

卒業論文 2016 年度（平成 28 年度）

Gyaon

## 音声活用の幅を広げる録音 / 共有システム

慶應義塾大学 環境情報学部

佐竹 紘明

増井俊之研究会

2017 年 1 月

# 卒業論文 2016年度（平成28年度）

Gyaon

## 音声活用の幅を広げる録音/共有システム

### 論文要旨

小型の録音装置は古くから販売されており、またスマートフォンには様々な音声録音アプリが存在するが、音声メモを記録したり様々な音を録音したりして活用している人は少ない。音声が広く活用されていないのは、録音の手間が繁雑であり、音の再生/検索が簡単でないためだと思われる。本研究では、既存の録音システムの不便を解消し、音声活用を促す録音システムを提案する。

### キーワード

音声、録音、インターフェース

慶應義塾大学 環境情報学部

佐竹 紘明

# 目 次

<b>第 1 章 序論</b>	<b>1</b>
1.1 背景 . . . . .	2
1.1.1 音声情報の有用性 . . . . .	2
1.1.2 音声が活用されない理由 . . . . .	2
1.2 本研究の目的 . . . . .	3
1.3 本論文の構成 . . . . .	3
<b>第 2 章 システムの提案</b>	<b>4</b>
2.1 設計指針 . . . . .	5
2.2 利用例 . . . . .	5
2.2.1 PC 版 . . . . .	5
2.2.2 Android 版 . . . . .	9
<b>第 3 章 実装</b>	<b>11</b>
3.1 システム構成 . . . . .	12
3.2 サーバ . . . . .	12
3.2.1 DB スキーマ . . . . .	12
3.3 PC 版クライアント . . . . .	12
3.3.1 WebAudioAPI . . . . .	12
3.3.2 音声リストの同期 . . . . .	13
3.3.3 プリレコーディング機能 . . . . .	13
3.3.4 Gyaon キー . . . . .	13
3.4 Android 版クライアント . . . . .	13
<b>第 4 章 関連研究</b>	<b>14</b>
<b>第 5 章 考察</b>	<b>16</b>
5.1 単純な録音操作による音声活用の促進 . . . . .	17
5.2 外部システムでの音声活用の促進 . . . . .	17
5.3 プライバシー問題 . . . . .	19
<b>第 6 章 結論</b>	<b>20</b>
6.1 本研究の成果 . . . . .	21

謝辭	22
參考文献	23

## 図 目 次

2.1 録音ボタン	5
2.2 音声リスト	6
2.3 Scrapbox のオーディオ記法	6
2.4 地図機能	7
2.5 著者の PC の Gyaon キー	8
2.6 ペダルの例	9
2.7 アプリケーションの起動画面	10
2.8 録音ボタンが配置された Android のホーム画面	10
5.1 音声素材ページ	18
5.2 音声つき単語帳	19

# **第1章 序論**

本章では、本研究の背景と目的、および本論文の構成について述べる。

## 1.1 背景

小型の録音装置は古くから販売されており、またスマートフォンでは様々な録音アプリケーションを利用できる。これらは音楽に携わる人やインタビュー/会議を記録するといった用途に重宝されるが、日常的にメモを記録するなどして活用している人は少ない。

### 1.1.1 音声情報の有用性

音声情報には、以下のような有用性が存在する。

#### 1. 嘁るだけで記録できる

2017年現在、ネット上での情報記録 / 情報交換はテキストベースのものがほとんどであるが、録音ならPC / スマートフォンへのキー入力を行う必要がなく、記録したいことを喋るだけで良い。日本の若者を中心に広く利用されているコミュニケーションサービスLINE<sup>1</sup>ではボイスメッセージ機能が利用でき、その簡単さから一部のユーザーに積極的に活用されている。

#### 2. 記録しながら同時に別のことを行える

長時間の録音でも、始めてしまえば記録に集中する必要がないので、同時に別のことを行える。これはテキストや写真、動画による記録とは異なる点である。

#### 3. 他のメディアとの相互補完性がある

また、写真や文章といった他のメディアとの組み合わせによって、互いの内容理解を助ける相互補完性の存在が知られている[6]。

以上のように、音声情報は優れた情報記録 / 情報交換の手段となりうる可能性を秘めている。

### 1.1.2 音声が活用されない理由

音声の有用性が知られているにも関わらず多くの人に活用されないのは、便利に使える録音システムが存在しないためだと思われる。

既存の録音システムには、以下のような問題点が見受けられる。

#### 1. 録音/再生の手順が煩雑

一般的に録音/停止/再生それぞれでボタン操作が必要であり、音声記録の単純さをインターフェースが阻害している。

---

<sup>1</sup><https://line.me/ja/>

## 2. 管理しづらい

録音時刻/再生時間といった基本的なメタデータや、自ら設定したタグ等をもとに音声を検索できるが、音声が増えていくほど、内容理解を助ける仕組みとして十分とはいえない。

## 3. 音声を再利用しづらい

録音システムの中で音声を管理しなければならないことが多く、外部アプリケーション / サービスでの音声活用の可能性を狭めている。

音声を最大限活用するためには、これらの問題を解決する新しい録音システムが必要である。

## 1.2 本研究の目的

既存の録音システムに見られる不便を解消し、音声を有効活用できるようなシステムを開発することが本研究の目的である。

## 1.3 本論文の構成

第1章では、本研究における背景と問題意識、目的について述べた。第2章では、第1章で述べた問題意識を踏まえ、新しい録音システムを提案する。第3章ではシステムの実装に関して述べ、第4章では関連研究について述べる。第5章ではシステムの考察を行い、第6章では本研究を総括する。

## 第2章 システムの提案

本章では、音声利用についての背景を踏まえ、新しい録音システム「Gyaon」を提案する。

## 2.1 設計指針

Gyaon では

- 単純な録音操作
- 音声管理の簡単さ
- 他システムからの音声の利用のしやすさ

を優先した設計を行うことで録音の不便を解消し、音声の有効活用を促すシステムを目指している。

## 2.2 利用例

様々な利用シーンに対応するため、PC 用 Web アプリケーションおよび Android スマートフォン用アプリケーションにて Gyaon システムが利用できる。それぞれのプラットフォームにおける利用例を述べる。

### 2.2.1 PC 版

#### (1) 基本操作

PC 版 Gyaon は <https://gyaon.herokuapp.com/> にて試験運用しており、Web ブラウザで動作する。図 2.1 の赤いボタンが録音ボタンであり、押している間だけ録音する。ボタンを離すと録音が停止し、プレビュー再生が行われる。同時に音声データがサーバへアップロードされる。アップロードが完了したものは音声リスト（図 2.2）に表示され、マウスカーソルを重ねるだけで再生できる。

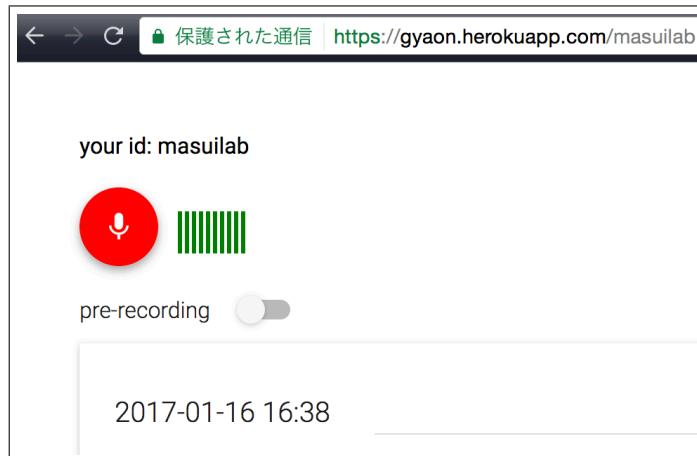


図 2.1: 録音ボタン

日付	音声ファイル名	長さ	X	削除
2016-10-25 14:42	Hikaruのピエゾで録音 at MacMini	00:01	X	□
2016-10-25 13:52	うさぎ算(hagino)	01:53	X	□
2016-10-25 11:52	ツインピーチも見よう	00:02	X	□
2016-10-24 20:37	シリコンバーを見よ Making the world a b	00:02	X	□
2016-10-24 19:43	スクランブル交差点	00:15	X	□
2016-10-24 19:41	研究会が盛り上がってます	00:05	X	□
2016-10-24 19:04	Gyaonで授業しちゃう (masui)	00:02	X	□
2016-10-24 19:04	しゃべらずに生活する (masui)	00:05	X	□
2016-10-24 18:55	ペダルはいいぞ (satake)	00:01	X	□

図 2.2: 音声リスト

## (2) コメントの追加

音声リスト内にコメント入力欄があり、クリックすると入力モードに移行する。音声の内容などを自由に記録することができる。

### (3) 音声 URL の取得

音声リストの右側のコピーボタンをクリックすることで、PC のクリップボードに音声データの URL がコピーされる。そのままアクセスすれば音声をダウンロードでき、ローカル環境のアプリケーションなどで利用できるほか、URL を経由して他人に音声を渡すことも可能である。また Nota.inc<sup>1</sup> の Wiki システム「Scrapbox」<sup>2</sup> ではオーディオ記法を利用でき、決められた書式で音声の URL を入力すると、Gyaon の音声リストと同様にマウスカーソルによる再生が可能となる。

- ④.。oO (今日もいい天気☆)♪
- ④.。oO (知らん!)♪
- ④.。oO (伊藤くんがやると言ってました)♪
- ④.。oO (一理ある)♪
- [早川匠.icon]。oO ([そうだね)  
<https://gyao.n.herokuapp.com/sounds/hykwtakumin/db6f222e6ae43f5b891ee830cc156e8f.wav>]
- ④.。oO (そうだよ(便乗))♪
- ④.。oO (そうかも)♪
- ④.。oO (感動した)♪
- ④.。oO (いい話だ)♪
- ④.。oO (では誰か発表をお願いします)♪
- ④.。oO (ここで一句)♪

図 2.3: Scrapbox のオーディオ記法

<sup>1</sup><http://www.notainc.com/ja/>

<sup>2</sup><https://scrapbox.io/>

#### (4) プリレコーディング機能

図 2.1 の「pre-recording」スイッチを ON にするとアプリ起動中は常時録音し、録音ボタンを押す 10 秒前から音声を記録できる。もし突然録りたい音が聞こえてきた場合でも逃さず録音することが可能である。

#### (5) 地図機能

音声データに紐付けられた位置情報をもとに、音声を地図上にマッピングして一覧できる<sup>3</sup>。音声の録音場所が分かることから、その場所の雰囲気や、景色がいい/この店は美味しいといったその場所ならではの有用な情報を記録/共有できる。

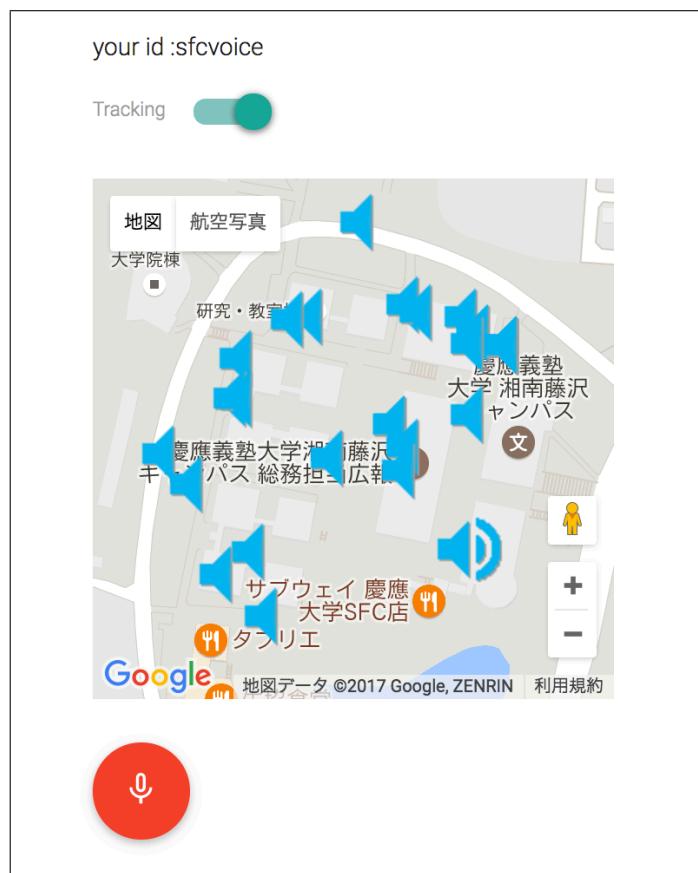


図 2.4: 地図機能

#### (6) ID の共有

Gyaon では ID ごとに音声を管理しており、URL 末尾でユーザ ID を指定できる。この ID を複数人で共有することで Gyaon を共有タスクリストのように利用したり、ボイスメッセンジャーのような使い方が可能である。

<sup>3</sup>地図機能は <https://gyaon.herokuapp.com/map/> にて試験運用中

### ユーザ ID の指定

<https://gyaon.herokuapp.com/>{ ユーザ ID}

### (7) Gyaon キー

Gyaon キーアプリケーション<sup>4</sup>を利用すれば、PC の任意のキーを録音ボタンに設定できる。これにより、アプリケーションを起動することなく録音可能である。Web アプリケーションと同様に押している間録音、自動アップロードが行われる。



図 2.5: 著者の PC の Gyaon キー

### (8) Gyaon ペダル

USB 接続の汎用なペダル(図 2.6)を接続すれば、楽器演奏/運転中/料理中といった手を使えないシーンでもハンズフリー録音が可能となる。また、ペダルとマイクを接続した Raspberry Pi 等を複数用意し、録音専用コンピュータとしてセットアップすることで、ユビキタスな録音環境を自宅や研究室等に構築可能である。

<sup>4</sup>Mac 用アプリケーションを <https://github.com/stkay/GyaonWithKarabiner> にて配布中



図 2.6: ペダルの例

## 2.2.2 Android 版

### (1) 基本操作

Android 用 Gyaon アプリケーションは録音のみ可能で、PC 版と同様に押している間録音/自動アップロードが利用できる。(図 2.7). 「Your ID」には任意のユーザ ID を入力する。PC 版で使用しているユーザ ID を入力すれば、スマートフォンでの録音を PC でも聞けるようになる。

### (2) 録音ボタンを常駐させる

図 2.7 の「STARTSERVICE」ボタンを押すと、録音ボタンをスマートフォン上に常駐させることができる(図 2.8)。

### (3) 利用シーン

スマートフォンを持ち運ぶる環境ならいつでもどこでも録音可能である。外出先で思いついたことを記録したり、鳥のさえずり/川のせせらぎといった環境音や、アナウンス/街頭演説といった情報を共有する用途にも活用できる。

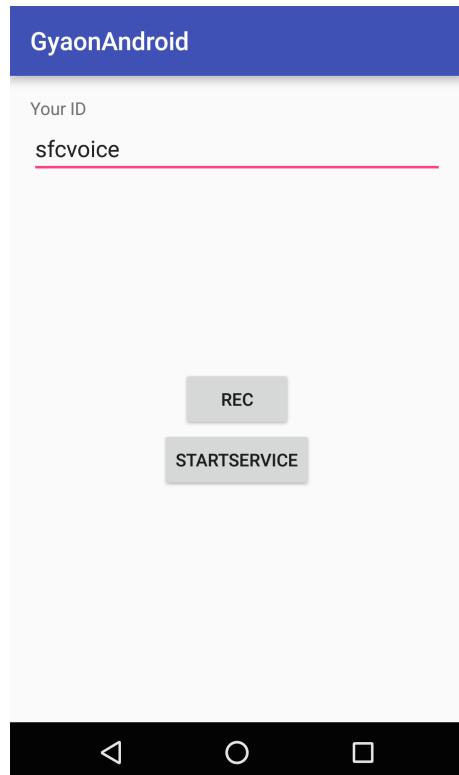


図 2.7: アプリケーションの起動画面

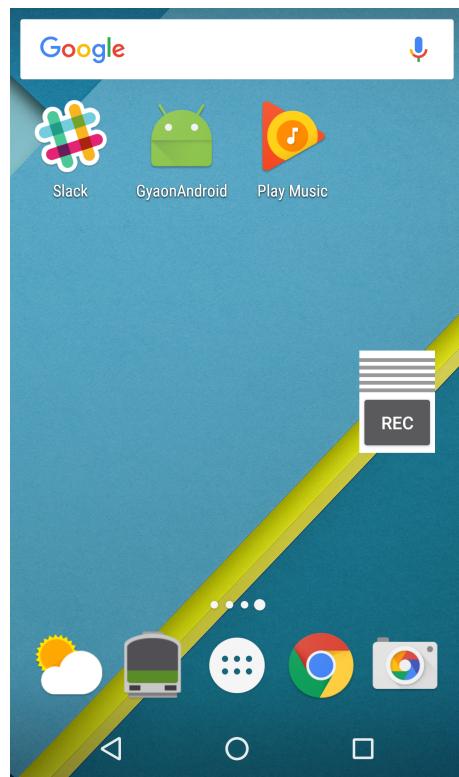


図 2.8: 録音ボタンが配置された Android のホーム画面

## 第3章 実装

本章では、第2章で述べたシステムの設計を受け、Gyaon の実装について述べる。

## 3.1 システム構成

Gyaon は、ユーザが実際に録音/再生するためのクライアントアプリケーションと、アップロードされた音声を保存/管理するサーバから構成される。構成図を図??に示す。

## 3.2 サーバ

サーバは Node.js<sup>1</sup> とその Web アプリケーションフレームワークである Express<sup>2</sup> によって実装されている。音声データの保存を Amazon S3<sup>3</sup>、DB 管理を mLab<sup>4</sup>にて行っている。

### 3.2.1 DB スキーマ

DB では以下のようなスキーマを定義し、音声データを管理している。

```
1 var soundSchema = mongoose.Schema({
2     lastmodified: Date, /* 更新時刻 */
3     user: String, /* ユーザID */
4     name: String, /* ファイル名 */
5     key: String, /* S3key */
6     size: Number, /* ファイルサイズ */
7     time: Number, /* 再生時間 */
8     comment: String, /* コメント */
9     location_x: Number, /* 位置情報 */
10    location_y: Number
11 })
```

## 3.3 PC 版クライアント

PC 版クライアントは HTML/CSS/JavaScript で実装されており、ブラウザ上の Web アプリケーションとして動作する。

### 3.3.1 WebAudioAPI

録音/再生といった主要な音声機能は、HTML5 audio<sup>5</sup> 並びに WebAudioAPI<sup>6</sup>によって実装されている。

HTML5 audio は HTML ドキュメント内に音声を埋め込むことができ、以下のような Audio 要素を宣言することで利用できる。

<sup>1</sup><https://nodejs.org>

<sup>2</sup><http://expressjs.com/ja/>

<sup>3</sup>Amazon Web Services によって提供されるオンラインストレージサービス。<https://aws.amazon.com/jp/s3/>

<sup>4</sup>MongoDB ベースのクラウド DB サービス。<https://mlab.com/>

<sup>5</sup><https://www.w3.org/TR/html5/embedded-content-0.html#the-audio-element>

<sup>6</sup><https://www.w3.org/TR/webaudio/>

---

```
1 <audio src="/test/audio.ogg"></audio>
```

---

WebAudioAPI は高度な音声処理や音声合成を行える JavaScript API であり、 Audio 要素を操作したり、 エフェクトを加えることができる。

### 3.3.2 音声リストの同期

複数のユーザが同じユーザ ID を使用していても、音声リストはリアルタイムに同期される。これは Node.js の WebSocket ライブライアリ「Socket.IO<sup>7</sup>」によって実装されており、サーバに音声がアップロードされると、音声リストの更新情報がブラウザに通知される。

### 3.3.3 プリレコーディング機能

---

```
1 preRecScriptProcessor.onaudioprocess = (event) => {
2   const channel = event.inputBuffer.getChannelData(0)
3   if(preRecAudioBufferArray.length * BUFFER_SIZE > SAMPLE_RATE * PREREC_SEC) {
4     preRecAudioBufferArray.shift()
5   }
6   preRecAudioBufferArray.push(new Float32Array(channel))
7 }
```

---

### 3.3.4 Gyaon キー

## 3.4 Android 版クライアント

Android 版クライアントは Java で実装されており、通常の Android アプリケーションとして動作する。

---

<sup>7</sup><http://socket.io>

## 第4章 関連研究

本章では、音声利用に関する関連研究を概観する。

日常的な体験や雰囲気を記録するための録音デバイスについて研究が行われている [2]. 録音は何かを記録、説明したり、思い出すのに重要な手段だが、日常的に行っている人は少ないと指摘されている。より気軽に録音できるよう日用品にセンサーを埋め込み、状況の変化に応じて録音/再生するデバイスが提案されている。日用品に溶け込み、単純明快な録音/再生インターフェースが実現されているが、本を閉じる/開くといった二値的な状況変化でしか操作できないため、デバイス単体で複数の音声を管理することが困難になっている。

人間生活に関わる情報を長期間にわたり記録するライフログ関連研究では、そのログの一部として音声データがよく利用されている [1]. 音声データの特徴量を計算し、閾値処理するインターフェースを提供することで再生区間を決定する手法 [9] や、講義ノートといった手書きメモに対して音声データを関連付ける手法 [3] などが提案されている。

研究ノートの補助的な記録手段として音声データを利用するシステムが提案されている [5]. 事前に用意された研究ノートから形態素解析などをを利用してキーワードを抽出し、音声データに付与する手法がとられている。

また、音声データの特定のタイミングにタグを付与できる録音システムが開発されている [8]. 学生がノートテイキングの補助として使うことを想定しており、重要な説明等を逃さないよう素早くタグ付けが行えるインターフェースが実装されている。

音声認識技術を活用したテキストによる音声検索手法が提案されている [4]. 試作されたアプリケーションでは音声データに含まれる単語からキーワード検索が可能となっているほか、音声認識の信頼度を文字色に反映させたり、ストップワードを見えにくくするなどしてユーザが録音の要点を思い出しやすくなるよう配慮されている。

写真と音声を組み合わせによる体験記録手法が提案されている [6]. 写真と手書きだけでは表現できない雰囲気を音声に記録し、写真によって音声データに一覧性を持たせる仕組みである。写真が音声の内容把握を助け、音声データの価値を高めることも確認されている。

また、同様の手法によって撮影された写真を、音声とともに閲覧できる Web サイトが公開されている [7].

## 第5章 考察

本章では、Gyaonについての考察を行う。

## 5.1 単純な録音操作による音声活用の促進

Gyaon では録音ボタンを押している間録音し、停止後直ちに音声のアップロードが開始される手順となっており、単純な録音操作を実現している。著者は開発時から継続的に Gyaon を使用しており、単純な録音操作によって、録音行為に対する心理的障壁が低くなったと感じている。このことから、今まで手書きやキーボード入力していたメモなどを全て録音するようになり、音声を活用する場面が増えた。その他にも、様々な場面で音声が活用された。

**研究会での活用** 著者の所属する増井研究会では様々な素材の叩打音を分析する研究が行われており、サンプリングのために Gyaon が活用された。音声をすぐに確認でき、コメントを付けられる点が好評であった。

また大学が主催する研究発表イベントではリアルタイムな情報共有に活用された。ブース展示がメインのイベントであったため、面白い展示を見つけた学生が Gyaon で共有したり、楽器のブースでは演奏を録音するなどした。増井研究会のブースでは常に Gyaon を起動していたので、有益な情報を見逃すことなく受け取ることができた。

**雰囲気記録システムとしての活用** スマートフォンでも簡単に録音できることから、その場の雰囲気を記録する用途に活用された。鳥のさえずりや川のせせらぎといった環境音を記録して自宅で聞いたり、研究会のミーティングを録音し、それを振り返りながらメンバーと懐かしむような場面もあった。

**楽器練習支援システムとしての活用** Gyaon ペダルとプレビュー再生機能の組み合わせによって、楽器練習支援システムとして便利に活用できた。楽器演奏において自分の演奏を客観的に評価することは難しいが、録音を行うことによって解決される。Gyaon ペダルを利用すれば演奏中でも録音でき、録音後のプレビュー再生で演奏をすぐに確認することができる。演奏/録音/確認を素早く行えることから、従来の録音システムでは実現できなかった新しい練習形態を生み出したといえる。

## 5.2 外部システムでの音声活用の促進

Gyaon は音声をアップロードし、URL によって音声データを簡単に取得できるため、Gyaon 以外のシステムでも音声活用が促進された。

**DTM ソフトでの活用** PC 上で音楽制作を行う DTM ソフトには、録音したり、取り込んだ外部音声ファイルを 1 つの楽器として利用するサンプラー機能が存在する。Gyaon で録音した様々な音声をダウンロードして DTM ソフトに取り込み、作曲に利用することができた。

**「Scrapbox」での活用** Scrapbox では音声を文章に埋め込むオーディオ記法が積極的に活用された。音声の指定には URL を利用するため Gyaon との親和性が高く、以下のような活用例が見受けられた。

### 1. 発表資料を作る

増井研究会ではミーティングで使用する発表資料を Scrapbox で作成している。文章や画像といった資料に加え、予め録音した発表原稿もページに埋め込むことで、発表者が喋らなくてもプレゼンテーションを行えるようになった。

### 2. 音声素材をまとめる

積極的に自分の声を録音する研究会メンバーがおり、図 5.1 のような音声素材ページが作成された。ミーティング中に再生して笑いを誘つたり、簡易的な合成音声として活用された。

何か書くと早川が録音して返す

- (t).oO (ますいけん、はっじまるよ～!)♪
  - 「ゆるゆり、はっじまつるよ～」と同じ感じで言ってほしい
- (t).oO (ここら辺で休憩にしましょう)♪
- (t).oO (伊藤くん、ちょっと静かにしてくれないかな)♪
- (t).oO (なるほど)♪
- (t).oO (なるほど2)♪
- (t).oO (なるほど!勉強になりました!)♪
- (t).oO (わからん)♪
- (t).oO (たしかに)♪
- (t).oO (ありがとうございます)♪
- (t).oO (デルタで)♪
- (t).oO (早川さんが)♪
- (t).oO (伊藤さんが)♪
- (t).oO (佐竹さんが)♪
- (t).oO (ドアをひらきました)♪
- (t).oO (電気をつけました)♪
- (t).oO (電気がつきました)♪

図 5.1: 音声素材ページ

### 3. 単語帳を作る

発音が難しい英単語等をまとめ、その音声を埋め込んだ図 5.2 のような単語帳が作られた。マウスカーソルを単語に重ねるだけで発音を確認できるので、より印象に残るものとなった。

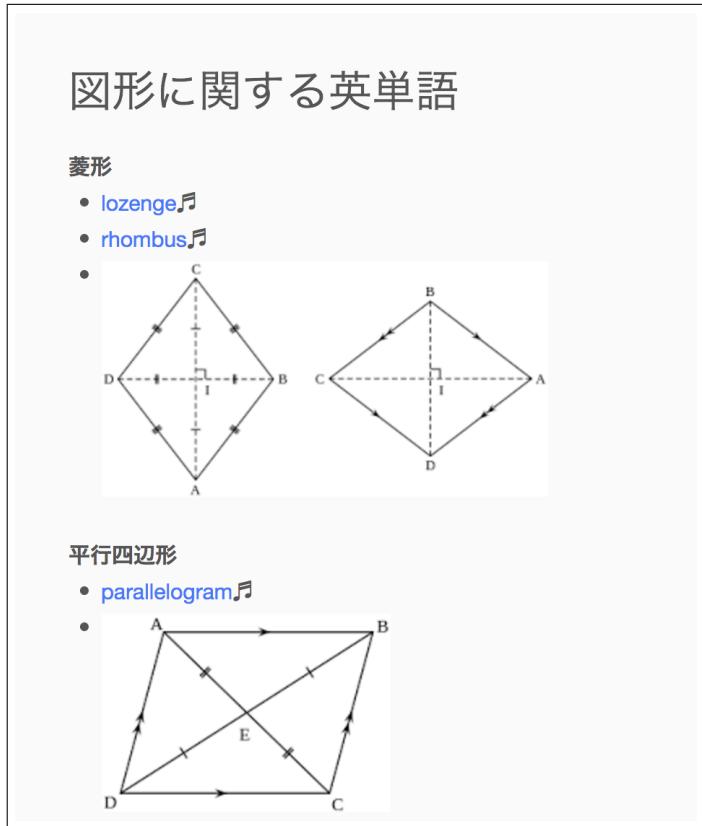


図 5.2: 音声つき単語帳

### 5.3 プライバシー問題

簡単に録音できるため他者に対するプライバシー侵害の可能性があり、社会で導入されることに対する心理的障壁は非常に高いと考えられる [9].

また、現在の Gyaon システムでは全ての音声が公開されており、誰でもアクセス可能である。正式なサービスとして運用する場合は、アカウント管理や閲覧権限の設定といった対策が必要である。

## **第6章 結論**

本章では、本研究で得られた成果について述べる。

## 6.1 本研究の成果

既存の録音システムの問題を明らかにし、新しい録音システム「Gyaon」を設計/開発した。そして、Gyaon の設計指針である

- 単純な録音操作
- 音声管理の簡単さ
- 他システムからの音声の利用のしやすさ

が録音機の不便を解消し、音声活用を促すことが明らかになった。

今後は、さらなる音声活用法の探求と録音機能の拡充を行っていきたい。

## **謝辞**

本論文の執筆において、担当教員である増井俊之教授に多大なるご指導と貢献をしていただきました。また、本システムの実装において、研究会OBの橋本 翔氏、桜井雄介氏に多くの貢献と助言を頂きました。環境情報学部の早川 匠氏には有用な音声素材を提供していただきました。この場を借りて感謝の意を表します。

## 参考文献

- [1] J. Gemmell, G. Bell, R. Lueder, S. Drucker, and C. Wong. MyLifeBits: Fulfilling the Memex Vision. In *Proceedings of the Tenth ACM International Conference on Multimedia, MULTIMEDIA '02*, pp. 235–238, New York, NY, USA, 2002. ACM.
- [2] I. Poupyrev, H. Oba, T. Ikeda, and E. Iwabuchi. Designing Embodied Interfaces for Casual Sound Recording Devices. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems, CHI EA '08*, pp. 2129–2134, New York, NY, USA, 2008. ACM.
- [3] L. Stifelman, B. Arons, and C. Schmandt. The Audio Notebook: Paper and Pen Interaction with Structured Speech. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '01*, pp. 182–189, New York, NY, USA, 2001. ACM.
- [4] S. Vemuri, C. Schmandt, W. Bender, S. Tellex, and B. Lassey. An Audio-Based Personal Memory Aid. In *UbiComp 2004: Ubiquitous Computing*, Vol. 3205 of *Lecture Notes in Computer Science*, pp. 400–417. Springer, 2004.
- [5] 川西 康介, 大平 茂輝, 松原 茂樹, 長尾 確. 研究活動における音声ログの整理と利用. 第74回全国大会講演論文集, 2012(1):723–724, mar 2012.
- [6] 中蔵 聰哉, 角 康之, 西田 豊明. 体験記録システムにおける写真撮影と音声録音の相互補完性. 人工知能学会全国大会論文集, 8:358–358, 2008.
- [7] 増井 俊之. 写録(しゃーろっく): 写真撮影と同時に録音した音を楽しむ. <http://www.pitecan.com/Sharoku/>.
- [8] 藤坂 祐史. Recture - 復習しやすい授業記録アプリ. <http://recture.com/>.
- [9] 河村 竜幸, 中西 英之, 石黒 浩. 連続音声録音を用いた会話体験の探索. 人工知能学会全国大会論文集, 21:1–4, 2007.