Vol.2023-HCI-202 No.16
2023/3/13

動的な音を用いたカジュアルな遠隔コミュニケーションの拡張

綿谷　晃太郎[1,a)]　塚田　浩二[1]

概要：Skype や Zoom 等に代表されるビデオ会議システムは会議等のフォーマルな場面だけでなくオンライン飲み会などのカジュアルな場面で利用されることが増えてきた．こうした場面では，細かな感情や雰囲気が共有されにくいことや，自宅から参加できることから辞め時が難しいといった様々な課題が生じてくる．本研究では，オンライン飲み会のような複数人でのカジュアルな遠隔コミュニケーション場面において，ユーザが任意のタイミングで効果音を挿入でき，かつ遠隔コミュニケーション中の発話数によって BGM のテンポを自動で調整するシステムを作成した．さらに，評価実験を通してシステムによるコミュニケーションへの影響やユーザの主観的な印象などを調査した．

Extending Casual Remote Communication with Interactive Sound

Wataya Koutaro[1,a)]　Koji Tsukada[1]

1. 序論

1.1 背景と目的

昨今のコロナウィルスの流行や技術の発展などから，Skype や Zoom 等に代表されるビデオ会議システムは，会議等のフォーマルな場面だけでなくオンライン飲み会などのカジュアルな場面で利用されることが増えてきた．こうした場面では，議題等がなく会話の停滞が発生しがちといったカジュアルなコミュニケーションの特性に加えて，細かな感情や雰囲気が共有されにくいといった遠隔コミュニケーションの問題や自宅から参加できることから辞め時が難しいといった社会的な問題が絡み合ってくる．

また，本研究では音とコミュニケーションの関係に着目した．これまでの研究では，音楽を聴くことによって人間の感情や行動が変容することなどが知られており，それを利用することで音楽がコミュニケーションのしやすさに効果があると考察されている [2]．音楽が利用される身近な例としては，カフェなどの飲食店では雰囲気づくりのために BGM をかけているところも多くみられる．

本研究では，オンライン飲み会のような複数人でのカジュアルな遠隔コミュニケーション場面において，インタラクティブに BGM や効果音を提示することで，会話の停滞や細かな感情の共有等の問題を軽減することを目的とする．

2. 関連研究

2.1 音が人の行動や感情に与える影響に関する研究

高口 [1] は，テンポの違う二つの BGM を使った実験により，人の行動にどのような違いを生むかを検討した．実験結果から，テンポの遅い BGM はテンポの速い BGM の時に比べゆっくりとした行動を促すことを確認した．

大野ら [2] は音楽のもつリラックス効果を利用して，会話者それぞれが独立して音楽を聴取しながらコミュニケーションをはかるという緊張緩和手法の提案を行い，実験によってコミュニケーションに及ぼす影響を調査した．その結果，音楽を聴取することによってコミュニケーション時の緊張が緩和され，さらに会話がしやすくなる傾向にあることを明らかにした．

高橋ら [3] は，音楽が情動にどのような影響を及ぼすかを明らかにするために，主観的な情動の指標である POMS

1 公立はこだて未来大学
Future University Hakodate
a) g2121063@fun.ac.jp

ⓒ2023 Information Processing Society of Japan

（Profile of Mood State）を用いた実験を行った．その結果，実験者側で選択した鎮静的音楽を聴取した場合と，被験者が選んだ好みの音楽を聴取した場合のどちらでも情動が変化することを明らかにしている．

2.2 遠隔コミュニケーションを支援する研究

SakeEffects[4] は，アルコールセンサを内蔵したマイクを試作した．オンライン飲み会中にアルコールセンサを用いて酔い度を評価しユーザの顔にビデオエフェクトを合成する．試用の結果，ビデオエフェクトの変化によって話題を広げることや，盛り上がりを増すことができていた．

小山ら [5] は，オンラインビデオ会議において参加者の同期的なジェスチャーをを用いて非言語コミュニケーションを支援するシステムを提案した．具体的には，ハイタッチとキャッチボールのジェスチャーを認識できる．評価実験の結果，ハイタッチでは一体感の増加，キャッチボールでは発話衝突の低減や全員の発言を促す効果が期待できるとした．

いざガヤ [6] は，背景雑音として他の部屋の会話を聞きながらビデオ通話ができるウェブアプリである． 複数の部屋があるウェブ会議システムで，部屋 A の BGM として部屋 B の会話を流すことで，コミュニケーションの促進を狙った．

徳原ら [7] は，ビデオ会議中に参加者が無意識に行なっている頷きなどをカメラで認識し，画面にその回数をリアルタイムで提示するというシステムを提案した．

二瓶ら [8] は，映像を伴う遠隔コミュニケーション場面において，会話参加者の顔映像と発話音声を肯定的に変容させるツールを実装した．具体的には，口角を上げたり，音声のピッチを高くする処理を行っている．さらに，実験を通してツールを適切に使えばコミュニケーションを活性化することを示した．

大槻ら [9] は，人の眼球を模倣した立体ディスプレイである ThirdEye を作成した．ThirdEye は，ビデオ会議システムと併用可能であり，遠隔の参加者の視線を提示する．実験によって，遠隔対話者の顔のみをディスプレイに表示する一般的な方法に比べ，正確に視線方向を提示できることを示した．

Shimizu ら [10] は，指向性スピーカーを用いて，グループ会話における現在の発話者と受動的な参加者にのみ音刺激を与え，両者の相互注視を誘導するシステムを作成した．評価の結果，受動的な参加者の発話量を増加させることができたと報告している．

2.3 本研究の特徴

本研究では，カジュアルな遠隔コミュニケーションを支援するために， BGM や効果音を提示するシステムを提案する．効果音を任意のタイミングで提示させることで感情の共有を狙い，また発話数をトリガーとして BGM のテンポを変更することで発話の停滞を抑止する効果を狙う．また，Web ベースシステムとして開発することで，一般的なパソコンから手軽に利用できる点も特徴である．

3. 提案システム

本章では，本研究で提案するシステムのコンセプトと構成について説明する．

3.1 コンセプト

本研究では，議題等がなく会話が停滞しがちである問題や，細やかな感情や雰囲気が共有されにくい問題，やめ時が難しいといった問題をコミュニケーション中に効果音や BGM を提示することで軽減したいと考えた．

細やかな感情や雰囲気が共有されにくい問題を軽減するために，カジュアルな遠隔コミュニケーション中にユーザが任意に効果音を提示できる機能を実装する．また，会話の促進ややめ時のきっかけを作るために，発話数などから BGM を自動提示する機能も合わせて実装する．

3.2 システム構成

提案するシステム構成図を図 1 に示し，想定しているシステムのフローを説明する．提案システムは，Zoom などの一般的なビデオ会議システムに対し，Bot として参加させて使用する．まず，オンライン会議システムの各参加者が，手動操作部・音声認識部からなるユーザ操作アプリに音声・操作を入力する．音声入力を発話文字列としてデータベース（Google スプレッドシート）に記録し，操作情報を P2P 通信で直接音提示アプリに送信する．音提示アプリは Bot 端末で動作しており，参加者に効果音と BGM を提示する．効果音は P2P 経由の操作情報を元に提示され，BGM はデータベース上の発話量を元に提示される．

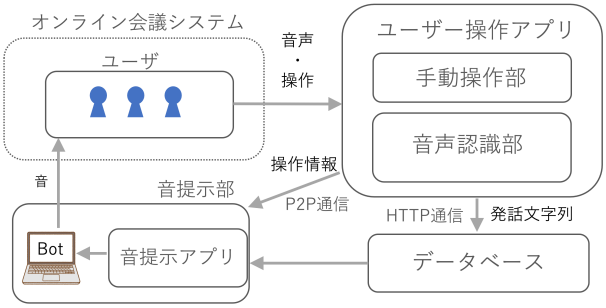

図 1 システム構成図

各アプリケーションは，ダウンロードなどが必要なく手軽に扱えるよう，ブラウザ上で動作する Web アプリケーションとして実装する．

4. 予備実験

4.1 目的

システムでの利用可能性を検討するために，Zoom を用いたカジュアルな遠隔コミュニケーション中に BGM のテンポを変化させた際の発話数の変化などを調査した．

4.2 手法

3 名の被験者を 1 グループとし，3 グループ（g1，g2，g3）に対し Zoom を使った雑談を行うよう指示した．雑談は，3 分の準備セッションと 9 分× 3 回の本番セッションから構成されている．各セッションの前にトークテーマを提示し，セッション中には BGM を再生した．その後，トークテーマの話しやすさなどについてアンケート調査を行なった．また，音声認識により発話数を記録した．

BGM として使用した曲は The Entertainer（作：Scott Joplin，1902）であり，曲のテンポとして遅い（bpm=88），普通（bpm=108），速い（bpm=128）の 3 種類用意した．セッションごとにテンポの組み合わせを変更し，

A) 普通→普通→普通
B) 普通→速い→普通
C) 普通→遅い→普通

の 3 セッション用意した．また，順序効果にも配慮した．この時，曲のテンポが変化することは被験者に伝えなかった．

4.3 結果

4.3.1 発話数取得率

まず，音声認識による発話数取得率の結果を述べる．発話数取得率は音声認識された発話数 ÷ 実際の発話数として算出した．実際の発話数は録音した実験時の会話を実験者が手動で文字起こしして計測した．算出した結果を表 1 として示す．

表 1 音声認識による発話数取得率

	発話数取得率 (音声認識された発話数/実際の発話数)
g1	79.6%（7596/9538）
g2	84.2%（6397/7596）
g3	86.8%（7389/8513）
全体	83.4%（21382/25647）

全体での平均が 83.4%となり，ある程度発話が取得できていることがわかったが，標準偏差は 3.6%であり，グループごとにややばらつきが見られた．

4.3.2 テンポによる発話数の変化

各セッションの発話数は，序盤／中盤／終盤に分けて計測した．参考として，セッション B での発話数を表 2 に示す．

テンポの変化によって発話数がどのように変化しているかを見るために，発話数の変化率を中盤の発話数 ÷ 序盤の発話数として発話数の変化率を算出した．算出した発話数の変化率をセッションごとに並べたグラフを図 2 として示す．

表 2 セッション B での発話数

	序盤（普通）	中盤（速い）	終盤（普通）
g1	1068	1077	1133
g2	937	912	957
g3	750	1042	864

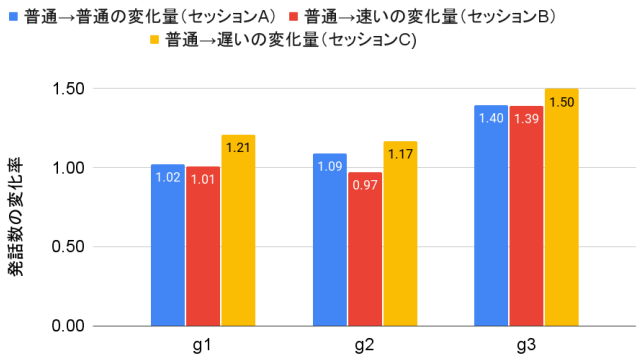

図 2 各セッションの前半と中盤での発話数の変化率の比較

図 2 より，セッション C の時どのグループも発話数の変化率が増加していることがわかった．また，アンケートの結果からは，セッションごとのトークテーマとの相関は見られなかった．詳細な原因の推定はできなかったが，BGM のテンポを遅くすることで発話がやや増加していたことがわかった．

5. 実装

5.1 システムの設計

本システムの設計として，システムに搭載される 2 つの主要機能について述べる．

1 つ目の機能として効果音ボタンを実装した．参加者が任意のタイミングでボタンを押すと，そのボタンに対応した効果音が挿入される．この機能で狙う効果として，参加者が会話中の感情等を手軽に他の参加者に伝える効果を狙った．本システムでは，「嬉しい」「悲しい」「怒り」の感情を表す効果音と，「拍手」といった動作を表す効果音を，それぞれ弱・中・強の 3 段階ずつ用意した．

2 つ目の機能として，参加者の発話数の変化やコミュニケーションの開始時からの時間経過といった情報をトリガーとし，BGM のテンポを自動で変化させる機能を実装した．この機能では，予備実験で得た知見などを活かし，BGM のテンポをシステムが自動で調節して発話の誘導を狙う．また，あらかじめユーザが設定した時間が経過した時や，発話が極端に減ってきた時に蛍の光を挿入することで，自然と終わるきっかけを演出する．BGM は，カフェな

ⓒ2023 Information Processing Society of Japan

どで利用されるジャズ調の曲を，テンポを調整して使用した．ジャズ調の BGM を選んだ理由として，予備実験などから会話の邪魔にならない程度の曲調であるものを使用するべきと判断したためである．テンポ（BPM）は，曲の速さを言葉で示した “速度標語” と呼ばれる基準を参考に設定した [11]．本稿では以降，BPM=108 の BGM を “普通テンポ”，BPM=88 の BGM を “遅いテンポ”，BPM=128 の BGM を “速いテンポ” とする．BGM の編集には，Audacity[*1]というオーディオ編集ソフトを使用した．その過程で，編集前の BGM の BPM を調べる必要があったため，MixMeister BPM Analyzer[*2]というソフトを使用し BPM を測定した．

これらの 2 つの機能を “ユーザ操作アプリ” と “音提示アプリ” の 2 つの Web アプリケーションとして実装した．ユーザ操作アプリは参加者全員が利用する．音提示アプリは Bot または参加者 1 人のデバイスで起動する．これらの 2 つのアプリ間での情報のやりとりは，Google スプレッドシートを経由した方法と P2P 通信による方法の 2 種類で行う．Google スプレッドシートは誰がどれだけ発話したかなどの情報を保存するために利用し，P2P 通信はユーザ操作アプリでボタンが押されたことを音提示アプリに素早く伝えるために利用する．動作環境として Google Chrome を想定した．システムによって鳴る効果音と BGM は全て，音量やテンポなどの編集が可能なフリー素材を用いた．[*3*4*5*6]

音の共有方法については，ソフトウェアミキサーを使った方法と，Zoom の音声共有機能を使った方法の 2 パターンを検討した．ソフトウェアミキサーを使った方法では，音をマイク入力として扱うことで常時利用が可能であるが，BGM として適切な音量（会話の邪魔にならない程度の音量）では Zoom のノイズ抑制機能のためか音が途切れて流れてしまうといったことが予備実験からわかった．よって，音量をある程度下げても安定して BGM が流れる，Zoom の音声共有機能を使った方法を採用した．しかし，音声共有機能は，画面共有機能と競合するデメリットがあるため，今後も継続して検討する必要がある．

5.2 ユーザ操作アプリ

ユーザ操作アプリの起動時の画面を図 3 に示す．ユーザ操作アプリは P5.js という Web 上で動作する Processing 環境で実装し，Web アプリケーションとして動作させる．ユーザが Web ブラウザ上で手軽に利用できることに配慮して実装した．

*1 https://www.audacityteam.org/
*2 https://www.mixmeister.com/bpm-analyzer.html
*3 STAR DUST BGM：https://stardustbgm.com/
*4 MMT38：https://mmt38.info/
*5 Fesliyan Studios：https://www.fesliyanstudios.com
*6 効果音ラボ：https://soundeffect-lab.info/

図 3 ユーザ操作アプリの外観

ユーザ操作アプリは，ユーザの発話数などの情報をデータベースとして使用する Google スプレッドシートに送信する機能や，ユーザが押した効果音ボタンの情報を音提示アプリに送信する機能を持つ．

Google スプレッドシートに情報を送信する方法として，Google Apps Script で作成したプログラムによってユーザ操作アプリと Google スプレッドシート間で HTTP 通信を行い，シート上に記録する．

ユーザが押した音ボタンの情報を音提示アプリに送信する方法として，ユーザ操作アプリと音提示アプリ間で Peer-to-Peer(P2P) 通信する方法をとった．P2P 通信によって，アプリ間の通信にかかる時間が僅かとなり，ユーザがボタンを押してから音提示アプリが音を提示するまでに感じるタイムラグを小さくすることが可能である．P2P 通信を実現するにあたって，PEER JS[*7]というライブラリを使用した．PEER JS は WebRTC（Web Real-Time Communication）を利用した P2P 通信を簡単に実現できるライブラリである．

UI の 1 つとして，グループ ID・ニックネームを入力可能なテキストボックスを実装した．音提示アプリ側で事前に取得した P2P 通信用 ID をチャットなどで共有してもらい，グループ ID の入力欄に入力し connect ボタンを押すことで通信を開始する．アプリ間で通信が確立できた時，画面に表示されている接続状況のステータスが “未接続” から “接続確立” に変化する．また，入力したグループ ID とニックネームを Google スプレッドシートのシート名に使用し，音提示アプリがシートから情報を取得する際に，同一グループ内のシートからのみ情報を取得するよう実装した．

また，発話数を取得するために，ユーザの発話を認識する．画面上部右にあるスイッチによってマイクのオン・オフの切り替えができる．発話数取得には p5.speech[*8]というライブラリを利用した．P5.speech は，P5.js にて音声認識 Web サービスの SpeechRecognition に容易にアクセスできる拡張機能である．また，日本語文字列をひらがなやカタカナに変換するひらがな化 API[*9]を利用した．音声認識の結果として得た漢字交じりの文字列をひらがな化した

*7 https://peerjs.com/
*8 https://idmnyu.github.io/p5.js-speech/
*9 https://labs.goo.ne.jp/api/jp/hiragana-translation/

ものの文字数を発話数として取得する．

想定するユーザ操作アプリの使用法を図 4 に示す．ユーザ操作アプリは図 4 のように，Zoom の画面上にブラウザのウィンドウサイズを小さくして重ねることで，参加者の顔を見ながら使用できるといった使用法を想定し実装した．そのため，アプリ画面が大きくなりすぎないようボタンの数を限定し，ボタンを押して離す時間の長さによって１つの効果音につき 3 段階の強弱を切り替えられるよう実装した．

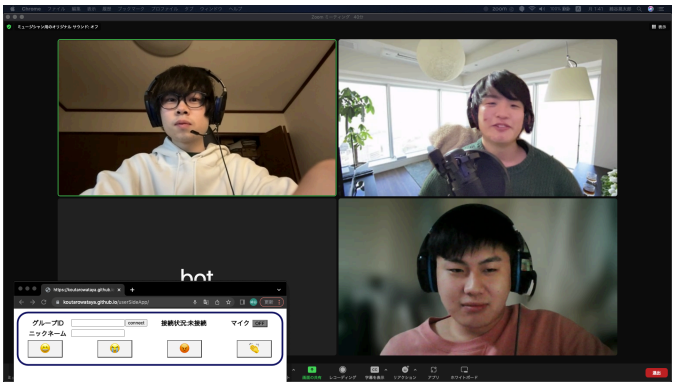

図 4　想定するユーザ操作アプリの使用法

各ボタンを押している時間が 0〜1 秒では “弱”，1〜2 秒では “中”，2〜3 秒では “強” の効果音が鳴り，3 秒以上押すとキャンセル状態となりボタンを話しても音が鳴らないよう実装した．その際のボタンの状態の変化を図 5 に示す．また，各音ボタンと効果音の対応表を表 3 に示す．

表 3　音ボタンと効果音の対応表

	弱	中	強
喜び	男性 1 人が「イェイ」と発する声	子供達が「イェーイ」と発する声	男性と子供達が「イェーイ」と発する声
悲しみ	気の抜けた「ショボーン」のような音	「チーン」というお鈴のような音	「ガーン」という音
不満	男性 1 人がブーイングする声	複数人がブーイングする声	より多くの人数でブーイングする声
拍手	1 人で拍手する音	複数人で拍手する音	複数の拍手に歓声などが合わさった音

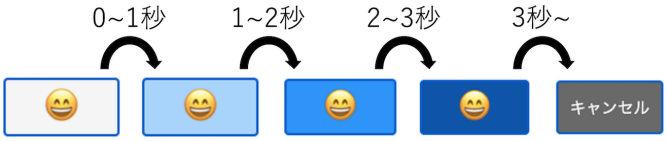

図 5　押す時間の長さによるボタンの状態変化

5.3　音提示アプリ

音提示アプリの起動時の画面を図 6 に示す．音提示アプリも上記と同様の理由で P5.js にて実装した．

図 6　音提示アプリの外観

主な使用ライブラリとして，P5.js で音声ファイルを利用可能とする P5.sound と，上述した PEER JS を利用した．まず，Google スプレッドシートにアクセスしユーザの発話数などを元にして BGM を変化させる機能を実装した．さらに，ユーザ操作アプリから P2P 通信を介して送られた操作情報を元に効果音を鳴らす機能を実装した．

P2P 通信には固有の ID が必要であるため，音提示アプリ起動時に PEER JS の機能を用いて，画面上にグループ ID を表示する．“ID をコピー” ボタンを押すことでクリップボードに ID がコピーされ，Zoom のチャットなどにペーストすることで，参加者に ID を伝えることができる．

ユーザの発話数によって BGM を変化させる機能には，基準となる 1 分あたりの発話数（基準発話数）が必要になる．基準発話数のデフォルト値は 225 語となっており，この値は予備実験の結果をもとに暫定的に算出した．また，基準発話数を再計測する機能も実装した．“基準発話数を計測” ボタンを押した瞬間から 3 分間グループ内の話数を計測し，1 分辺りの平均値を基準発話数とする．

また，効果音や BGM はデフォルトで OFF になっている．それぞれのボタンを押すことによって効果音や BGM を提示する機能が ON に変更される．

同様に，発話数によって BGM のテンポを自動で変更する機能もボタンで ON ／ OFF を切り替えることができる．BGM のテンポ自動切り替えは，基準発話数と一定時間内の発話数との差分によって遅い・普通・速いのどれかに切り替わり，閾値は基準発話数の ±0 ∼ 50% の範囲で選択できる．また，BGM を切り替える頻度も，1〜60 分の間で選択できる．

また，任意の時間を設定し，その時間に到達した際に「蛍の光」を提示する機能を実装した．この機能によって，オンライン飲み会などの辞め時が難しい問題を解消するきっかけを作りたいと考えた．

6.　評価実験

6.1　目的

カジュアルな遠隔コミュニケーション場面を想定して本

ⓒ2023 Information Processing Society of Japan

システムを利用することで，コミュニケーションへの影響や主観的な印象を調査する．

6.2 手法

3 名を 1 グループとし 6 グループ (以降 g1〜g6) を被験者とする．Zoom を使って雑談をしながらユーザ操作アプリを使うよう指示した．手順として，最初にユーザ操作アプリの操作方法を参加者に教示し，操作に慣れるための 3 分間の練習セッションを行うよう指示した．練習セッションの後，BGM あり・なしの本番セッションを各 6 分間行った．なお，実験順序は順序効果に配慮して入れ替えた．また，各セッションでは予備実験同様に話題を提示した．BGM のあるセッションでは，システムのテンポ自動変更機能によって，発話数の変化に合わせて BGM のテンポを変化させた．具体的には，2 分毎に基準発話数と発話数の比較を行い，g1，g2 では基準発話数の ±20% を閾値とし，g3〜g6 では基準発話数の ±10% を閾値とした．予備実験での知見より，今回の実験の基準発話数は 225 語とした．また，2 回目のセッションの後，蛍の光を提示した．その後アンケート調査を行なった．

6.3 結果と考察

6.3.1 BGM による発話数の変化

実験中にテンポが遅く変化したのはグループ 5 の 1 つのみのグループ，速く変化したのは．グループ 3，4，6 の 3 つのグループだった．テンポが変化した際の発話数の変化率を，変化後 2 分間の発話数/変化前 2 分間の発話数として算出した．算出した変化率や関連するアンケートの結果を表 4 に示す．

表 4 各グループの発話数の変化率と関連するアンケート結果（1. ネガティブ-5. ポジティブ）

	変化率	BGM の テンポ変化	テンポ変化に 気がついたか	BGM や その変化に 影響されたか
g5	1.18 (376/319)	普通→遅い	4.67 (0.58)	4.00 (0.00)
g3	0.92 (585/637)	普通→速い	1.00 (0.00)	1.00 (0.00)
g4	1.21 (676/557)		3.67 (2.31)	1.67 (1.15)
g6	1.14 (582/510)		3.67 (1.53)	1.00 (0.00)

遅いに変化したグループでは，予備実験の結果と同じく発話数が増加した．自由記述では，「会話が詰まったような感じがした」といった記述が見られ，会話の詰まりによって生まれた間を埋めるために発話が増えた可能性がある．

一方，速いに変化した 3 グループでは発話数の増減がどちらも見られた．その中でも，発話数が減少したグループ 3 では，「テンポ変化に気がついたか」のアンケート結果が 1.00 であったことから，テンポが変わったことに全く気がついていなかった．

次に，各グループのセッション毎の発話数とその際の話題の話しやすさのグラフをそれぞれ図 7 と図 8 として表す．

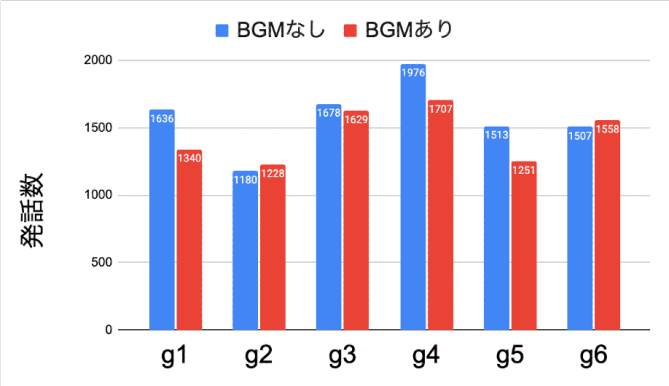

図 7 各グループのセッション毎の発話数

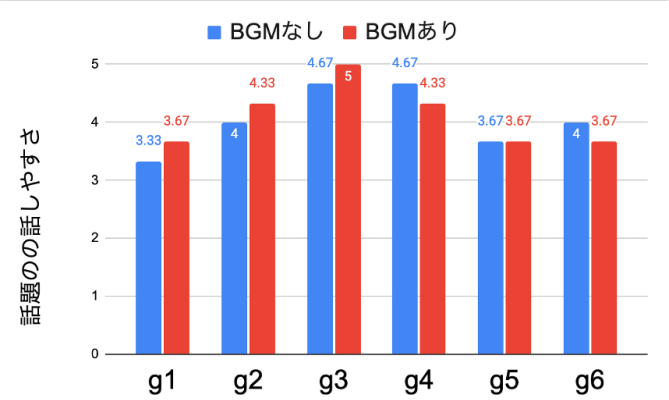

図 8 各グループのセッション毎の話題の話しやすさ（1. ネガティブ-5. ポジティブ）

図 7 と図 8 より，発話数と話題の話しやすさについて明確な相関は発見できなかった．また，グループ 1，4，5 では BGM が無いセッションに比べて BGM があるセッションでの発話数が小さくなり，他のグループでは発話数に大きな差が見られなかった．この時，自由記述には「BGM があると間があっても気分的に楽」と行った記述が見られ，間を気にしなくなった分，発話数が減少したと考える．

6.3.2 効果音ボタンについて

効果音ボタンの使いやすさと各効果音のイメージついてのアンケート結果を表 5 に，1 セッション中に効果音ボタンが押された回数を強弱別にまとめた結果を表 6 に示す．

「音ボタンの使いやすさ」の結果の平均と標準偏差は 2.67(1.14) と低い結果になった．自由記述から，「会話を遮ってしまう」「無意識に押すのが難しい」といった記述や，「長押しによって会話に対する反応が遅れてしまう」というボタンの仕様についてのマイナスな記述が多く見られ

表 5 「音ボタンの使いやすさ」と各効果音について「イメージとあっていたか」のアンケート結果 (1. ネガティブ-5. ポジティブ)

	平均 (標準偏差)
音ボタンの使いやすさ	2.67(1.14)
「喜び」の効果音の適切さ	4.22(0.65)
「悲しみ」の効果音の適切さ	3.89(1.08)
「不満」の効果音の適切さ	3.67(1.50)
「拍手」の効果音の適切さ	4.72(0.75)

表 6 1 セッション中に効果音ボタンが押された回数（強弱別）

	弱	中	強
押された回数 平均（標準偏差）	5.25 (4.37)	1.42 (2.11)	0.92 (1.16)

た．各効果音ボタンについてセッション毎に押された回数を強弱別で見ると，弱の効果音が平均 5.25 回に対し，中が 1.42 回，強が 0.92 回となった (表 6)．この点からも，長押しによる音の切り替えは使いにくかったと推察される．

次に，各効果音のイメージについての結果は，表 5 から，「喜び」が 4.22「拍手」が 4.72 と高めの評価となった．「不満」と「悲しみ」についても，それぞれ 3.89，3.67 とやや高めとなったがばらつきが見られた．「不満」の音はブーイングの効果音を元に作成したため，日本人にはややなじみにくかった可能性がある．また，1 セッション中に効果音ボタンが押された回数を種類別にまとめた結果を表 7 に示す．「喜び」の音はイメージとは合っていたがあまり利用されておらず，一方「悲しみ」の音はイメージとはさほどあっていないが，最も多く利用されていた．また，「不満」と「悲しみ」の効果音を提示した際に笑いが生まれるといった場面も見られた．

表 7 1 セッション中に効果音ボタンが押された回数（種類別）

	平均 (標準偏差)
喜び	0.33(0.65)
悲しみ	3.00(2.17)
不満	1.67(2.46)
拍手	2.08(3.70)

6.3.3 蛍の光と自由記述

「蛍の光がかかった時にコミュニケーションに影響があったか」というアンケート (1. ネガティブ-5. ポジティブ) の結果は，被験者全体の平均が 4.56，標準偏差が 0.78 と高い値となった．自由記述では，「そろそろ終わりの時間だということを認識できた」なとの記述が多く，コミュニケーションの終わりを演出できていることがわかった．

また，自由記述として，「カメラオフの状態で使用する際，表情などの代わりに効果音を挿入することで他の参加者に言葉のニュアンンスやリアクションを伝えやすいのではないか」というような記述が複数見られた．

6.3.4 まとめ

BGM と発話数に関するまとめとして，BGM のテンポが遅くなった時，発話数が増加しており，テンポが遅くなることで生まれた間を埋めるために発話が増えた可能性がある．また，テンポが速くなった時は発話数の増減がどちらも見られた．テンポの変化別に見ると，遅い変化は速いテンポに比べ気づかれやすく，変化による影響も多く感じられていた．

効果音に関するまとめとして，効果音のイメージは全体としてやや高めの評価となり，一部の音では笑いが誘発されていたことから，効果音とその提示される場面によっては，コミュニケーションを楽しくする効果などが期待できる．一方，効果音ボタンの長押しによる効果音の切り替えが使いづらいことがわかったため，フリックによる効果音の切り替えなどの改善手法を実装していきたい．

また，蛍の光は「コミュニケーションに影響があったか」のアンケート結果が平均 4.56(0.78) と高い値であり，自由記述も踏まえてコミュニケーションの終わりをうまく演出できていたと考える．

7. 議論

7.1 効果音提示のタイミングについて

効果音に関するアンケートでは，「会話を遮ってしまう」と行った記述が複数見られた．実験の様子からも，発話している最中に効果音が鳴ることで，話者が発話を中止し効果音に対するリアクションを取ると行った場面が観察された．効果音の提示はコミュニケーションへの割り込み感が強く，提示タイミングによってはコミュニケーションを阻害する可能性がある．現状のシステムでは，効果音を提示する回数やタイミングに制限がないが，例えば話者の発話が認識されている間は効果音の音量を抑えたりミュートするといった実装を検討していきたい．

7.2 システムの利用場面の可能性

自由記述から，カメラオフの時にリアクションなどを伝えやすいといった記述が複数見られた．それを踏まえ，例えばオンライン講演会などの，講演者と聴衆がいるようなシチュエーションで本システムを利用することで，講演者と聴衆の間でのコミュニケーションを支援することができるのではないかと考えた．現状のオンライン講演会は，基本的にマイクによる音声や，チャット機能を利用した文字によってコミュニケーションが行われる．本システムによる効果音挿入によって，コミュニケーションの幅を広げられるのではないかと考えた．一方で，上で述べたように，効果音による割り込み感により講演を阻害してしまうことも考えられるため，システムの制限についても考える必要がある．今後は，カジュアルな遠隔コミュニケーション場面に限らず，システムの利用場面について視野を広げてい

きたい.

8. まとめと展望

本研究では，オンライン飲み会のような多人数でのカジュアルな遠隔コミュニケーション場面において，議題等がなく無く会話が停滞しがちと行った問題や，細やかな感情や雰囲気が共有されにくい問題，やめ時が難しいといった問題を軽減するために，インタラクティブに BGM や効果音を挿入する手法を提案した．この提案に合わせ，予備実験の知見を元にユーザの発話数やボタン操作などの情報を取得するユーザ操作アプリと，発話量に応じて BGM のテンポを自動変更する音提示アプリの作成を行なった．システムを使用した実験では，カジュアルなオンラインコミュニケーション中に，一部では BGM のテンポ変化による発話の増加や効果音提示によって笑いを生む場面が見られた.

今後の展望として，BGM 変更機能の効果を多グループで検証し，テンポによる発話誘導の可能性を引き続き調査していく．また，効果音ボタンを，フリックによる効果音の切り替え機能の実装や，効果音提示による割り込み感への配慮などの改善を進めていきたい.

参考文献

[1] 高口央.BGM は購買意欲をかき立てるか，流通経済大学社会学部論叢，2011，22.1:p.89-100.
[2] 大野直紀，徳久弘樹，中村聡史. 自身のみ聴取可能な音楽を用いたコミュニケーション円滑化手法の提案，情報処理学会研究報告グループウェアとネットワークサービス（GN），Vol.2019-GN-106，No.22，pp.17，2019.
[3] 高橋幸子，山本賢司，松浦信典，伊賀富栄，志水哲雄，白倉克之. 音楽聴取が情動に与える変化について: 音楽聴取前後の POMS スコアの変化を中心として. 心身医学，Vol.39，No.2，p.167-175，1999.
[4] 長井彩乃，相澤裕貴，須田翔悟，渡邊恵太．SakeEffects：アルコールセンサを用いたオンライン飲み会を盛り上げるための顔エフェクトシステム，インタラクション 2021 論文集，2021，187-188
[5] 小山環，塩澤秀和.”参加者の同期的なジェスチャーを支援するビデオ会議システム.”マルチメディア，分散，協調とモバイルシンポジウム 2022 論文集，2022，1041-1049.
[6] 田坂佑太，大井翔，国山字弘，佐野睦夫． いざガヤ：リモートコミュニケーション環境における背景雑音が臨場感に与える影響の考察，研究報告ヒューマンコンピュータインタラクション，2021-HCI-192，38，2021
[7] 徳原耕亮，荒川豊，石田繁巳.”頷きのリアルタイムフィードバックによるビデオ会議支援手法の提案.”マルチメディア，分散，協調とモバイルシンポジウム 2021 論文集，2021，953-959.
[8] 二瓶芙巳雄，田口和佳奈，中野有紀子，深澤伸一，赤津裕子．“多人数遠隔コミュニケーションにおける肯定的感情表出支援の効果と支援適用タイミングの決定”，情報処理学会論文誌，2021，Vol.62，No.2，p.761-771.
[9] 大槻麻衣，丸山啓太，葛岡英明，鈴木雄介.“ThirdEye: 遠隔対話システムにおける対話者の視線を提示する眼球ディスプレイの開発と評価”，ヒューマンインタフェース学会論文誌，2017，Vol.19，No.2，p.211-218.
[10] Shimizu, Rurika., Takase, Yutaka., and Nakano, Yukiko. I., “Audio Based Group Conversation Support System”, Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers，2017，UbiComp’17，p.197–200.
[11] セイコーインスツル株式会社．“やってみよう・メトロノーム記号の読み方”．SEIKO．2022-06-13. https://www.sii.co.jp/music/try/metronome/01.html,（参照 2023-0108）