 (1)	I 型 (4), II 型 (1)
第 2 群 (11)	i 型 (2), ii 型 (2), iii 型 (5), iv 型 (5)	I 型 (4), II 型 (5), III 型 (4), IV 型 (1)
第 3 群 (9)	ii 型 (2), iii 型 (5), iv 型 (3)	I 型 (3), II 型 (2), III 型 (5)
第 4 群 (4)	i 型 (3), ii 型 (2)	II 型 (3), III 型 (2)
第 5 群 (15)	i 型 (4), ii 型 (11), iv 型 (1)	II 型 (2), III 型 (2), IV 型 (10)
第 6 群 (10)	i 型 (1), ii 型 (3), iii 型 (6), iv 型 (2)	I 型 (2), II 型 (5), III 型 (6)
第 7 群 (9)	i 型 (2), ii 型 (5), iii 型 (2)	I 型 (2), II 型 (6), III 型 (4)
第 8 群 (14)	i 型 (5), ii 型 (8), iii 型 (1), iv 型 (4), v 型 (1)	I 型 (4), II 型 (6), III 型 (6), IV 型 (1)
第 9 群 (23)	i 型 (9), ii 型 (12), iii 型 (3), iv 型 (3)	I 型 (4), II 型 (8), III 型 (9), IV 型 (7)

() 内は分類された個数を指す

7 議論

本章では、効果音と対応付くオノマトペの収集方法、抽出した音響特徴数、デンドログラムを分割する高さ、デンドログラムの分析の観点の 4 つについての議論を行う。

7.1 効果音と対応付くオノマトペの収集方法

本稿では、実験プログラムを用いて、効果音に対応付くオノマトペを収集した。この実験で実験参加者に与えられた指示は、聴取した効果音を他人が回答するであろうオノマトペで表現し、それを可能な限り数多く回答することである。しかし、聴取した効果音からその効果音は何の音であるかを推測し、それはどのようなオノマトペで表現されるかを回答された可能性がある。例えば、「カッカッ」という効果音を聴取した際、実験参加者はそれを「時計」の音であると推測し、時計の音を表現するために一般的に使用されることが多い「チクタク」というオノマトペを回答するようなことである。このような回答が多い場合、その効果音に本来付与されるべきであったオノマトペが対応付いていない可能性がある。これは、他人が回答するであろうオノマトペを回答させたことと、可能な限り数多くのオノマトペを回答させたことの 2 点が起因していると考えられる。

前者は個人によるオノマトペの表記揺れを統制するために、実験参加者に指示された内容である。他人が回答するであろうオノマトペを回答させたことにより、その効果音は何の音であるかを推測し、効果音からは聴取されないオノマトペの回答を誘発したと考えられる。

後者は数多くのオノマトペを収集することで共通したオノマトペを発生させやすくし、効果音に対応付くオノマトペの選出を容易にするために、実験参加者に指示された内容である。可能な限り数多くのオノマトペを回答させたことにより、聴取した効果音から連想されるあらゆるオノマトペを回答された可能性がある。

これらの回答は、聴取した効果音の波形特徴を反映したオノマトペでないため、デンドログラムの分析には適さない。これらの回答を減少させるための解決策として、まず、他人が回答するであろうオノマトペの回答をするように指示をしないことが挙げられる。この指示により、どこまで表記揺れを統制できているかが不明であり、効果音から聴取されないオノマトペの回答が懸念される。さらに、オノマトペの回答難易度を上げていることも考えられる。しかし、本稿で行った実験は、効果音を検索する際の、クエリと一致しているかどうかの判断材料として使用するタグを収集するためには、多くのオノマトペが収集される点で有効であると考える。

7.2 抽出した音響特徴量数

音響特徴量の類似度ごとに効果音をクラスタリングした。クラスタリングを行うために効果音から抽出した音響特徴は、音色的印象的側面と文献 [6] を参考にして抽出した、5 節の表 5.1 の特徴量 8 項目である。この特徴量を用いた分類では、同一のクラスタ内の効果音の内、50% の効果音が類似している音であることが確認された。しかし、提案システムを実装する際、本稿で使用した効果音数よりも多くの効果音を扱うため、類似している音であると感じにくくなることが考えられる。そのため、提案システムの実現へ向けて、精度を高める必要があると考える。

清水らが行った効果音の音響特徴を用いた分類では、効果音から 33 項目の特徴量が抽出されている [10]。この研究では「水」の音が扱われており、「波」の音同士や「激しい流水」の音同士で分類されることが確認されている。音の発生源が同一であっても詳細なニュアンスが適切に反映された分類になっていることから、音響特徴を用いたクラスタリングで、効果音の詳細なニュアンスでの分類を行うことが可能であることがわかる。このことから、人が類似していると感じる効果音同士での分類をするためには、効果音から抽出する音響特徴数を増やす必要があると考える。

7.3 デンドログラムを分割する高さ

デンドログラムの分割位置は、分類内容を確認した分析者の主観評価によって定められる。本稿では、デンドログラムを分析する際、単一の効果音が存在しないようにデンドログラムの高さを調整して分割した。しかし、この方法では、最多で 23 個の効果音が 1 つの群に分類されていた。これは、本研究で提案するシステムを実現する際、本稿で行った分割方法では、提示する効果音数が多すぎる可能性が考えられる。23 個の効果音がすべて類似した音同士の分類であれば、本稿で行った方法は問題なく機能することが考えられる。しかし、そうでない場合、関係のない効果音が混入する割合が高くなる。

デンドログラムを、本稿で設定した分割数である 9 群に分割した場合と、10, 15, 20 群に分割した場合の群ごとに分類される効果音の個数の例を、表 7.1 に示す。この表から、デンドログラムを 15 群以上に分類すると、1 つの群内に分類される効果音数が 20 個を超える群がないことが見てとれる。20 群以上に分類すると、1 つの群内に分類される効果音数が 1 個の群が見られたり、10 個以下の群が増えたりすることも確認された。

本研究で提案するシステムは、文脈、音響、オノマトペの表象、の 3 つの観点それぞれの類似性から効果音を提示する。そのため、デンドログラムを単一の効果音が存在する方法で分割しても、他の 2 つの観点で類似した効果音が提示されるため、提示される効果音数に影響がないことが考えられる。例えば、第 5 群がさらに 2 つの群に分割可能であることが示唆されたように、デンドログラムを分割する高さを下げる、より詳細に分類する方が良い結果が得られる可能性があると推測される。

7.4 デンドログラムの分析の観点

本稿の検証は効果音に対応づいたオノマトペの比較のみであり、効果音の発生源や音質が類似しているか否かという観点を考慮していない。オノマトペの表象特徴が類似していても、人がその音を聴取した際に、その効果音同士が類似していると感じない可能性がある。そのため、クラスタリングされた効果音集合を人が聴取した場合、類似していると感じるのかについて聴取実験を行い、検証する必要があると考える。

表 7.1: デンドログラムの分割数による群ごとの効果音分類個数の違い

	9 群 *	10 群	15 群	20 群
第 1 群	5	1	1	1
第 2 群	11	4	4	1
第 3 群	9	11	11	3
第 4 群	4	9	3	2
第 5 群	15	4	6	9
第 6 群	10	15	4	3
第 7 群	9	10	5	6
第 8 群	14	9	10	4
第 9 群	23	14	10	5
第 10 群	—	23	4	10
第 11 群	—	—	5	3
第 12 群	—	—	1	7
第 13 群	—	—	13	4
第 14 群	—	—	6	5
第 15 群	—	—	17	1
第 16 群	—	—	—	5
第 17 群	—	—	—	8
第 18 群	—	—	—	6
第 19 群	—	—	—	4
第 20 群	—	—	—	13

* 本稿で採用した分割数

8 結論

本稿では、文脈、表象、音響という3つの要素で可視化された類似性により、効果音の探索的検索が可能なシステム実現の端緒として、音響特徴量によってクラスタリングされた効果音集合に見られるオノマトペの表象情報を考察した。

1章では、本研究に至った背景と問題点を述べ、解決すべき課題を明らかにした。効果音CDや効果音検索サイトを用いて効果音を検索する際の問題点を3点に集約し、それらの問題点は付与する効果音の決定までに多大な聴取時間を要することに繋がると述べた。これらを解決するために、効果音の文脈、音響、オノマトペの表象の類似性を可視化し、各々の類似性を辿れるようにすることで、効果音の探索的検索が可能なシステムを実現するとした。そして、システム実現の端緒として、音響特徴量を用いた効果音の階層的クラスタリングと、クラスタリングされた効果音集合に見られるオノマトペの表象情報との関係性について考察することを述べた。

2章では、本研究で提案するシステムの実現に関する研究を述べた。まず、効果音を検索する際、検索する音を言語化する必要があるため、人がイメージした効果音を表現する方法について述べた。次に、効果音検索システムに関する研究を述べた。最後に、不明確な要求を明確にしていく情報検索モデルである Exploratory Search について説明し、その有効性について述べた。

3章では、対象ユーザを明らかにし、本研究で提案するシステムのデザイン指針を立てた。効果音検索のモデルケースを考察することで、効果音検索に必要となるデザイン指針を5点に集約した。集約したデザイン指針に基づいて、効果音検索システムの概要を述べた。提案システムを実現するためには、文脈、音響、オノマトペの表象の3種類の観点の類似性を提示することが必要である。本稿では、音響の実現に向けて、音響特徴量に基づいて分類された効果音の特徴を、効果音に対応づいたオノマトペによって考察することとした。

4章では、効果音に対応付くオノマトペを収集するために、作成したプログラムを用いてインターネット上で実験を行った。収集されたオノマトペを統制し、最も共通して付与されたオノマトペを効果音に対応付けた。

5章では、効果音から音響特徴量を8項目抽出し、抽出された音響特徴量を用いて Ward 法で階層クラスタリングを行った。階層クラスタリングでは、効果音 100 個を分類した。

6章では、第5章で得られたデンドログラムを高さによって9群に分類し、分析を行った。その結果、類似したオノマトペごとの明確な分類ではなかったが、群ごとの分類がおおよそ意味することが判断可能な分類であることがわかった。加えて、その群に属する効果音集合を聴取すると、類似していると感じる効果音同士の分類であることが示唆された。

7章では、効果音と対応付くオノマトペの収集方法、抽出した音響特徴数、デンドログラムを分割する高さ、デンドログラムの分析の観点の4つについての議論を行った。効果音と対応付くオノマトペの収集方法では、効果音からは聴取されないオノマトペを回答された可能性と、その解決策を述べた。抽出した音響特徴数では、人が類似していると感じる効果音同士での、より詳細な分類をするためには、効果音から抽出する音響特徴数を増やす必要性を述べた。デンドログラムを分割する高さでは、分割する高さを下げる、より詳細に効果音を分類することを検討した。デンドログラムの分析の観点では、分類された効果音群を人が聴取した際の印象を考慮していないことから、聴取実験を行う必要性を述べた。

今後は、クラスタリングされた効果音集合を人が聴取した場合、類似していると感じるのかについて聴取実験を行い、検証すると共に、文脈とオノマトペの表象の2種類の観点それぞれの類似性提示技術の実現を目指す。

謝辞

本研究を纏めるにあたり，関西大学総合情報学部総合情報学科の松下光範教授に多大な御指導，御鞭撻を賜りました．また，インターンなど，貴重な機会をたくさん頂きました．大変お世話になったことを深く感謝致します．

本研究を進めるにあたり，多大なる御尽力を頂きました，立命館大学情報理工学部メディア情報学科の山西良典助教に心より感謝致します．

実験手続きについて，御指導を賜りました，日本電信電話株式会社 NTT コミュニケーション科学基礎研究所の松田昌史氏に深く御礼申し上げます．

研究を通じて，貴重な御意見を賜りました，関西大学大学院総合情報学研究科知識情報学専攻の上間大生氏，大塚直也氏，山下諒氏に感謝の意を表します．

実験で使用するプログラムの実装にあたり，関西大学総合情報学部総合情報学科の盛山将広氏に御協力を頂きました．心より感謝致します．

研究の相談に乗って下さった，松下研究室に所属している先輩，同期，後輩に，深く御礼申し上げます．

効果音とオノマトペを対応付ける実験に御参加下さった皆様，実験の宣伝活動を行って下さった皆様の御協力が頂けなければ，本研究は成り立ちませんでした．感謝の意を表します．

関大立命合同研究会では，有意義な御指摘を賜りました．参加者の皆様に御礼申し上げます．

本研究では，効果音を配布しているサイト「フリー効果音」の効果音を使用しました．効果音を提供して頂いたことを，感謝申し上げます．

苦楽を共にし，いつも支えになってくれた，関西大学総合情報学部総合情報学科の軽音楽サークル Banditz に所属しているバンドメンバーである，北田佳子氏，松岡龍太郎氏に，深く感謝致します．

研究に行き詰った際の気分転換として，大きな音の鳴るエレクトリックアコースティックギターを用いての弾き語りを，深夜にも関わらず行っていたことに対して，苦情ひとつ入れること無く受け入れて下さった近隣にお住まいの方々にお詫びをすると共に，心より御礼申し上げます．

最後に，お世話になりました友人，家族に深く感謝致します．

参考文献

- [1] Lartillot, O., Toiviainen, P. and Eerola, T.: A Matlab Toolbox for Music Information Retrieval, *Data Analysis, Machine Learning and Applications*, pp. 261–268 (2008).
- [2] Taylor, R. S.: Question–Negotiation and Information Seeking in Libraries, *College & Research Libraries*, Vol. 29, No. 3, pp. 178–194 (1968).
- [3] Wake, S. and Asahi, T.: Sound Retrieval with Intuitive Verbal Expressions, *Proc. 5th International Conference on Auditory Display* (1998).
- [4] Ward, J. H.: Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, Vol. 58, No. 301, pp. 236–244 (1963).
- [5] White, R. W. and Roth, R. A.: *Exploratory Search : Beyond the Query–Response Paradigm*, Synthesis Lectures on Information Concepts, Retrieval, and Storage, Morgan and Claypool Publishers (2009).
- [6] 草間かおり, 伊藤貴之: MusCat : 楽曲の印象表現に基づいた一覧表示の一手法, 情報処理学会研究報告, Vol. 2009-MUS-81, No. 19, pp. 1–6 (2009).
- [7] 田守育啓: オノマトペ擬音・擬態語をたのしむ (もっと知りたい!日本語), 岩波書店 (2002).
- [8] 田守育啓, ローレンススコウラップ: オノマトペ—形態と意味 (日英語対照研究シリーズ), くろしお出版 (1999).
- [9] 帆足啓一郎: 音楽情報の検索, 電子情報通信学会技術研究報告. SP, Vol. 110, No. 143, pp. 19–24 (2010).
- [10] 清水敬太, 北原鉄朗, 駒谷和範, 尾形哲也, 奥乃博: OnomaTree : 擬音語と木構造を併用した環境音検索インターフェース, 情報処理学会第 69 回全国大会, Vol. 69, No. 2, pp. 193–194 (2007).
- [11] 盆子原健太, 大塚直也, 松下光範: EC サイトにおける商品探索プロセスに着目したプレゼント探索支援システムの提案, 人工知能学会インタラクティブ情報アクセスと可視化マイニング研究会, Vol. 6, No. 6, pp. 34–39 (2014).
- [12] 和氣早苗, 旭敏之, 井関治: 効果音検索システム～「音」の表現方法に関する実験と考察～, 情報処理学会第 48 回全国大会, pp. 261–262 (1994).
- [13] 青木直史, 伊藤博之, 佐藤隆文, アレキサンダー・ブルガー: キーワードによる効果音データベース検索システムの開発, 2003 年電子情報通信学会総合大会, p. 42 (2003).
- [14] 岩宮眞一郎: 音色の感性学—音色・音質の評価と創造, コロナ社 (2010).