

情報工学コース卒業研究報告

**研究活動における音声ログの
整理と利用に関する研究**

2012年2月

川西 康介

概要

近年、人間の活動を記録し利用するライフログが注目されている。研究者は、研究活動におけるログとして研究ノートに記録することが多いが、より簡単な方法である音声による記録はまだ一般的ではなく、その理由として、利用が困難であるという問題点が挙げられる。音声ログの利用をより容易にするためには、音声ログに何らかの情報を付与し、適切な状況で利用できるように整理する必要がある。一般的な録音アプリケーションにおいては、タイトルやタグをユーザーが付与することで整理できるが、それらを手動で入力する場合、ユーザーにかかるコストは高い。また、音声認識技術を用いてテキストに書き起こし、検索を容易にする方法があるが、音声認識により得られるテキストのみでは、整理に必要な言葉が音声に含まれていない場合、必ずしも有効な整理を行うことができない。

そこで本研究では、研究活動における音声ログに着目し、音声ログに付与するタイトルとタグを、研究者が日頃から記録している研究ノートから抽出することで、容易に付与できる仕組みを実現した。

提案システムでは、まず研究者が記録するテキストによるノートの情報をノートサーバーで蓄積・管理する。ノートサーバーはノートの本文に形態素解析を施すことによりタグを抽出し、データベースに保存する。次に、録音アプリ iRec がノートのタイトルや抽出されたタグをダウンロードし、アプリの内部に保存する。ユーザーが iRec を利用して録音すると、録音終了時にノートタイトル一覧が表示され、ユーザーがその中から音声ログのタイトルとして適切なものを選択すると、タグ一覧が表示される。ユーザーは、選択したノートから抽出されたタグとユーザーが任意に作成したタグの2種類の中から、音声ログのタグとして適切なものを複数個任意に選択することができる。選択されたタイトルとタグは音声ログに付与され、ボイスサーバーで蓄積・管理される。

提案システムを評価するために2種類の被験者実験を行った。まず、タイトルとタグの入力にかかる時間を評価するため、提案システムを利用する場合と、手動でタイトルとタグを入力する場合の2通りの記録方法により音声ログの作成を行った。次に、音声ログに付与されたタイトルとタグの有効性を評価するため、前述した2通りの記録方法ごとに、音声ログの内容を利用した問題を出題し、それに回答する時間を測定した。その結果、提案システムを利用した方が、タイトルとタグの入力時間は短くなり、問題の回答時間は長くなることが分かった。また、アンケート調査により、音声ログを利用する際に役に立ったタグのうち、約77.8%が音声認識では得られないタグであり、提案システムによるタグの付与が有効であることが確認された。

目次

第1章	はじめに	1
第2章	研究活動における音声ログ	5
2.1	音声ログ	5
2.2	音声ログの整理に関する従来方法	6
2.2.1	タイトルやタグの付与	7
2.2.2	音声認識	8
2.3	研究ノートを利用した整理手法	8
第3章	音声ログの記録・整理手法	11
3.1	システムの概要	11
3.2	研究ノートからのタグ抽出	11
3.2.1	研究ノート作成システム	11
3.2.2	タグの抽出方法	14
3.3	録音アプリ iRec	15
3.3.1	音声ログの種類	15
3.3.2	iRec のシステム	16
3.3.3	タイトルの手動入力	21
3.4	音声ログの閲覧インタフェース	21
3.5	システムの利用	23
第4章	評価実験	27
4.1	音声ログ記録時のメタデータ入力に関する実験	27
4.1.1	実験の目的と方法	27
4.1.2	実験結果	28
4.1.3	考察	28
4.2	音声ログの検索に関する実験	34
4.2.1	実験の目的と方法	34
4.2.2	実験結果	36
4.2.3	考察	39
4.3	まとめ	43

第5章 音声ログの利用に関する展望	45
5.1 音声ログの研究ノートへの自動関連付け	45
5.2 適切なシーンでの音声ログのリマインダー	47
第6章 関連研究およびアプリケーション	49
6.1 音声ログの記録と利用に関する研究	49
6.2 タイトルやタグの付与可能な録音アプリケーション	49
6.3 写真撮影と音声録音の相互補完性に関する研究	50
6.4 音声認識を用いた音声の文字化と検索に関する研究	51
6.5 音声認識を用いたスケジューリングに関する研究	51
6.6 ミーティングなどの長時間の音声記録に関する研究	52
第7章 まとめと今後の課題	53
7.1 まとめ	53
7.2 今後の課題	53
7.2.1 大量のデータに基づく本手法の検証	53
7.2.2 タイトルおよびタグの付与方法	54
7.2.3 音声認識の利用	55
7.2.4 音声ログの検索	55
7.2.5 ミーティングなどの音声記録	56
謝辞	56
参考文献	59
付録	60

第1章 はじめに

コンピュータやインターネットなどの情報技術の急速な発達に伴い、人間の生活や活動などの記録をデジタルデータとして蓄積・管理する機会が増えてきている。これらはライフログと呼ばれ、その原型は Vannevar Bush が 1945 年の The Atlantic Monthly 誌の記事 “As We May Think” において提唱した Memex というシステム概念までさかのぼることができる [1]。Memex は、蓄積された情報が電氣的にリンクされており、ある情報から別の情報に容易にアクセスできるデバイスである。

ライフログの代表的な研究として、Microsoft 社の BARC Media Presence Group に所属する Gordon Bell らが 2002 年に初めて発表した MyLifeBits プロジェクトがある [2]。MyLifeBits は、ユーザーにカメラや音声レコーダー、各種センサーを取り付け、生活のすべてを記録するというプロジェクトである。具体的には、電子メール、閲覧した Web ページやスキャンした紙書類、本、写真、音声記録、音楽 CD などが対象であり、実際に記録したデータから試算すると、それらを記録するのに必要な記録媒体の容量は、一人当たり 5 年間で約 80GB になると言われている。記録媒体が大容量化しており、一生分の記録を残すことも可能である。また、相澤らによる FoodLog という特定の行動に限定したライフログのシステムが提案されている [3]。FoodLog は、日々の食事を画像として解析することにより、ユーザーの食生活を改善する手助けとなることを目標としている。2010 年 4 月にプレリリースされており、Web¹にアクセスすることにより、誰でも無料で利用できる。

ライフログという言葉は広く浸透しており、ブログのような日記などもライフログとして扱う人がいる。ライフログサービスとして Life-X²などが挙げられるが、Evernote³や Twitter⁴といった、日記やつぶやきを記録するサービスもライフログを記録することを目的として扱われることがある。また、携帯電話の高機能化やスマートフォンの登場により、ライフログを利用する機会は増えてきている。自身の活動を記録するという点で、常に持ち歩くような携帯電話やスマートフォンはライフログに適したデバイスであると言える。ライフログにより、人は自身の生活や活動の記録を振り返ることが可能となり、その記録を利用することで、今

¹FoodLog : <http://www.foodlog.jp/>

²Life-X : <http://life-x.jp/>

³Evernote : <http://www.evernote.com/>

⁴Twitter : <http://twitter.com/>

後の生活や活動をより良いものにすることができる。

我々は、MyLifeBits よりも狭く、FoodLog よりも広いライフログのドメインとして、研究生活に着目している。我々研究者は、様々な研究活動の記録をログとして残すことがある。たとえば、備忘録のような簡単なメモや長時間に及ぶミーティングの映像や議事録などである。研究者はこれらの記録を利用することで、研究活動をより充実したものに行っている。

研究者が記録を残す方法として、テキスト、音声、ビデオなどの手段が考えられるが、その中でも一般的に利用されるのはテキストである。テキストによる記録の場合、非常に簡単に利用することができ、検索も容易であるが、記録にかかる手間は少なくない。また、会話中のメモやふと思いついたアイデアなどを記録する際、漏らさず書き残すことは困難である上、要点だけを記録することに集中した結果、重要なことを聞き漏らしたり、思いついたことを部分的に忘れてしまうといった危険性もある。ビデオによる記録では、デモの様子や議論風景など、言葉で伝えるよりも多くの情報を同時に伝えることができるが、長時間にわたるビデオの視聴やデータ処理(ファイル転送や編集を含む)には時間がかかるため、目的の情報にすばやくアクセスするためには、内容に基づく構造化や適切なインデックスの付与などの工夫が欠かせない。また、「今日サーベイを行った」「明日はシステムの実装を行う」などの簡単なメモを記録するためにビデオを利用するのは少し冗長と思われるし、ビデオカメラによる撮影行為を日常的に行うことについて、撮影する/されることに抵抗感を感じる人は依然として多い。

一方、音声による記録の場合、テキストによる記録のような手間がかからず、ビデオよりも小さなデータサイズで長時間の記録が可能という特長がある。音声時間が時間情報とともに記録されたものを、本研究では音声ログと呼ぶ。2012年現在、民生用のビデオカメラ(内蔵96GBメモリ)では最大約40時間(LPモード時)、スマートフォン(内蔵64GBメモリ)では約6時間(H.264形式)のビデオ撮影が可能であるのに対し、ICレコーダー(内蔵8GBメモリ)では2000時間(WMA形式LPモード)を超える録音が可能であり、音声ログの記録の敷居は非常に低いと考えられる。しかし、長時間の記録が可能であるがゆえに、記録した音声ログをこまめに整理しておかないと、膨大な数の音声ログが溜まった際に、必要な情報がどの音声ログに含まれているかを短時間で判断できないなど、利用が非常に難しいという問題がある。

大量の音声ログが記録された場合に、音声ログの利用を容易にするための従来方法として、音声ログに対してタイトルやタグなどの補足情報をユーザーが付与する方法がある。しかし、ユーザーが手動でタイトルやタグを入力する場合、その手間は非常に多く、記録が容易であるという音声ログの利点を損なっている。また、音声認識技術を用いて音声ログの内容を文字化する方法があるが、「～の実装において」「～の目的で」などの重要な情報が、記録時のユーザーにとっては当然の条件として省略されて発話された場合、音声認識により正確に文字化

されたとしても、必ずしも十分な整理を行うことはできない。

そこで本研究では、音声ログの利点である記録の容易さを損なわずに音声ログを整理することを目的とし、研究活動における音声ログの整理を容易かつ有効に行うためのシステムを提案する。研究者は、一般的にテキストによる研究活動の記録を行っているが、常にテキストを残せる状況にあるとは限らず、そのような場合には、音声ログを記録することが望ましい。しかし、単に音声ログを記録するだけでは利用が困難になることから、テキストによる研究活動の記録(以降、研究ノートと呼ぶ)を利用する。すなわち、音声ログに付与するタイトルとタグを研究ノートから抽出するシステムを提案する。

具体的には、まず、ユーザーが日常的に作成・更新する研究ノートの情報を Web 上のノートサーバーで管理する。ノートに変更が加わるたびに、サーバーではノートの本文に形態素解析を施して名詞からなるタグを抽出する。次に、iRec と呼ばれるスマートフォン向けの録音用アプリケーションを開発し、それを用いてユーザーが録音を行う。すると、iRec が事前にノートサーバーからダウンロードしておいた、ユーザーが普段書きためた研究ノートのタイトル一覧が表示される。記録した音声ログと関連が深いと思われるタイトルをユーザーが選択すると、当該ノートから抽出されたタグと、ユーザーが任意に作成したタグの一覧が提示され、その中から任意のタグを複数個選択することにより、タイトルとタグが音声ログに付与される。記録された音声ログはサーバーで蓄積・管理され、Web ブラウザからアクセスすることによって利用・管理できる。

提案システムでは、音声ログのタイトルとして、研究ノートのタイトルのいずれか一つが付与される。音声ログのタイトルとして適切かどうかはユーザーが研究ノートにどのようなタイトルを付与しているかによるが、少なくとも、音声記録時に手動でタイトルを入力する場合と比較して、自由度の低いタイトルになると予想される。そのため、研究ノートから抽出される多様なタグを簡便に付与する仕組みをユーザーに提供することによって、音声ログの整理を目指している。

提案システムを評価するために、2種類の被験者実験を行った。1つ目は、タイトルとタグの入力コストを評価するための実験である。被験者には、一定期間、提案システムを利用する場合とタイトルとタグを手動入力する場合の2通りの方法によって音声記録を行ってもらい、タイトルとタグの付与に要した時間を測定した。2つ目は、提案システムによる音声ログの整理の有効性を評価するための実験である。被験者に音声ログの再生を必要とする問題を出題し、問題に回答する時間と回答中に再生した音声ログの数を測定した。被験者を2グループに分け、一方は提案システムにより記録された音声ログを、もう一方はタイトルとタグを手動入力して記録された音声ログを対象にして問題に回答してもらうことで、提案システムによるタイトルとタグの付与の有効性を評価した。2つの実験により、提案システムを利用した方が、タイトルとタグの入力時間は短くなり、問題の回答時間は長くなることが分かった。

本論文の構成を以下に示す．まず，第2章で音声ログの整理に関する従来手法について述べ，その問題点と本研究における解決法について述べる．次に第3章で音声ログを記録・整理する具体的なシステムについて述べる．第4章では提案システムの有効性を記録時のコストと，整理されているかどうかの2点から検証する実験と考察について述べる．第5章では，提案システムを利用して記録された音声ログを有効に利用するための方法について述べる．第6章では関連研究やアプリケーションを紹介し，第7章でまとめと今後の課題について述べる．

第2章 研究活動における音声ログ

2.1 音声ログ

研究者は研究活動において、備忘録、ミーティングの記録、活動の記録、活動の成果など、様々な記録をとる。そしてそれらの記録を見返して重要なことを思い出したり、活動の変遷を確認したりし、自身の活動にフィードバックして、研究活動をより充実したものに行っている。

一般的に利用される記録方法はテキストによる記録である。テキストは短時間で内容が把握できることから、利用は非常に容易である。しかし、記録する手間は比較的多く、時間のない時や手の離せない状況では記録が難しいという問題がある。そのような場合には、より記録に手間のかからない音声による記録が有効である。しかし、音声ログを利用する、すなわち音声による記録から必要な情報を見つけるためには、膨大な数の音声の中から必要なものを探し出す必要があり、一覧を見返すだけでは適切な情報にアクセスすることができないといった問題点がある。

そこで本研究では、音声による記録を音声ログと定義し、音声ログの利用を容易にするための方法を提案する。研究活動において記録される音声ログとして、実際に記録された例を図 2.1 に示す。

このように記録された音声ログの利用を容易にするためには、音声ログに対して何らかの補足情報を付与し整理することが有効である。音声ログに自動で付与される補足情報には、録音日時、録音時間などがある。これらは、2～3 日前などの最近記録された音声ログを利用する上では重要な情報となるが、より過去に記録された音声ログを利用する場合にはあまり有効ではないため、音声ログのタイトルやタグなど、より細密な情報をユーザーが手動で付与する必要があると考えられる。また、音声ログの利点は記録が容易であることから、ユーザーが手動で情報を付与する手間はできる限り少ないことが望ましい。

また、音声ログの利用が容易であるためには、ユーザーが音声ログを利用する状況に合わせて音声ログが整理されていることが望ましい。例えば、図 2.1 で示した例にある「今週中に、音声再生機能を実装する」という音声ログの場合について説明する。この場合の実装とは、PC でプログラミングをして機能を作成する行為であることから、ユーザーが PC を利用している状況において思い出したい内容であることが分かる。つまり、数多くある音声ログの中で、この音声ログに対

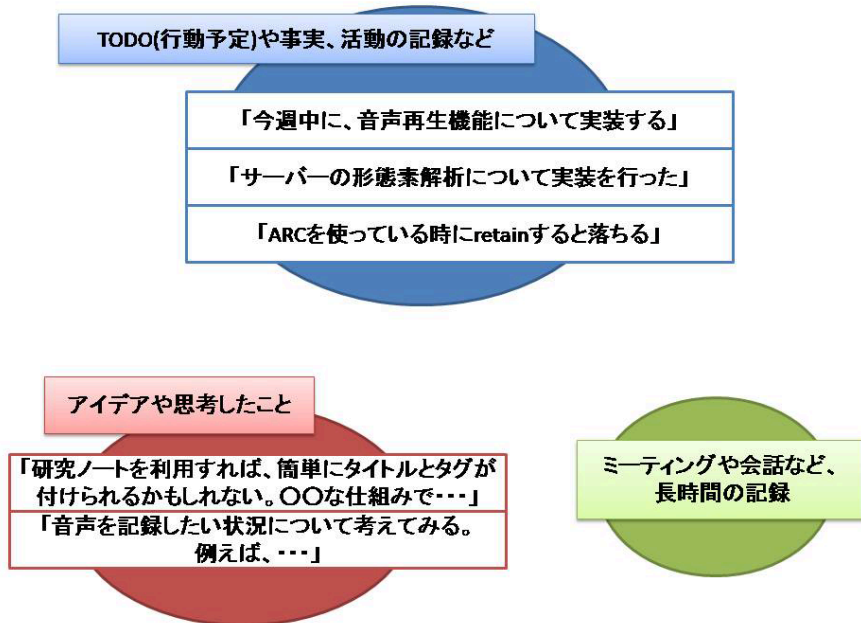


図 2.1: 音声ログの記録例

して「実装」というテキストが情報として付与されていれば、ユーザーは音声ログを容易に利用できる。また、ユーザーが何か忘れていないか、今日やるべきことは何か、といった動機から音声ログを探す場合、今後やるべきことを記録した音声ログなのか、やり終えたことを記録した音声ログなのかを判断するための情報が重要である。例えば、「今日、サーバーの形態素解析について実装を行った」という音声ログは、やり終えたことを記録したものであるが、「サーバー」「形態素解析」「実装」などの情報だけが付与されていても、その音声ログを再生すべきかどうかの判断は困難である。行動予定に対しては「TODO」などのテキストが情報として付与されていれば、ユーザーにとって音声を再生すべきかどうかの判断が容易になると考えられる。そこで本研究では、ユーザーが音声ログに対して利用する状況に適したタイトルとタグを付与することで音声ログを整理するとともに、その整理にかかる手間が極力少なくなるような方法を提案する。

2.2 音声ログの整理に関する従来方法

従来の音声ログの整理方法として、ユーザーがタイトルやタグを音声ログに付与したり、音声認識によって音声ログを文字化するなどの方法が挙げられる。本節では、これらの従来方法について述べ、その問題点を提示する。そして、本研



図 2.2: ボイスメモのラベル一覧 (左) とカスタム入力 (右)

究で提案する，研究ノートを利用した音声ログの整理方法について説明する．

2.2.1 タイトルやタグの付与

一般的な録音機器として，IC レコーダーが挙げられる．IC レコーダーは音声録音するための機器であり，主に会議やインタビューなどにおける音声の記録に利用される．またスマートフォンに録音機能がある場合も多く，録音するためのアプリケーションも数多く存在する．ほとんどの IC レコーダーや録音アプリにおいて，ユーザーは音声ログに対してタイトルやタグなどを付与し整理することができる．例えば，Olympus 社から発売されている VoiceTrek DM-4¹ や，Sony 社から発売されている ICD-BX312² では，ファイル名を編集したり，ファイルを管理するフォルダの作成などができる．iPhone の標準アプリケーションであるボイスメモ³ では，図 2.2 に示すようにラベルという呼称で，予め用意された数種類から選択するか，あるいはユーザーが音声ログのタイトルを任意のテキストで入力できる．

しかし，これらの方法では，タイトルやタグは何種類かの用意されたものの中から選択するか，手動で入力する必要がある．細密な整理を行うために手動で入力を行う場合，ユーザーにかかる手間は多くなってしまう．また，録音時間の短い音声ログを記録した場合，ほぼ全ての内容に相当するタイトルやタグを手動で

¹Olympus VoiceTrek DM-4 : <http://olympus-imaging.jp/product/audio/dm4/>

²SONY ICD-BX312 : <http://www.sony.jp/ic-recorder/products/ICD-BX312/>

³ボイスメモ : <http://www.apple.com/jp/iphone/built-in-apps/more-built-in-apps.html>

入力してしまうと、テキストによる記録を行う場合と比べて記録時にかかる手間はほとんど変わらないか、むしろ多くなってしまう。

また、これらの方法では、音声記録時ではなく後から、既に保存されている音声ログに対して任意のタイトルやタグを付与する場合はほとんどであり、記録時にタイトルやタグを付与するための手間がかからないようになっている。ユーザーは自分にとって重要な音声ログや、後から利用したいと考えるような音声ログに対してタイトルやタグを付与することにより、ユーザーにとって必要な音声ログのみを整理することが可能である。しかし、録音直後にその音声ログにタイトルやタグを付与することに比べて、長い時間が経過してから録音日時のみを手掛かりに音声ログの中から目的の音声ログを探すことは困難である。特に日常的に大量の音声ログを記録する状況では、効率的とはいえない。

2.2.2 音声認識

音声認識を利用した整理方法もある。例えば、音素表記によるマッチングを行う検索方法 [4] や、認識誤りに強い検索方法 [5] がある。また近年では、複数人が音声共有し認識誤りを修正し合うことで、正確な書き起こしを行い、さらに音声認識精度を高める方法 [6] などもある。これらの方法により、音声認識の精度はどんどん良くなってきており、今後も音声ログの文字化への利用は増えていくと考えられる。

しかし、仮に音声認識により音声ログが正確に文字化されたとしても、必ずしも音声ログの整理に役立つわけではない。例えば、図 2.1 において例に出した「今週中に、音声再生機能を実装する」という音声ログの場合、一体何の音声再生機能を実装するのかが分からず、音声再生を行うアプリケーションを複数同時に開発している場合には、どのアプリケーションにおいて機能を実装すればよいのか判断できない。ユーザーは音声録音時には当然どのアプリケーションにおいて実装すればよいかを分かっており、だからこそ省略して話して録音してしまったのだと考えられるが、このような音声ログを整理するためには、補足情報としてアプリケーション名などが必要となる。そのため音声認識によって音声ログが文字化されただけでは、このような音声ログを整理することは難しいと言える。

2.3 研究ノートを利用した整理手法

研究者は一般的に、デジタルかアナログかに関わらず、テキストによる記録を日常的に行っている。本研究では、研究者が作成する研究活動に関するテキストの記録を研究ノートと定義する。

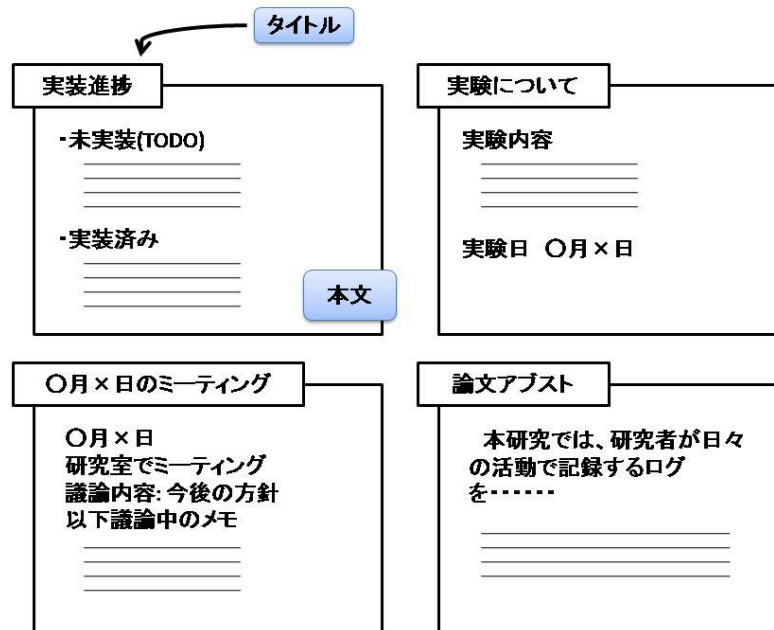


図 2.3: 研究ノートの例

研究ノートに記録される内容は、TODO をリスト化した記録、新たな知見や事実の記録、活動の記録、あるいは議論で決定したことのまとめなど、主に、研究活動を支援・補助するものであり、研究者は研究ノートを見返すことで次の活動にフィードバックしている。すなわち、研究者は、研究ノートへの記録と研究ノートの利用を繰り返すことによって、研究活動をより充実したものに行っている。本研究で想定している研究ノートの例を図 2.3 に示す。

研究活動における音声ログは、研究ノートに記録できないような場面において多く記録され、その内容は研究ノートと同様の、TODO や簡単なメモといった研究活動を支援・補助するものである。この性質から、音声ログに付与するタイトルとタグとして適切なものが、研究ノートに含まれている可能性が高いと考えた。

さらに、研究活動は、断続的な話題ではなく、1つのある大きな研究テーマに沿って行われる活動である点が特徴と言える。つまり、断続的な記録を行っても、記録内容のテーマは基本的に同じであり、内容が大きく変わることは稀である。また、研究ノートは図 2.3 のように、研究におけるいくつかのサブテーマごとに作成されるのが一般的である。この特徴から、音声ログに付与するタイトルとタグとして適切なものが研究ノートにまったく存在しないという状況は考えにくく、特に、直近に記録された研究ノートほどより関係性の高い話題である可能性が高いと考えた。

そこで本研究では、この研究ノートに着目する。すなわち、研究ノートを利用

することにより、音声ログに対して簡単にタイトルとタグを付与し整理するという方法である。具体的には、音声ログのタイトルとして研究ノートのタイトルを利用し、さらに音声ログに付与するタグをその研究ノートから抽出して付与する仕組みを実現した。ただし、この仕組みを利用することによって、音声ログに付与されるタイトルは、手動で入力する場合に比べて自由度の低いタイトルとなり、細密な整理はより困難であると考えられる。しかし、豊富なタグを選択することにより、多様なタグによる整理が実現されることを期待している。

第3章 音声ログの記録・整理手法

本章では、研究ノートを利用して音声ログを整理するシステムについて説明する。

3.1 システムの概要

提案システムの構成を図3.1に示す。提案システムは3つの要素から構成される。ノートサーバー、録音アプリ iRec、ボイスサーバーである。

ノートサーバーは、研究者が日頃から作成している研究ノートのデータを保管するためのサーバーである。研究ノートを作成するシステムは汎用的なもので構わない。ただし、ノートの情報をノートサーバーに保管することが重要である。ノートサーバーは、研究ノートを保存する際、そのノートからタグを抽出し、ノートのメタデータとしてデータベースに保存する。

録音アプリ iRec は、iPhone/iPod touch, iPad 用に開発した音声記録を作成するためのアプリケーションである。iRec は音声の録音と再生が可能だけでなく、予めノートサーバーから、ノートタイトル及びその研究ノートから抽出されたタグをダウンロードし、録音時に音声に対するタイトルとタグとして付与することができる。また、iRec で作成した音声ログは、ボイスサーバーにアップロードすることができる。

ボイスサーバーは、音声ログを保管しておくためのサーバーである。iRec から音声ログとしてアップロードされた音声ファイルと、そのタイトル、タグ、録音日時および録音時間などのメタデータを保存し、管理する。

3.2 研究ノートからのタグ抽出

3.2.1 研究ノート作成システム

提案システムに必要な研究ノートを構成する要素は、タイトル、本文、作成日時、更新日時である。ノートサーバーに最新のノート情報を保管するために、これらの要素をデータ構造として保持し、ノートサーバーとデータを送受信可能な研究ノート作成システムが必要である。本研究では、研究ノート作成システムとして iSticky[7] と、ブラウザでのノート編集機能を利用する。

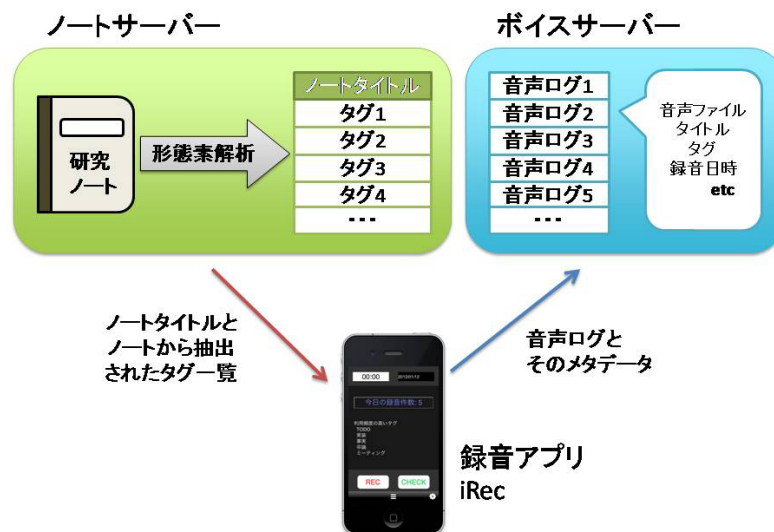


図 3.1: システムの構成

iSticky とは、著者らの研究グループで開発した、ミーティングの支援を目的とした iPad 用アプリケーションである。iSticky には、主な機能として、研究ノート執筆、画像管理、スケッチ描画・管理など、研究活動におけるコンテンツの蓄積・管理を行う仕組みが備わっている。また、それらのコンテンツを利用したミーティングを行う機能を持っている。さらに、ミーティング内容を保存することで、過去のミーティングを振り返ることができる仕組みが実現されている。提案システムでは、iSticky の研究ノート執筆機能のみを利用する。iSticky における研究ノートの閲覧インターフェースを図 3.2 に示す。

図 3.2 のインターフェースの左に表示されているノートリストには、今までに書き溜めてきたノートのタイトルが一覧となって表示されており、タップして選択することで、その本文が右側のテキストエリアに表示される。なお、タップとは iPad などのタッチパネル端末における、タッチパネルを指 1 本で 1 回叩く動作をいう。このテキスト内容は行単位で管理されており、行ごとに選択してコピーやカットなどの編集を行うことができる。図 3.3 にその様子を示す。

テキスト編集ボタンをタップすると、ノートを執筆する画面に遷移し、ノートの本文を編集することができる。このとき、一番上の行がノートのタイトルとなる。また、作成されたノートの行数や作成時間などの詳細を閲覧することもできる。ノートタイトルをロングタップすると図 3.4 に示すようなメニュー画面が表示され、ノートの削除などの編集ができる。なお、ロングタップとはタッチパネルを指 1 本で長押しする動作をいう。



図 3.2: iSticky の研究ノート閲覧インターフェース

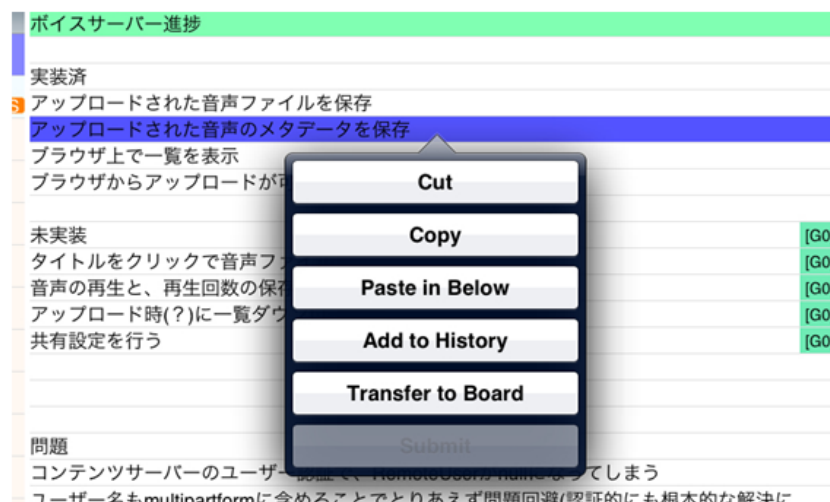


図 3.3: iSticky における行ごとの編集



図 3.4: iSticky のノート詳細およびメニュー画面

図 3.2 左下部に配置された Update ボタンをタップした後にノートを選択すると、ノートサーバーに保管されている情報と同期することができる。ノートは ID で管理されており、既にノートサーバーに同一の ID を持つノートが保管されている場合、更新日時の新しい方を優先して保管し、古い方の iSticky またはノートサーバーの情報を更新する。iSticky からノートサーバーにアップロードされる情報は、タイトル、本文、作成日時、更新日時、その他メタデータである。

また、ノートはブラウザからも閲覧編集することができる。ノートサーバーに保管されているノートの情報は、図 3.5 のように表示される。ノートタイトルをクリックすると、本文が書かれたページに遷移する。編集ボタンをクリックすると、タイトルと本文を編集するページに遷移し、内容を更新することができる。

3.2.2 タグの抽出方法

研究ノートが新たに作成されたり、更新された際に、ノートサーバーはその研究ノートの本文に形態素解析を施すことにより特定の品詞をタグとして抽出し、データベースに保存する。

提案システムでは、形態素解析器として lucene-gosen¹ を利用する。lucene-gosen とは、Java で利用可能な日本語用の形態素解析ライブラリであり、jar ファイルに辞書を内包している。外部の辞書を指定したり、複数の辞書を利用することもで

¹lucene-gosen : <http://code.google.com/p/lucene-gosen/>

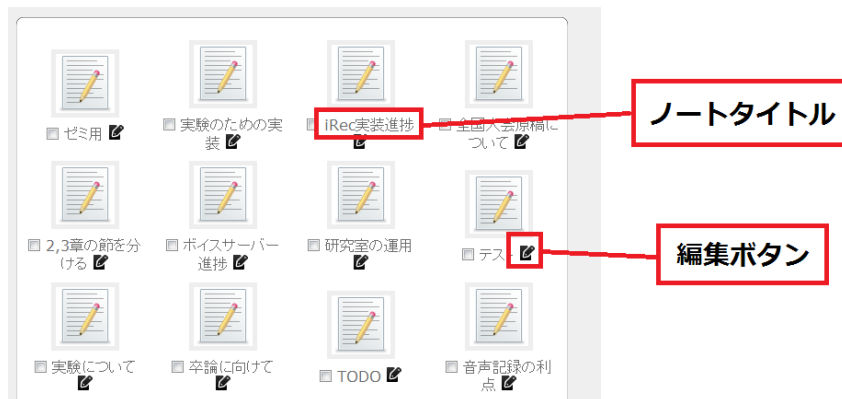


図 3.5: ノートサーバーのノート一覧

きるが、提案システムでは新たに辞書を作成せずに、jar ファイル内包の辞書のみを利用する。lucene-gosen を利用する理由は、導入、運用が容易であり、また、現在も開発が行われて頻繁に更新されているためである。形態素解析により抽出するタグは、名詞-一般、名詞-サ変接続、名詞-未知語の 3 種類であり、連続してこれらが出現する場合は一つの複合名詞として連結して扱う。また、形態素解析を施す際に同じタグが抽出された場合、その数をカウントしてデータベースに保存する。抽出されたタグとそのカウントは、研究ノートごとに分けて保存される。

3.3 録音アプリ iRec

3.3.1 音声ログの種類

研究活動において記録される音声ログには、簡単なメモから長時間に及ぶ議論まで、様々な内容がある。それらが分類されていれば、音声ログをより簡単に利用できると考えられる。研究室の大学生 6 名に一定期間自由に音声ログを記録してもらったところ、ミーティングなどの長時間の音声ログが記録される数に比べて、簡単なメモなどの短時間の音声ログが記録される数の方が明らかに多いことが分かった。本研究では、膨大な音声ログの中から目的の音声ログを探すための整理を目的としているため、短時間の音声ログが整理されていれば記録された音声ログ全体として十分な整理ができていると考えられる。そこで、記録された短時間の音声ログの内容を分析し、以下の 3 つのカテゴリーに分類した。

1. アイデア思考

新しいアイデアや、頭の中で整理できていない事柄などを口に出して言語化することで、メモをしたり考えをまとめたりするために記録した音声ログ



図 3.6: iRec の録音画面

2. 事実

行動したことや、調査などにより新たに分かったことなどを記録した音声ログ

3. TODO

自身が今後やらなければならないことなどの備忘録を記録した音声ログ

3.3.2 iRec のシステム

音声ログを整理する手法を実現した録音アプリケーション iRec について説明する。iRec は、iPhone/iPod touch 用に開発した録音アプリである。iRec の録音画面を図 3.6 に示す。

iRec は、起動時にノートサーバーに接続し、更新されたノートの情報を取得する。取得する情報は、ノートのタイトルと、サーバーが予め抽出しておいた音声ログ用のタグ、及びノートのメタデータである。iRec で録音ボタンをタップすると、録音が始まる。再び録音ボタンをタップすると録音を終了することができる。録音中は、録音時間が赤色で表示される。また、CHECK ボタンは本研究では利用しないが、録音中に CHECK ボタンをタップすることで、その時間をマーキングすることができる。この機能は主に長時間の記録において利用されるものである。具体的には、長時間の記録の中で重要だと思われる箇所や、後から聞き



図 3.7: タイトルの選択画面

返そうと思った箇所で、このボタンをタップすることで、その時間情報が音声ログに付与され、再生時にその時間に容易にアクセスすることが可能になる。録音・再生切り替えボタンをタップすることで、録音画面と再生画面を切り替えることができる。設定ボタンをタップすると、設定画面が表示される。設定画面の詳細については後述する。ユーザーが録音を終了すると、図 3.7 に示すように研究ノートタイトルの一覧が更新日時順に表示される。

一覧の中からノートのタイトルを一つ選択すると、そのタイトルが記録した音声ログのタイトルとして保存される。なお、iRec 内にノート情報が一切存在しない場合は、タイトルの入力画面は表示されず、録音日時がタイトルとして保存される。タイトルを選択すると、図 3.8 に示すようなタグ一覧が表示される。

タグ一覧の一番上には、図 3.9 に示すようなカテゴリーの選択インタフェースが表示される。前述した通り、短時間の音声ログである、アイデア思考、事実、TODO の 3 つのカテゴリーが並んで表示されており、3 つのうちいずれか 1 つを選択し、タグとして付与することができる。

タグ一覧には、推薦タグとユーザータグの 2 種類がある。推薦タグは、選択したノートの本文から抽出されたタグの一覧であり、形態素解析によってノートから抽出された数が多いほど重要なタグであると仮定し、抽出数の多いタグから順に表示される。ユーザータグは、ユーザーが任意に作成して追加したタグ一覧である。タグ作成ボタンをタップするとタグの入力画面が表示され、新たに必要なタグを作成することができる。作成したタグは、以後ユーザータグとして、音声記録時に一覧に表示され、簡単に付与できるようになる。iRec 内にノート情報が



図 3.8: タグの選択画面



図 3.9: カテゴリーの選択



図 3.10: iRec の再生画面

一切存在せずにノートタイトルを選択できなかった場合は、ユーザータグの一覧のみが表示される。

推薦タグ及びユーザータグは複数個選択でき、一覧にあるタグに対してタップを繰り返すと、付与するかどうかを切り替えることができる。選択中のタグは、一覧のタグの横にチェックマークが表示され、画面下部の選択中のタグ一覧に表示される。また、タグの選択時には、タイトル選択画面に戻り、ノートのタイトルを選び直すことができる。

タグの選択後、保存ボタンをタップすると、選択中のタグが、カテゴリーを含めて、全て音声ログに付与される。この操作により、録音の一連の作業は終了する。

iRec の再生画面を図 3.10 に、チェックマーク機能についてのインタフェースを図 3.11 に示す。

再生画面には記録された音声ログ一覧が表示される。表示される情報は、音声ログのタイトルと付与されたタグおよび、音声ログの長さである。一覧の中から任意の音声をタップすることによりその音声ログを選択することができる。再生ボタンをタップすると、選択中の音声再生される。また再生位置を示すシークバーを操作することで、任意の時間から再生することができる。

チェックボタンをタップすると、再生中の音声ログの現在の再生時間に対してマーキングすることができる。これは、音声記録時に CHECK ボタンを押すのと同等の操作を再生中に行う機能であるが、本研究においては使用していない。シー

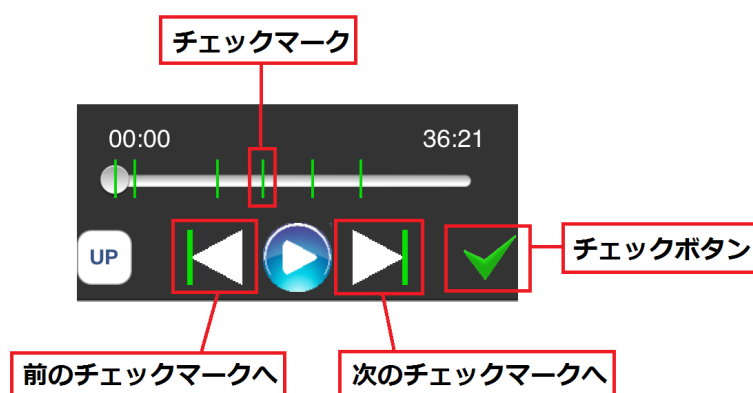


図 3.11: チェックマーク機能についてのインターフェース

クバー上に表示されているチェックマークは、記録時、または再生時にマーキングされた時間を表している。「次のチェックマークへ」、または「前のチェックマークへ」ボタンをタップすることで、現在再生中の時間から前後のチェックされた時間へ容易にアクセスすることができる。また、チェックマークをロングタップすることでそのチェックマークを削除することができる。

アップロードボタンをタップすると、選択中の音声ログをボイスサーバーにアップロードすることができる。アップロードされる情報は、音声ファイル、その音声ログのタイトルとタグ、および録音日時などのメタデータである。音声ログを選択中に詳細ボタンをタップすると、その音声ログの詳細なメタデータを確認することができる。音声ログの詳細画面の例を図 3.12 に示す。この画面では、録音日時やファイルサイズを確認するだけでなく、音声ログのタイトルを編集することもできる。

録音画面と再生画面に共通する設定ボタンをタップすると、図 3.13 に示すような設定画面が表示される。設定画面では、ノートサーバーの URL とボイスサーバーの URL、またそれらのサーバーにアクセスするためのユーザー情報を入力する。ノートの情報をダウンロードしたり、音声ログをアップロードする際には、ここで設定された各サーバーに接続して行われる。

ボイスサーバーは、アップロードされた音声ログを蓄積・管理する。タイトルやタグ、メタデータだけでなく、選択された研究ノートの情報も同時に保存することで、今後様々な応用が可能であると考えられる。



図 3.12: 音声ログの詳細画面



図 3.13: iRec の設定画面

3.3.3 タイトルの手動入力

iRec は、一般的な録音アプリと同様に、手動によるタイトル入力機能も備えている。図 3.14 に手動入力画面を示す。また、タイトルを手動入力する場合、タグ一覧にはユーザータグのみが表示される。

3.4 音声ログの閲覧インターフェース

音声ログの再生は、Web ブラウザで行うこととした。Web ブラウザで音声ログを閲覧するインターフェースを図 3.15 に示す。

一番上に検索フォームが表示され、その下に音声ログ一覧が表示される。検索フォームでは、カテゴリーの絞り込み機能と、キーワード入力によるハイライト検索機能を提供している。前者では、アイデア思考、事実、TODO のカテゴリーのいずれかをチェックし、絞り込みボタンをクリックすることで、一覧を絞り込むことができる。後者では、フォーム内に入力したキーワードと一致するタイトルとタグの部分で黄色くハイライト表示する。また、空白を入れることで、複数キーワードの指定ができる。これらの検索を行った結果を図 3.16 に示す。

各音声ログには、図 3.17 に示すように、タイトル、タグ、録音日時、録音時間、および再生ボタン、停止ボタン、音量を変更するボタンが表示される。タイトル



図 3.14: タイトルの手動入力画面



図 3.15: 音声ログの閲覧インターフェース



図 3.16: カテゴリーの絞り込みとハイライト検索を行った結果

タイトル : 実験について
 タグ : 事実, タグ, 実験, 実装, 研究, 準備,
 録音日時 : 2012-01-18 03:01:32
 再生時間 : 0時間0分22秒
 再生 停止 音量アップ 音量ダウン

図 3.17: 音声ログの表示

およびタグが表示されない設定の問題では、図 3.18 のように表示される。この場合、音声ログを探す手掛かりとなる情報は録音日時と録音時間のみとなり、音声ログの検索はできない。

3.5 システムの利用

ここまで、提案システムによる音声ログの記録およびタイトルとタグの付与の仕組みについて述べた。本節では、研究者が提案システムを実際に利用する流れを、具体的な例を用いて説明する。

提案システムによる音声記録、およびタイトルとタグの入力の流れを図 3.19 に示す。まず、研究者が記録したい内容(ここでは「今週末に被験者にアンケートをとる」)を iRec に向かって発話する。次に、iRec の画面に表示された研究ノート



図 3.18: 音声ログの閲覧インターフェース (タイトルとタグの非表示)

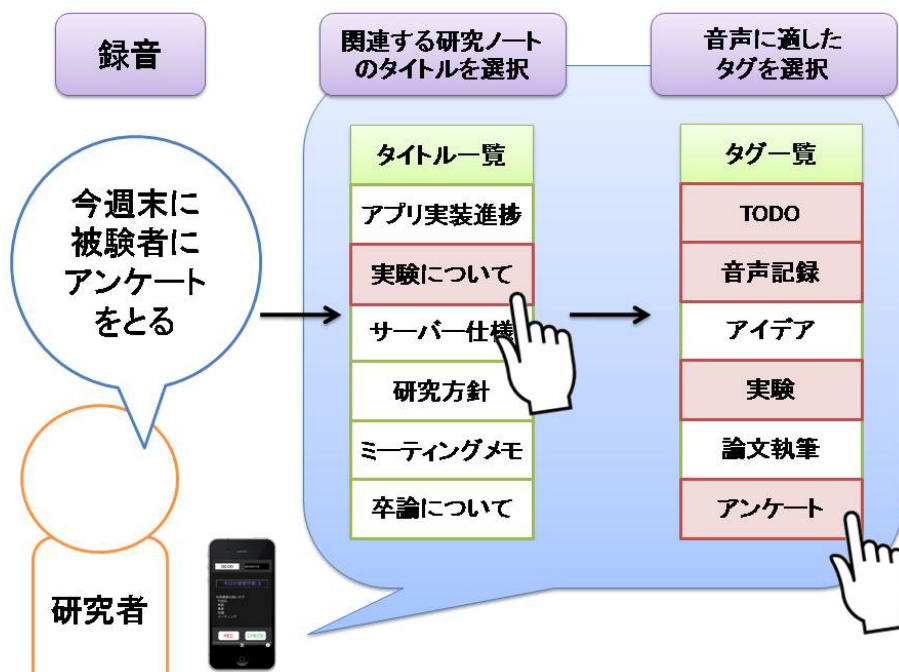


図 3.19: タイトルとタグの入力例

のタイトル一覧から、発話内容に最も関連する「実験について」を選択する。最後に、タグの選択画面上で、発話内容に関係のある「音声記録」「実験」「アンケート」を選択した後、カテゴリーとして「TODO」を選択し、音声ログの記録操作を終了する。

以上のように、研究ノートと当該ノートから抽出されたタグを選択することで、容易に音声ログを記録し、整理に必要な情報を付与することができる。また、この例では、「音声記録」や「実験」など、発話内容そのものには含まれない情報をタグとして付与しており、音声認識による音声ログの文字化のみでは困難なタグ付けを行っている。

第4章 評価実験

提案システムを利用した音声ログに対するタイトルとタグ(メタデータと呼ぶ)の付与を効率的に行えること、およびそこで付与されたメタデータが音声ログの整理に有用であることを評価するために、2種類の被験者実験を行った。それぞれの実験の目的、方法、結果、考察について述べ、最後に両実験のまとめを述べる。音声ログ記録時のメタデータ入力に関する実験と、音声ログの検索に関する実験である。

4.1 音声ログ記録時のメタデータ入力に関する実験

4.1.1 実験の目的と方法

本実験の目的は、音声ログを効率的に整理する上で提案システムの利用が効果的であるかどうかを評価することである。具体的には、音声ログのタイトルとタグの付与に要した時間を計測する。被験者は研究室に所属する大学生6名であり、実験期間は約1カ月間とした。比較システムとして、被験者がタイトルを手動で入力し、さらに任意で作成したユーザータグをもとにタグを付与する方法を設定した。なお、タイトルを手動で入力するシステムでは、1文字以上入力されていない場合は保存できないようにした。タイトルとタグの入力時間に関する個人差を吸収するために、被験者には、それぞれの方法を一日ごとに交互に利用してもらった。被験者6名を3名ずつの2グループに分け、それぞれグループI(被験者A, B, C)、グループII(被験者D, E, F)とした。グループIは提案システムを使用する方法から、グループIIは手動入力を使用する方法から開始した。

被験者は、1日9件以上音声ログを記録することとした。その内訳としてアイデア思考を1件以上、事実を4件以上、TODOを4件以上とした。音声ログ記録時には必ずこれらのカテゴリーのうちどれか1つを選択しなければ、保存できないようにした。実験日数は実験期間中の16日間であり、一人当たりの音声ログ記録数は最低で144件、最高で148件であった。被験者は各々が所有するiPhone, iPod touchのいずれかにiRecをインストールし、音声ログの記録を行った。

なお、実験後に口頭のアンケート調査を行った。提案システムを利用した場合と手動入力の場合のどちらが容易に記録を行えたかについて質問し、その回答および自由意見を録音した。

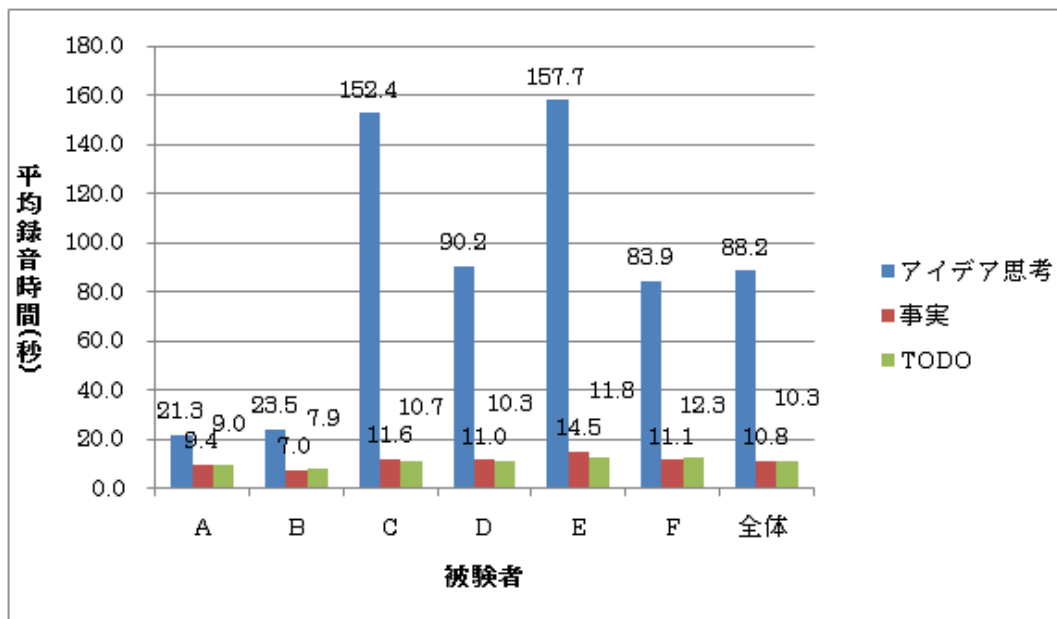


図 4.1: 音声ログのカテゴリー別平均録音時間

4.1.2 実験結果

カテゴリー別の音声ログの平均録音時間を被験者ごとにグラフにしたものを図 4.1 に示す。図 4.1 より、アイデア思考に関する音声ログの録音時間は、被験者による差が大きいことが分かる。

次に、タイトルの平均入力時間を図 4.2 に、タグの平均入力時間を図 4.3 に示す。また、それらの合計時間を図 4.4 に示す。タイトルの入力時間は、提案システムを利用したときの平均入力時間は 6.1 秒であり、手動入力 of 33.1% であった。一方、タグの入力時間は、提案システムと手動入力に大きな差はなかった。タイトルとタグの入力時間を合計すると、いずれの被験者においても提案システムを利用した場合の方が入力時間は短かった。被験者全体では、提案システムの方が手動入力 of 60.8% の時間で入力を終えている。

4.1.3 考察

本項では、タイトルとタグの入力時間、入力タイトルの多様性、タグの抽出効果と多様性、およびアンケートから得られたその他の意見について考察する。

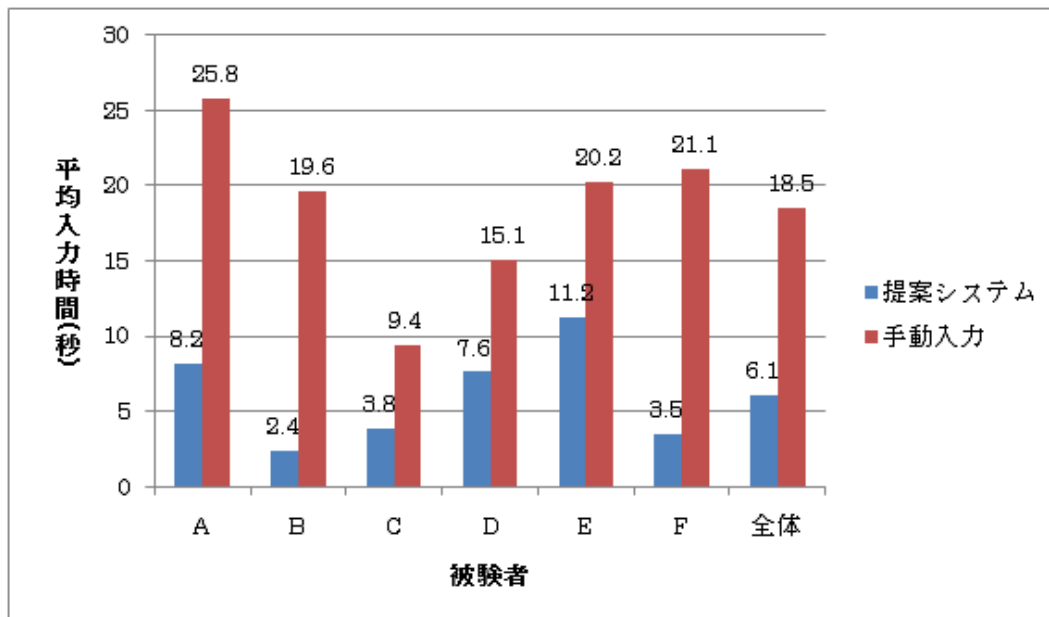


図 4.2: 平均入力時間 (タイトル)

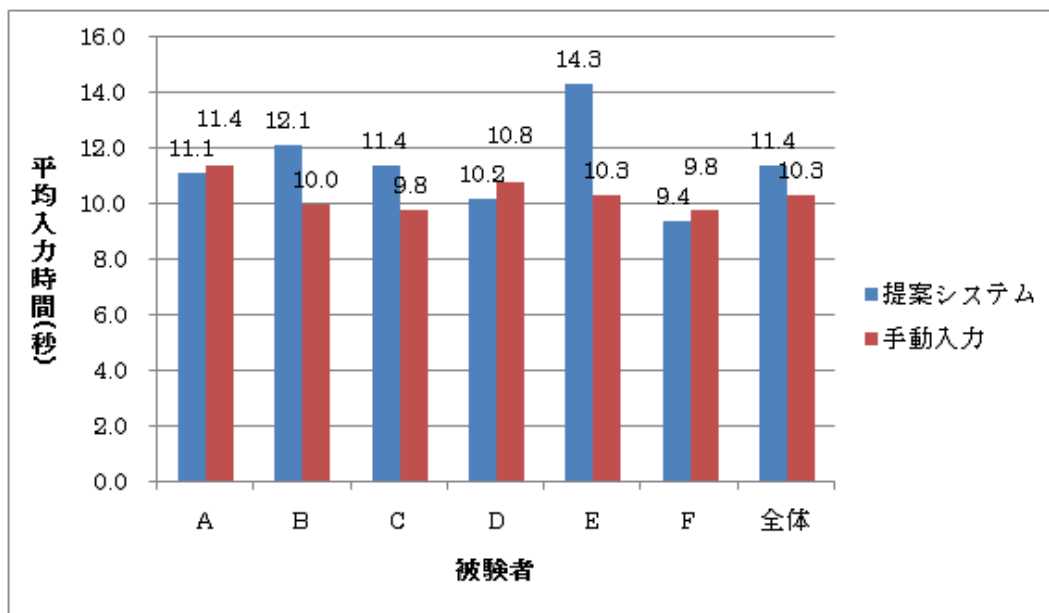


図 4.3: 平均入力時間 (タグ)

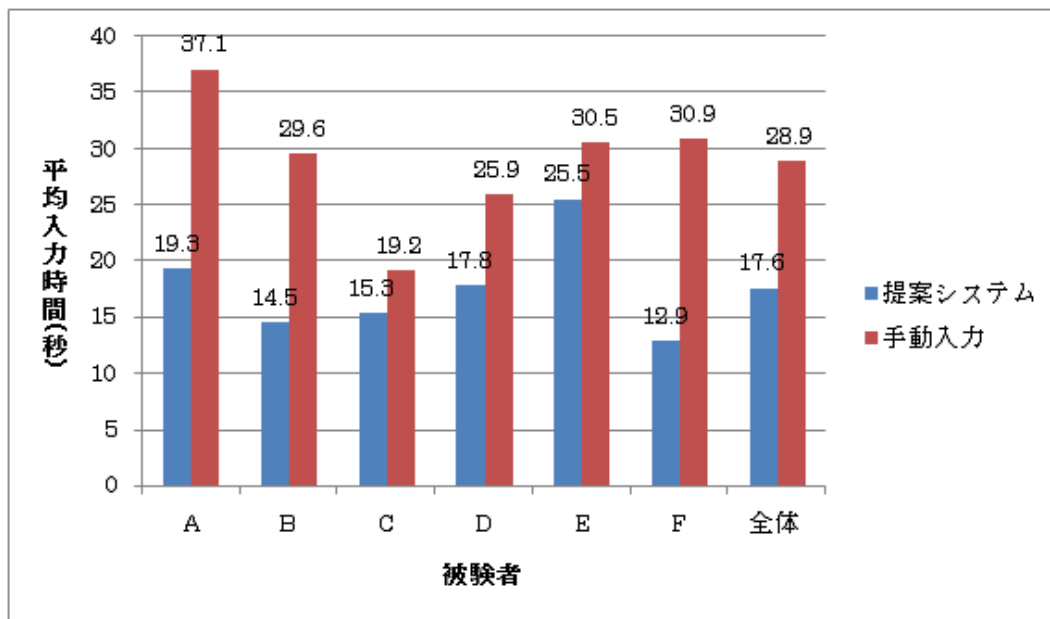


図 4.4: 平均入力時間 (タイトル+タグ)

タイトルとタグの入力時間

本研究の目的は、提案システムを利用することにより、タイトルとタグを短い時間で入力できるようにすることである。図 4.2 より、タイトルの入力時間は、個人差はあるものの全被験者において減少していることが分かる。一方、図 4.3 より、タグの入力時間は、提案システムを利用した場合とそうでない場合との差がほとんどなく、個人差もほとんどないことが分かる。手動で作成したタグから選択する方が提案システムを利用するよりも短時間で入力を終えている。これは、自分で作成したタグのみが候補として現れる場合と、さらに追加でノートから抽出されたタグが表示された場合とで、表示されるタグの数が異なることが理由として考えられる。人によって研究ノートに書かれている文章量は異なるが、平均 13.9 個のタグが抽出されて表示されている。ユーザータグは一人当たり平均 18.5 個作成されていたことから、合計すると、提案システムを利用した場合の方が手動入力の場合と比較して 1.8 倍のタグが表示される。また、図 4.4 にタイトルとタグの入力時間の合計を示す。すべての被験者において、提案システムを用いた場合のほうが入力時間が短くなったことが分かる。これらの結果から、手動で入力する場合と比較して、記録コストの面で有効であることが確認された。

入力タイトルの多様性

アンケートでは、手動によるタイトルの入力が面倒であったという意見が3名から得られたが、ノートタイトルを選ぶ方が面倒であったという意見も1名から得られた。その理由は、音声ログの内容に最も関連しているノートタイトルを探すより、単にタイトルを入力したほうが、考える手間が要らないというものであった。さらに、タイトルの手動入力と、関連するノートタイトルの選択のどちらかを行える、あるいは両方とも行ってもよいような仕組みがあれば嬉しいという意見が得られた。提案システムでは、研究ノートに含まれないタグは抽出できないため、ユーザーが選択可能なノートがそもそも存在しない状況もあり得る。そこで、手動入力する部分と、簡単に入力を行える部分が両方組み合わさった仕組みを実現し、そのバランスをうまく調整することが重要であると考えた。

ノートタイトルから音声ログのタイトルを選択する場合には、有限のノートの中から選択することで、同じようなタイトルばかりになることが容易に想像される。しかし、タイトルを手動入力する場合にも、同じようなタイトルを音声ログに付けてしまうという意見が得られた。その原因として考えられる2つの意見が得られた。1つは、音声ログの記録時に手間をかけたくない、あるいは手間がかけられないような状況において、短時間で入力できるタイトルを入力したという意見である。もう1つは、実験の設定として1日9件以上の記録を条件としたため、その条件を満たすために、とにかく記録を行い、簡単にタイトルの入力を済ませたという意見である。なお、このように同じようなタイトルを入力していた被験者は6名中3名であり、半数がタイトルの入力を手間だと考えていた。

タイトルの多様性の確保とユーザーの利便性には、トレードオフの関係があるため、有用なタグの抽出が重要であると言える。

タグの抽出効果と多様性

アンケート結果には、ノートから抽出されたタグから必要なタグを探すのが面倒であったという意見や、そもそも付与したいタグがその中に無かった、などの意見があった。付与したいタグがなかった理由は2つ考えられる。

1つめの理由は、提案システムにより研究ノートから抽出されるタグが、音声ログのタグとして適当でなかったということである。ユーザーにとって必要なタグが何であり、どのように研究ノートからタグを抽出すれば、タグ候補として一覧に表示されるのかをさらに分析する必要があると考えられる。

2つめの理由は、実験期間の短さである。提案システムが支援するタグの付与は、ユーザータグに含まれない、研究ノートから抽出されたタグがある場合に大きな効果を発揮する。本実験を行った約1ヶ月の実験期間中に、作成または更新された研究ノートの数は一人当たり平均3.3件と少なく、提案システムが有効な

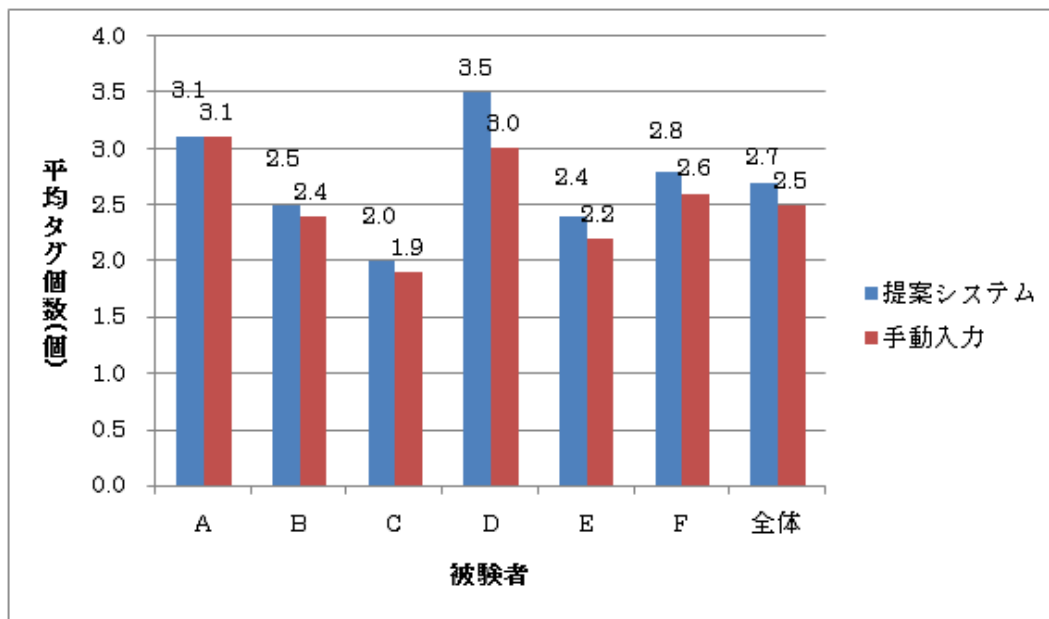


図 4.5: 音声ログ1件当たり付与された平均タグ個数

タグを抽出できなかった可能性がある。そのため、継続して運用を行い、提案システムが有効なタグを抽出できるかどうか分析を行う必要があると考えられる。

付与したいタグを探すのが面倒であったという意見に対しては、表示を工夫するなど、簡単に探せるような仕組みを開発する必要があると考えられる。また、提案システムでは音声ログのタイトルとして利用するため、選択するノートの数に1つに限定している。そのため、必要なタグが選択したノートに含まれず、他のノートに含まれている場合が考えられる。この問題を解決するために、複数のノートを選択するなどの方法により、多様なタグを抽出できる仕組みを実現する必要がある。

また、付与されたタグの内容や、個数についても分析を行った。音声ログに付与されたタグの個数を比較した結果を図4.5に示す。個人差があるものの、すべての被験者において提案システムを利用した場合に付与するタグの個数が多くなる傾向にあった。これは、ノートから抽出されたタグが音声ログ記録時のタグ一覧に表示されることにより、タグの選択候補が増えたことが理由であると考えられる。

次に、新しいタグが音声ログに付与される個数について分析をした。1名あたり144件から148件の音声ログが記録されたが、最初に付与されるタグは当然新規タグになってしまうため、そのうち最初の40件を初期記録期間として除外した。それ以降の64件から68件の音声ログに対して付与されたタグのうち、新規のものを数えたところ以下の図4.6に示すような結果となった。個人差はあるものの、すべての被験者において提案システムを用いた場合の方が新しいタグを付与する

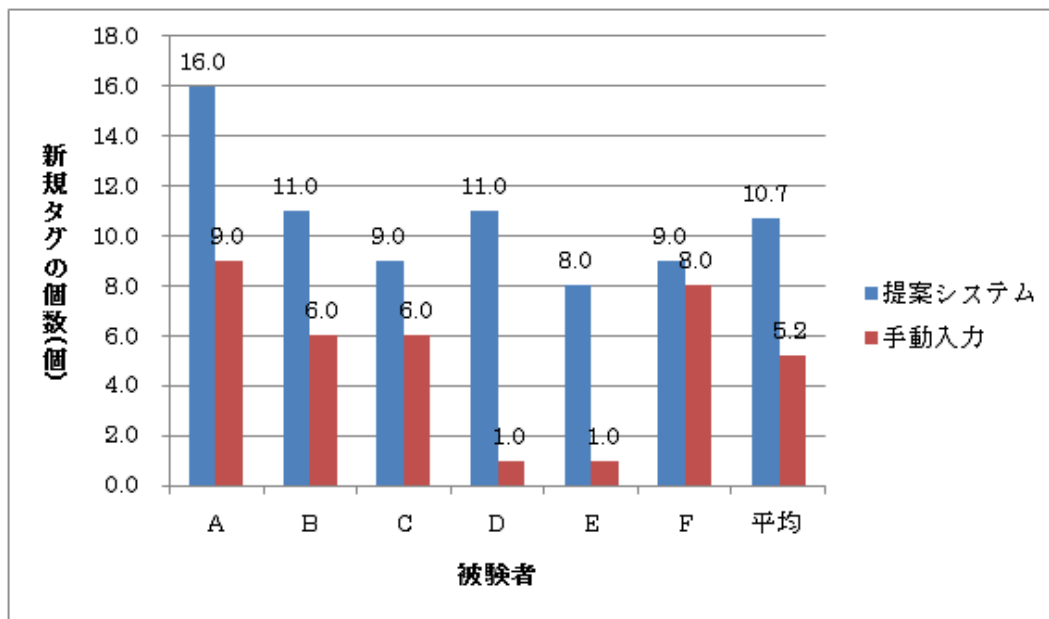


図 4.6: 音声ログに付与された新規タグの個数

傾向にあることが分かる。これは、タグ候補が増えるだけでなく、わざわざユーザータグとして作成するまでもないタグや、咄嗟に出てこないタグが一覧に表示されることが理由として考えられる。

以上の結果から、提案システムは多様なタグの付与を可能にしていると言える。

その他の意見

TODO で宣言した内容を行動に移した後、それを事実で記録している場合が多かった。つまり、TODO と事実が対応付いている場合がほとんどであった。このことから、TODO の音声ログに対して、チェックを付けるなどの方法で行動したことを記録したほうがよいのではないかという意見があった。そのような場合には、対応する TODO の音声ログを探す仕組みを提供することにより、容易に記録できることが望まれる。

また、研究活動以外の音声ログを記録してもよいのではないかという意見があった。もっと日常的な記録を残すことにより、研究活動に関係する音声ログを探す手がかりとなるのではないかという意見である。本実験では、研究活動に関係する音声ログのみを記録してもらったが、その他の音声ログについても自由に記録した場合、利用の際にどのような結果が得られるのかについて分析する必要があると考えられる。

その他にも、音声ログにはテキストには含まれない情報が含まれているという

意見があった。音声を聴き返すことにより、記録時の自分の感情や状態、周囲の状況が思い出せるということである。これらの情報を分析することは困難であるが、テキストには含まれない有用な情報として、利用できると思われる。

4.2 音声ログの検索に関する実験

4.2.1 実験の目的と方法

本実験の目的は、提案システムによるタイトルとタグの付与が、整理に有効であるかどうかを評価することである。具体的には、被験者に音声ログの再生を必要とする問題を出題し、回答するのに要した時間、および、回答中に再生した音声の数を計測する。4.1 節の実験の終了後、1 週間が経過した後に本実験を行った。被験者に一人ずつ問題を出題し、回答してもらった。被験者には、記録した音声ログを聞き返すことにより回答できる問題を出題し、タイトルやタグを見たり、問題文に含まれるキーワードで検索しても、目的となる音声が一意に定まらないようにした。

被験者とそのグループ分けは 4.1.1 項と同様である。提案システムを利用して記録した音声ログのみを表示するグループ I と、タイトルとタグを手動で入力した音声ログのみを表示するグループ II に分けることにより、それぞれのタイトルとタグの有効性を比較した。また、両グループの被験者には、利用していない半分の音声ログを用いて、タイトルとタグを非表示にした場合の問題にも回答してもらい、タイトルとタグの有無による差を比較した。両グループの実験の流れを図 4.7 に示す。

カテゴリーごとに問題を作成した。アイデア思考に関する問題を 1 つ、事実に関する問題を 3 つ、TODO に関する問題を 2 つ用意した。また、タイトルとタグが表示されない場合の問題については、事実に関する問題を 2 問用意した。

アイデア思考に関しては、全ての被験者に共通で「これまでに記録したアイデア思考の中で、最も大事な話題は何ですか」という問題を出題した。これは、特に正解のない問題であり、あくまでも被験者の主観によって答えられる問題であるため、回答の内容が音声ログに含まれる内容であることを回答の正当性としての条件とした。

事実に関しては、記録された音声ログを聞いて、被験者ごとに適切な問題を作成した。具体的には「〇〇を行ったのはいつですか」「〇〇が分かったのはいつですか」といった、一度だけ記録された事実を問う問題とした。問題設定には、ある程度恣意性が含まれてしまうが、回答中の検索で絞られる候補ができる限り同数となるような問題とした。また、事実に関する問題は 3 問出題したが、図 4.8 に示すように、3 問の回答に必要な音声ログが、できる限り離れて表示されるような問題とした。

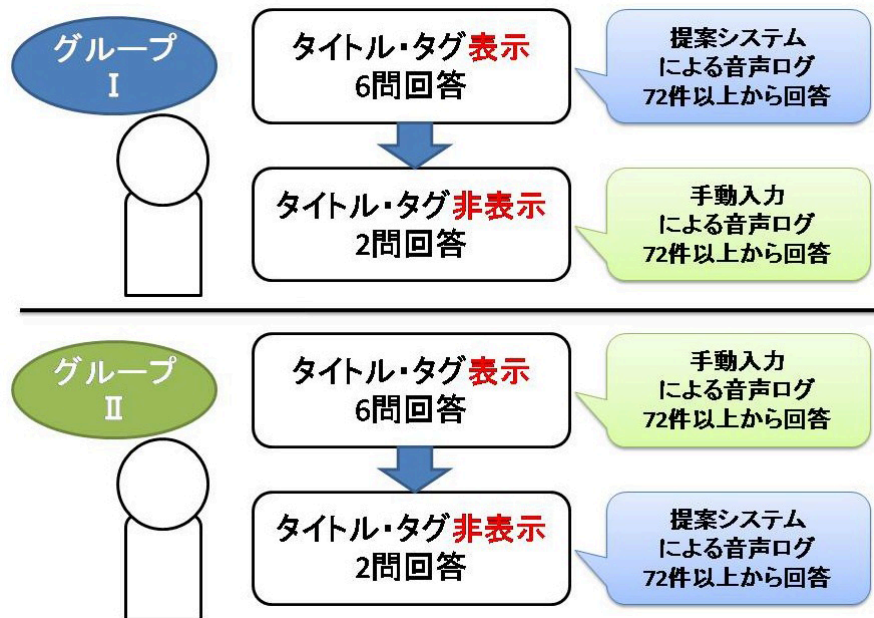


図 4.7: 実験の流れ

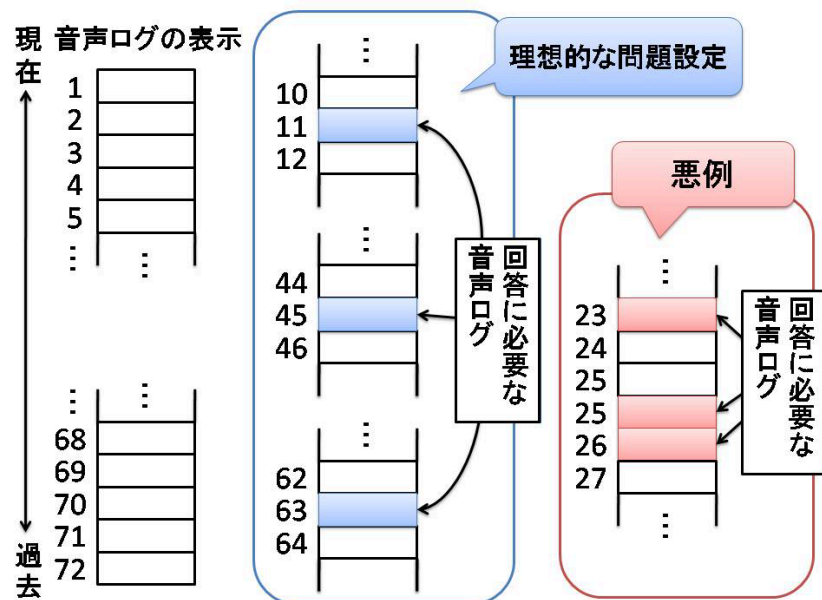


図 4.8: 事実に関する問題設定例

TODO に関しては、「〇月〇日～〇月〇日の期間の TODO で、現在も達成できていないことはありますか？あればその内容を教えてください」という問題とした。この問題もアイデア思考と同じように特に正解がなく、被験者の主観によって答えられる問題である。問題の対象とする期間内における TODO の音声ログが被験者ごとにほぼ同数となるようにした。具体的には、約 3 日分の TODO である約 12 件が対象となるような問題とした。また、TODO に関する問題は 2 問出題したが、1 問目は実験日のちょうど一週間前から過去の約 12 件、2 問目はさらにその時点から過去の約 12 件を対象とした。ちょうど 12 件とならない場合があるのは、各被験者によって一日の記録件数にばらつきがあるためである。

アイデア思考、事実、TODO の問題はどれも、音声ログの一覧から必要なものを再生することにより回答できるような問題設定となっているため、回答時間から音声ログが整理されているかどうかを検証する。

なお、実験後に口頭および紙面によるアンケート調査を行った。口頭アンケートでは、「タイトルとタグが音声の整理に有効か」「タイトルとタグのどちらが整理に役立っているか」について質問し、その回答および自由な意見を録音した。紙面アンケートでは、事実の 3 問に関して、「役に立ったタグ」「あればよかったタグ」「検索したキーワード」について質問した。

4.2.2 実験結果

各カテゴリーの問題に対する全被験者の平均回答時間を図 4.9 に、回答中の音声ログの平均再生回数を図 4.10 に示す。なお、各被験者ごとのデータは付録に記載する。

図 4.9 から、アイデア思考の問題に対する回答時間に大きく差が出ており、提案システムの方が早く回答できていることが分かる。また、事実・TODO の問題に対する回答時間にも差があり、どちらも手動でタイトルとタグを入力した音声ログの方が回答時間が短いという結果になっている。

図 4.10 から、アイデア思考の問題に対する回答中の音声ログの再生回数には差がないことが分かる。また、事実・TODO の問題に対する回答中の音声ログの再生回数は、どちらも手動入力の方が少数であることが分かる。

次に、被験者ごとの各カテゴリーの問題に対する平均回答時間を、アイデア思考は図 4.11 に、事実は図 4.12 に、TODO は図 4.13 にそれぞれ示す。どのカテゴリーの問題においても、被験者ごとの個人差が大きく出ていることが分かる。

事実に関する問題は、タイトルとタグが表示されていない場合についても 2 問出題した。タイトルとタグが表示されている場合と比較した平均回答時間についてのグラフを図 4.14 に、回答中の音声ログの平均再生回数についてのグラフを図 4.15 に示す。図 4.14 から、グループ II の方がグループ I より平均回答時間が短いことが分かる。提案システムを用いた方が、手動でタイトルとタグを入力し

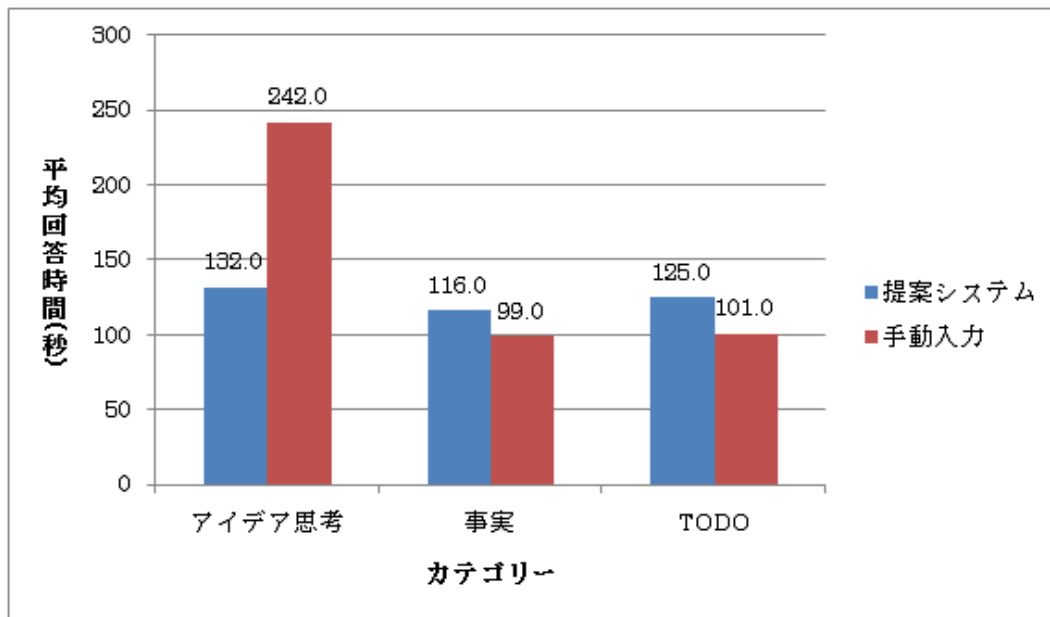


図 4.9: 平均回答時間

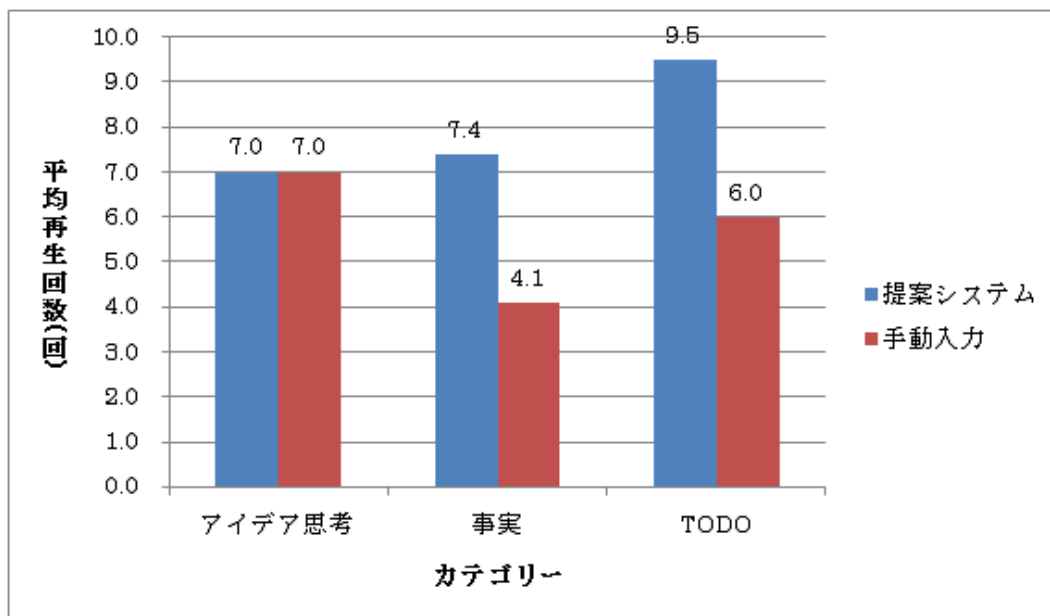


図 4.10: 音声ログの平均再生回数

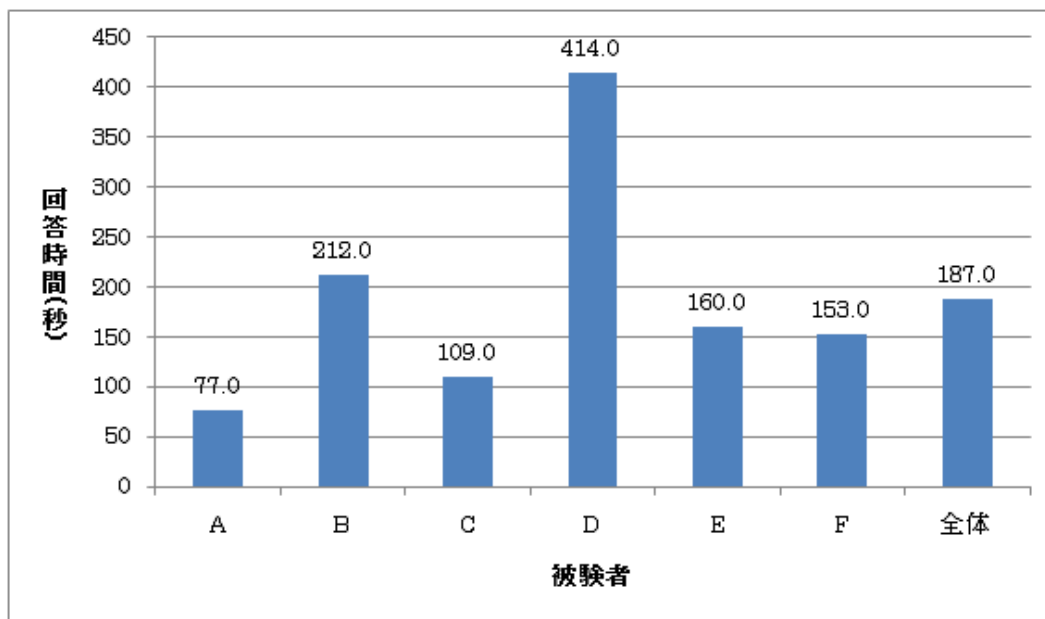


図 4.11: 回答時間 (アイデア思考)

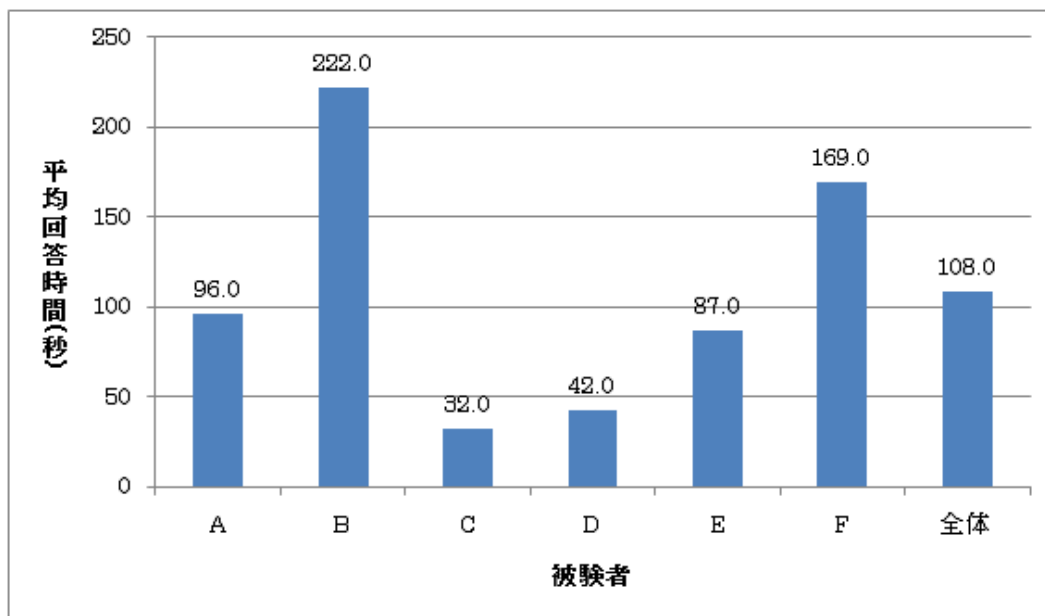


図 4.12: 平均回答時間 (事実)

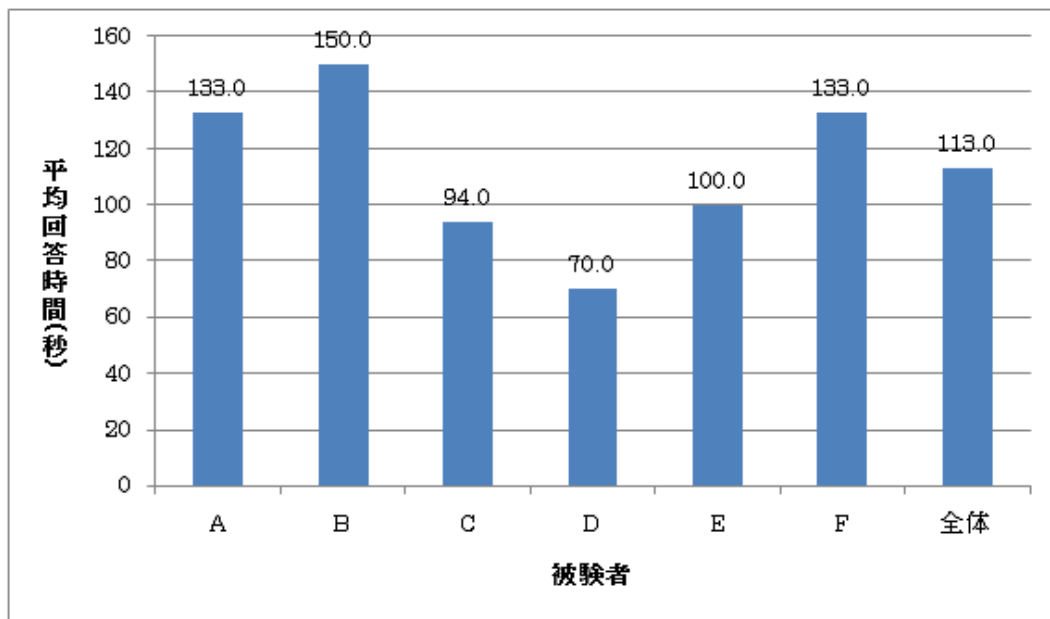


図 4.13: 平均回答時間 (TODO)

た場合に比べて、117.2% の時間がかかっている。また、回答中に再生した音声ログの数は、グループ II がグループ I に比べてかなり少ないことが分かる。

4.2.3 考察

本項では、検索時間、タイトルの有効性、タグの有効性、およびアンケートから得られたその他の意見について考察する。

検索時間

図 4.9 より、アイデア思考に関する問題において、回答時間に大きな差が見られるが、回答の仕方に個人差がかなり見受けられたため、それによる差であると考えられる。すなわち、タイトルやタグから内容を思い出すことで、すべての音声を聴き返すことなく回答する被験者がほとんどであったのに対し、すべての音声を聞き返し、その中から回答を吟味する被験者が居たことによる差である。また、アイデア思考の音声ログは、被験者によって録音時間が大きく異なる (図 4.1) が、回答時間と比較したところ、録音時間の短い被験者が回答時間も短いとは限らなかった。記録された音声の長さとは、その中から最も重要な話題を探し出す時間との間には、関係性は見られなかった。

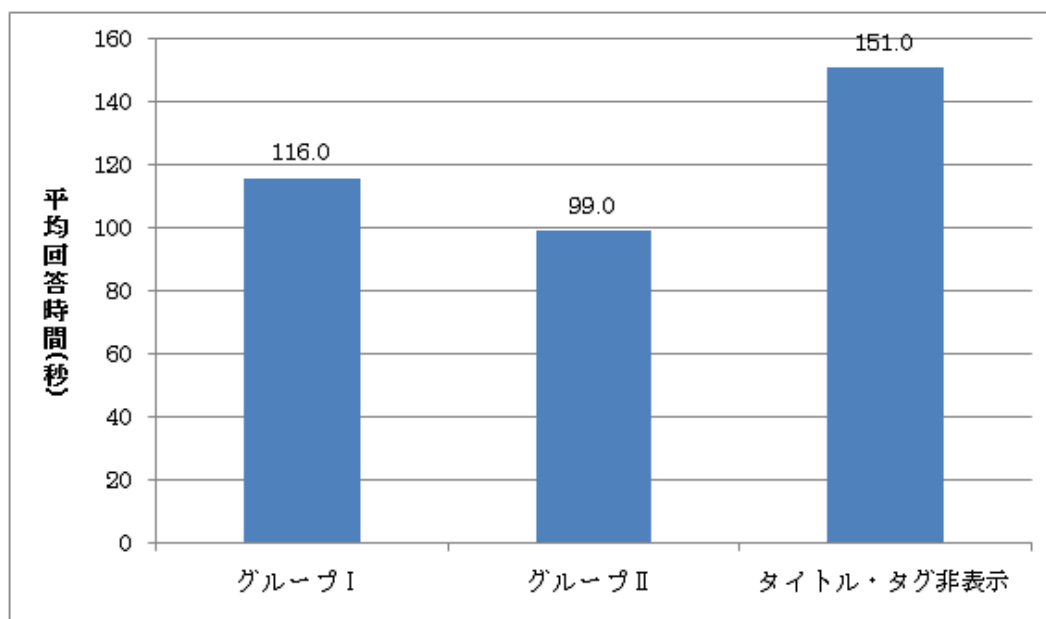


図 4.14: タイトルとタグの表示の有無による平均回答時間の比較

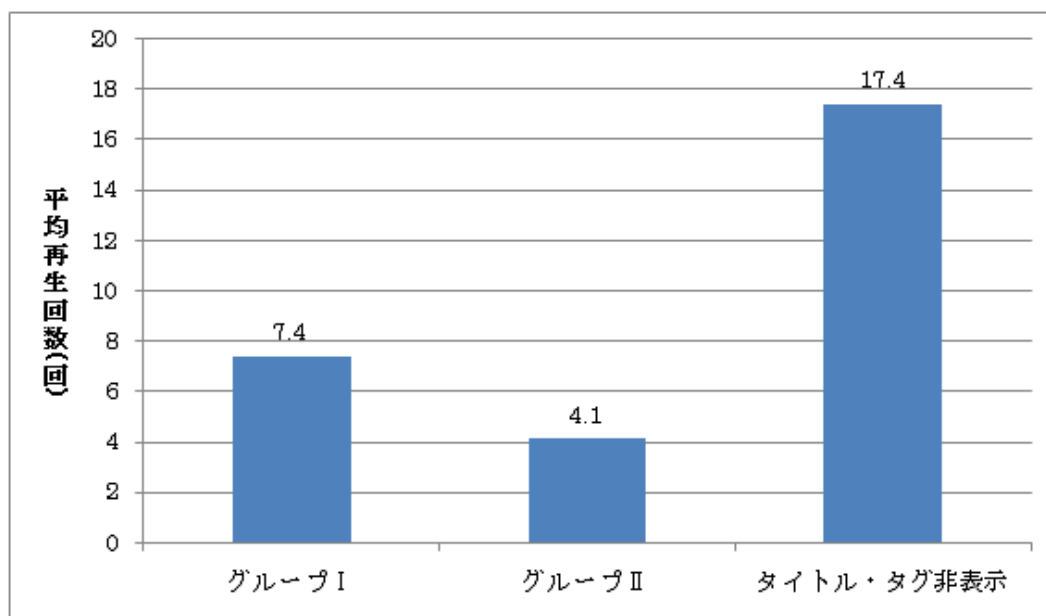


図 4.15: タイトルとタグの表示の有無による音声ログの平均再生回数の比較

事実に関する問題や、TODO に関する問題についても、回答時間に個人差がかなりあることが分かる (図 4.12, 4.13). しかし、これらはタイトルやタグの付与方法が影響したものであるとは言い切れない. グループ I とグループ II とで比較した場合、平均的にグループ II の方が回答時間は短いものの、グループ I にも回答時間の短い被験者がいるためである.

全体の結果として、提案システムを利用した場合の音声ログと、タイトルとタグを手動入力した場合の音声ログを比較して、手動入力の方が検索のコストが低いことが分かる. これは、タイトルを手動で入力したほうが正確な情報が付与され、閲覧時に内容を想起しやすいということが理由として考えられる. しかし、タイトルの自由度が低い提案システムにおいても、タグが付与されていることによって、検索が容易になり、タイトルを手動入力する場合と比べて、検索コストがある程度の差に収まっていることも事実である. 本実験では、音声ログに対してタイトルとタグが付与されていない場合と比較して、回答時間にそれほど大きな差は見られなかった (図 4.14). 研究活動が長期にわたると、録音日時はさらに役に立たなくなり、音声ログに付与された音声ログの有効性が際立つと予想される.

タイトルの有効性

タイトルとタグが音声ログの整理に有効かというアンケートでは、被験者全員が、タイトルとタグのある方が必要な音声ログに容易にアクセスできたと答えた. しかし、タイトルとタグが表示されている場合に、録音日時を重視せずにタイトルとタグから検索を行って目的の音声ログを探したために、回答が遅れてしまうケースがあった. このことから、音声ログにおいて、録音日時が重要なメタデータであることが分かる. このようなケースは、音声ログを利用する機会が日常的でないために、音声ログの利用に対する習熟度が低いことが原因で発生したと考えられる. 音声ログを利用する機会が増えれば、このようなケースは減るであろう.

タイトルとタグのどちらが整理に役に立っているかという質問では、タイトルと答えた被験者が 2 名、タグと答えた被験者が 2 名、両方と答えた被験者が 2 名であった. タイトルと答えた被験者は、タイトルを手動入力したグループの 2 名であり、タグと答えた被験者は、提案システムを利用したグループの 2 名であり、両方と答えた被験者は、それぞれのグループから 1 名ずつであった. タイトルを手動入力すれば、タイトルが役に立ち、提案システムを利用すればタグが役に立つことがわかる. しかし、タイトルを手動入力したグループから、タグが役に立たなかったという意見はなかったが、提案システムを利用したグループからは、タイトルが役に立たなかったという意見があった. 提案システムを利用した場合の検索時間が長いことから、提案システムより手動入力の音声ログの方がよく整理されていることが分かる. この問題を解決するために、情報の付与をより容易に行うための仕組みを考案する必要がある.

タグの有効性

事実の3問に関しては、回答時に役に立ったタグ、あればよかったタグ、検索したキーワードについて質問した。グループIのうち2名、グループIIのうち2名が、事実というタグが役に立ったと回答した。また、それ以外のタグで役に立った数は、合計でグループIからは9個、グループIIからは4個得られた。このことから、グループIの回答にはタグが重視されていることが分かる。さらにその内容を分析したところ、グループIの9個のうち7個、グループIIの4個のうち1個が、音声の内容には含まれないタグであった。これにより、音声認識では困難な情報の付与と整理が行えていることが分かる。ただし、「あればよかったタグは何か」という質問では、両グループの合計で、音声の内容には含まれないタグが4個であるのに対し、音声の内容に含まれるタグが15個であった。また、実際に検索されたキーワードも、音声の内容に含まれるキーワードが8個あったのに対し、音声の内容に含まれないキーワードはわずか1個であった。これらのことから、タイトルとタグの付与だけでなく、音声認識によるタグの抽出も併用することが望ましいと言える。

その他の意見

タイトルを手動で入力したとしても、後で役に立たないようなタイトルになってしまっている場合があるという意見が得られた。例えば、記録したときは最も適していると考えて入力したタイトルに固有名詞が含まれていた場合、しばらく時間が経った時に、その固有名詞が何なのかをそもそも覚えておらず、役に立たないという場合である。このような場合に、タグにより大まかな内容を知ることが、非常に重要な手がかりであると考えられる。

アイデア思考の音声ログについては、タイトルを見ても内容が分からない、という意見があった。記録されている内容が多岐に渡る場合や、単純に長時間の場合など、きちんと聴きなおさなければ内容を正確に把握できないということである。アイデア思考は、短時間の音声であっても、内容の複雑さから、タイトルとタグによる情報の補足はあまり効果が無いと考えられる。しかし、本研究で目的としているのは、音声ログの整理である。アイデア思考というカテゴリーに属し、そのうえで複数のタグが付与されており、容易にアクセスが行えていることから、アイデア思考の音声についても整理はできていると言える。今後は、このような音声ログの中身についても分析し、ユーザーが簡単に情報を獲得できるような仕組みを考える必要がある。

4.3 まとめ

音声記録時にタイトルとタグを付与する方法として、提案システムは手動で入力する場合と比較して、60.8% の時間で付与することが可能であった。音声検索時には、提案システムを利用した音声ログから探す場合は手動で入力した音声ログから探す場合と比較して、117.2% の時間が必要となることが分かった。記録時のコストが低くなる代わりに検索時のコストは高くなるというトレードオフの関係にあることから、記録と検索のどちらの側面を重視するかによって、どちらが良い記録方法であるかは変わってくる。今後、長期的な運用を行いつつ、適切なタイトルの入力方法やタグの抽出・表示方法について、更なるシステムの改良を行う必要がある。

第5章 音声ログの利用に関する展望

本研究では、蓄積された音声ログから必要な情報を獲得するまでの利用しか実現していない。しかし、提案システムによって付与されるタイトルやタグを用いることにより、様々な応用が期待できる。本章では、音声ログの利用方法について、今後実現可能であると考えられる仕組みについて、具体例を交えて説明する。

5.1 音声ログの研究ノートへの自動関連付け

提案システムを利用して記録された音声ログは、研究者が日常的に作成している研究ノートのタイトルや本文の情報を保持している。これを利用することで、研究ノートを拡張することができると考えられる。つまり、研究ノートの適切な箇所に音声ログが自動で関連付くような仕組みである。

音声ログに付与されたノートタイトルから、関連付けるべき研究ノートは一意に決まる。さらに、音声ログに付与されたタグが、当該ノートから抽出されたタグである可能性が高いことより、研究ノートの本文中のどの箇所に音声ログを関連付ければよいかの判断を、ユーザーが任意にタグを付与した場合と比べて、より容易に行えると考えられる。関連付ける箇所は、図5.1に示すように、行、段落などであり、その判断基準は、出現するタグの数などが考えられる。

研究ノートの中で内容的に関連している箇所に音声ログが付与され、図5.2に示すような簡単なインタフェースを用意することにより、研究者は研究ノートの読み書き中に関連する音声ログを再生し、必要な情報を容易に発見することができる。

このシステムが実現すれば、より自然な形式で音声ログの利用が促進され、利用されずに埋もれてしまう音声ログが減ったり、わざわざ音声ログを探し出す手間がなくなるなどの効果が期待できる。

自動で関連付けることにより、システムがユーザーの意図しない箇所に音声ログを関連付けることもあるだろう。そのような場合に、ユーザーが任意に関連付ける箇所を変更したり、あるいは関連付けそのものを削除できる必要がある。また、そのような関連付けの修正や削除の履歴情報をサーバーにアップロードし、管理することで、以後の自動関連付けの精度を高めていくなどの方法が考えられる。

本研究において提案したシステムでは、ボイスサーバーで音声ログを蓄積・管理するため、研究ノートから直接ボイスサーバーにアクセスすることにより、音

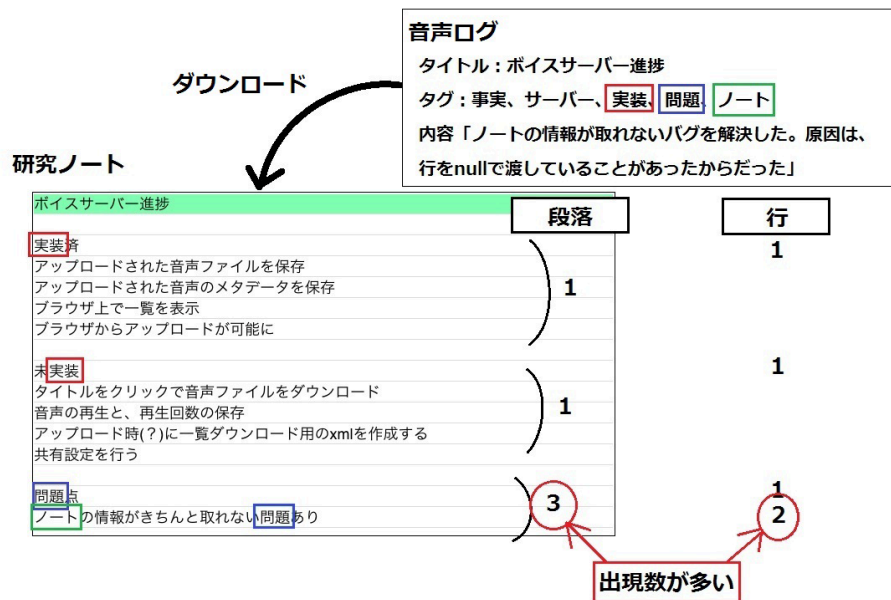


図 5.1: 音声ログの研究ノートへの自動関連付け例



図 5.2: 研究ノートの行に関連付いている音声ログの例

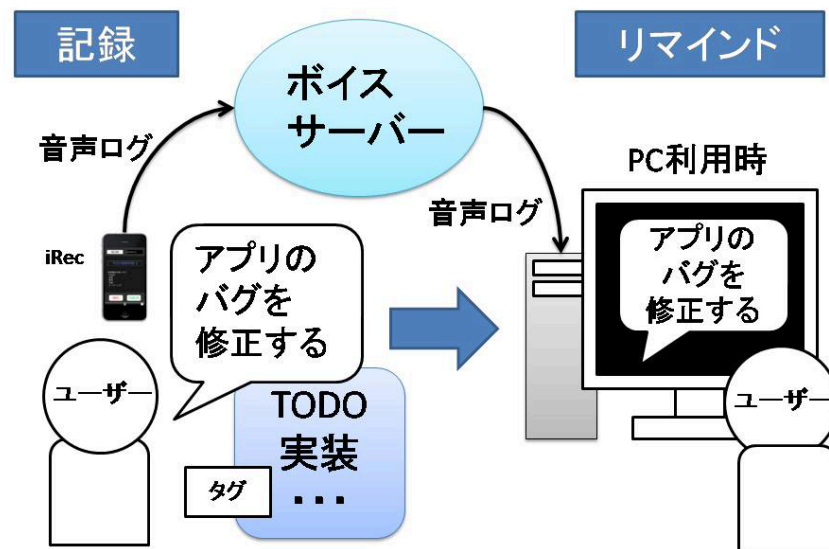


図 5.3: リマインドの流れ (PC 利用時)

声ログをダウンロードすることができる。また、提案システムにおいて利用した研究ノート作成システムである iSticky は、研究ノートの本文を行ごとに管理しているため、音声ログを関連付ける箇所として行単位を指定することができる。今後実現すべきこととしては、研究ノート上での音声ログの再生・編集インタフェース、関連付けの編集機能などが挙げられる。

5.2 適切なシーンでの音声ログのリマインダー

本研究では、研究活動における短時間の音声ログを、アイデア思考、事実、TODO の 3 種類に分類した。その中で、TODO に関する音声ログは、記録される件数が多く、また適切な状況で思い出す必要のある内容が記録されている。音声ログにタグが付与されていることで、TODO を適切な状況で想起させる仕組みを実現できると考えられる。

例えば、図 5.3 に示すように、PC を用いた研究活動に関する TODO の音声ログを記録した後、研究者が PC を利用する際に、PC 上でその音声ログが自動再生されることにより、実行すべき TODO を想起させるという仕組みである。

音声ログを利用する状況は、PC での作業の他にも、場所や時間などを指定することも可能である。タグからリマインドさせる場所や時間を自動で判断し、研究

者に TODO を想起させる仕組みである。例えば、「実験」というタグが付与された音声ログであれば実験室に移動した時に、「1 週間後」というタグが付与された音声ログであれば音声記録されてから 1 週間後に、スマートフォン上でリマインドさせるなどの方法である。リマインドしたいタイミングをユーザーが意識してタグとして付与することにより、適切なリマインドが可能となる。また、記録時に、タイトルとタグだけでなく、予めリマインドしたい日時や場所を指定する機能を追加することにより、さらに正確なリマインドを行うことも可能である。

この仕組みにより、研究者は忘れてしまっていたことを容易に思い出すことができ、利用されずに埋もれてしまう音声ログが減るなどの効果が期待できる。iPhone には GPS 機能や、通知機能と呼ばれるリマインダーに相当する機能があるため、場所や時間によるリマインドは、本研究で実現した iRec を若干改良することにより実現可能である。PC での作業に伴ったリマインドをさせるには、PC 上で動作し、ボイスサーバーから音声ログをダウンロードして、自動再生するクライアントを新たに開発する必要がある。

この仕組みの問題点は、「今週中に」「〇月×日までに」といった期限を指定する TODO の場合、最終日にリマインドされても効果的ではないため、いつ、どのタイミングでリマインドを行うべきかを判断する必要があるということである。この問題の解決法として、単純に毎日リマインドさせるという方法が考えられる。しかし、あまりに頻繁にリマインドされれば、ユーザーはその提示を煩わしく感じ、さらには注意しなくなる恐れがある。十分に時間間隔を開けてリマインドさせる方法では、リマインドが十分行えているのか、あるいは過多なのかは、ユーザーの判断に委ねられる。ユーザーごとに個人適応する仕組みを持ったリマインダーを実現すべきか、それとも機械的に判断をするリマインダーを実現すべきなのか、今後さらに分析・考察する必要がある。

第6章 関連研究およびアプリケーション

6.1 音声ログの記録と利用に関する研究

音声ログを記録し利用する先行研究として、連続音声録音による記憶補助システムの研究がある [8, 9]. ユーザーが IC レコーダーなどの音声記録装置を身に付けて常時録音し続け、記録された音声を記憶の補助のために利用するというものである. 同様の研究として、ウェアラブルコンピュータによる音声録音などが挙げられる [10]. しかし、これらの音声記録は、他者のプライバシーを侵害する恐れがあり、常時録音可能な場面は限定されてしまう. また、長時間の音声ログが記録されることで、その中で重要な箇所のみを抽出するなどの構造化が必要となる. 本研究では、ユーザーは記録したいと考えた時だけスマートフォンを取り出して録音を行う. ユーザーにとって意味のある音声を取捨選択することにより、記録される音声ログの1つ1つは特定の目的、すなわち、会議音声や、講演音声、思いついたアイデアや後ですべき TODO、見たり聞いたりその場で行った行動などの意味を持つため、会議など長時間にわたる一部の音声ログを除いて、セグメンテーションの必要はない.

6.2 タイトルやタグの付与可能な録音アプリケーション

iPhone/iPod touch には標準アプリとして、ボイスメモ (ボイスレコーダー) が搭載されている. ボイスメモには、ラベルという呼称でタイトルやタグに相当するものを音声ログに付与することができる. ボイスメモにおけるラベルは、音声記録時ではなく、後から任意に付与することができる. 予め用意された Podcast, インタビュー, レクチャー, アイデア, ミーティング, メモの中から選択するか、カスタムを選択してユーザーが任意の文字を入力してラベルとすることができる. 予め用意されたラベルは、本研究におけるカテゴリーに相当すると考えられる. しかし、ボイスメモにおいては、ラベルを1つしか選択することができず、その他タイトルやタグなどを付与することはできない. カスタム選択によってラベルが手動入力された音声ログ以外、全て同じようなラベルが付与されることとなり、大量の音声ログを扱うには有効なアプリではない.



図 6.1: RecEver の設定画面

また、RecEver¹ は iPhone/iPod touch 用の録音アプリケーションであり、録音した音声をすぐに Evernote にアップロードして保存することができる。図 6.1 に示すような設定画面から、予め保存する Evernote のノートや付与するタグを選択しておくことができる。

RecEver は、ノートに音声ログを付与するという点で、5.1 節で述べたノートへの自動関連付けを、ノート自体。しかし、Evernote の仕様上、研究ノートのどの位置に関連付けるかまでは選択できず、またボイスメモと同様に付与できるタグは1つであり、手動で入力する必要がある。

6.3 写真撮影と音声録音の相互補完性に関する研究

中蔵らは、写真と音声を組み合わせによる記録手法を提案している [11]。写真と手書きだけでは記録することのできない会話の盛り上がりや、音声録音によって保存するとともに、写真によって、音声記録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、音声記録の価値を高めることも確認されていることから、視覚情報が音声記録にとって重要であることが分かる。

本研究では音声ログに対してタイトルとタグというテキストによる視覚情報を付与した。タイトルとタグが音声ログの整理に役立つことは本研究でも確認され

¹RecEver : <http://itunes.apple.com/us/app/receiver/id478113345>

たが、視覚情報をより含んだ写真などを、音声ログのメタデータとして付与することにより、さらに有用な整理が可能になると考えられる。

6.4 音声認識を用いた音声の文字化と検索に関する研究

音声認識を利用して音声を文字化することにより、検索を容易にする手法がある。音声の検索時には、未知語が大きな問題となるため、Wechsler らは、単語列による書き起こしではなく、音素表記による書き起こしによって、検索語の音素表記とのマッチングを行っている [4]。また、日本語では、前田らが音素表記を用いて検索実験を行っている [12]。

しかし、単語を利用した方が検索性能が高いことから、単語と音素の両方を利用した検索システムが開発されている。Renals らは単語インデックスと音素インデックスを組み合わせ、検索実験を行っている [5]。

このように、音声認識技術は年々発達してきており、音声ログを音声認識によって文字化することには、検索において大きな意味がある。本研究では、音声認識では得られない整理に着目し、音声認識を利用しなかったが、今後、音声認識を併用したシステムの拡張を行っていくことで、検索時の利用をさらに容易にできると考えられる。

また、音声認識による音声ログの整理において、最も致命的な問題となるのが、音声認識誤りである。これを解決する方法として、認識誤りを減らすために、複数の音声認識システムを利用して、結果を統合するという方法が提案されている [13]。また、音声認識した文字と音声そのものを Web 上に公開し、不特定多数のユーザーが認識誤りを修正することで、音声の検索性能および認識性能をシステムの運用中に向上させるような仕組みも提案されている [6]。

本研究で提案するシステムでは、音声ログをボイスサーバーで蓄積・管理するため、音声ログを複数ユーザーで共有し、認識誤りを修正する仕組みを参考にできると考えられる。

6.5 音声認識を用いたスケジューリングに関する研究

Wong らは、ウェアラブルコンピュータを用いることにより、スケジューリングを容易に行うシステムである The Calendar Navigator Agent(以下 CNA)を提案している [14]。ユーザーはヘッドアップディスプレイを装着し、モニターに表示されたカレンダーから、記入された予定に素早くアクセスできる。また、常にユーザーの会話を録音し、音声認識にかけて文字化することにより、会話に含まれる約束や予定に関する発話を読み取り、リアルタイムにカレンダーの情報を更新する。この研究においては、音声認識技術を用いることで、リアルタイムに話され

た予定をスケジューリングしている．ユーザーの手間が要らず，しかも誤ったスケジューリングをユーザーが即座に確認できる点も重要である．Wong らは，誤ったスケジューリングを容易に修正する仕組みについても開発中としている．また，Starner は CNA に触れつつ，ウェアラブルコンピュータにおける音声入力的重要性を説いている [15, 16]．

本論文の 5.2 節において述べた TODO のリマインダーとは，録音を常に行うかどうかという点が大きく異なる．録音を常に行う場合，会話を正確に文字化し，スケジューリングするために，かなり精度の高い音声認識技術が必要となる．また，Olympus 社から発売されている VoiceTrek DM-4 という IC レコーダーは，実際に音声認識を利用したスケジューリング機能を持っている．日時の情報や予定の内容を発話することでスケジューリングを行うが，この機能の利点は，ユーザーが録音時に，音声認識をかけやすい発話を意識的に行うことができる点である．しかし，CNA や VoiceTrek は，あくまでスケジュールの管理を行うものである．ウェアラブルコンピュータを利用する場合，常に情報は提示されるが，慢性的に表示される情報をユーザーがどの程度意識するかといった問題が考えられ，また，そもそもそのような装置を装着することがためられるユーザーも依然として多いと思われる．

6.6 ミーティングなどの長時間の音声記録に関する研究

ミーティングなどの長時間に及ぶ音声ログは，セグメンテーションや構造化が必要であると考えられる．ミーティング音声の構造化については先行研究が多く存在する．Wilcox らは，ミーティング中に電子ノートにメモをとることで，メモの記録時間から録音した音声との対応付けを行い構造化を図る Dynamite というシステムを開発した [17]．また，Stifelman は，講義音声を記録する際に，ユーザーがノートを記述する行動と，話し手のピッチや中断，抑揚などから音声を構造化する The Audio Notebook を提案している [18]．長時間の音声ログを対象とする場合，これらの構造化の方法を利用したり，更なる構造化の方法を考案する必要がある．

著者らの研究室では，TimeMachineBoard[19]と呼ばれる，大型液晶ディスプレイによる電子ホワイトボードを用いたミーティングシステムを利用して，比較的小規模なミーティングを日常的に行っている．TimeMachineBoard では，iSticky の研究ノートに書かれたテキストをアップロードすることで，ボードにテキストを表示する．そして，テキストがボードにアップロードされた際の時間はログとして記録されるため，この情報を利用することにより，議論を記録した音声ログを構造化できると考えている．

第7章 まとめと今後の課題

7.1 まとめ

本論文では、研究活動における音声ログに着目し、研究ノートを利用することにより、音声ログに対して容易にタイトルとタグを付与して整理するシステムを実現した。

提案システムは、ノートサーバー、録音アプリ iRec、ボイスサーバーの3つの要素から構成される。ノートサーバーは3.2節で説明した通り、研究ノートを蓄積・管理し、ノートの本文に形態素解析を施すことにより、タグを抽出し、データベースに保管する。録音アプリ iRec は3.3節で説明した通り、音声ログの録音、再生、研究ノートの情報を利用したタイトルとタグの付与を行うことができる。ボイスサーバーは3.3節で説明した通り、iRec からアップロードされた音声ログを蓄積・管理する。

提案システムの有効性を検証するために、2種類の実験を行った。1つは、被験者に一定期間、提案システム及びタイトルとタグを手動入力するシステムにより音声ログを作成してもらう音声ログ記録時のメタデータ入力に関する実験である。この実験において、提案システムを利用することにより、音声ログに対してタイトルとタグを容易に付与できることが確認された。もう1つは、音声ログを記録してもらった被験者に問題を出題し、音声ログを再生することで回答してもらう音声ログの検索に関する実験である。この実験により、タイトルとタグの付与が音声ログの整理に役立つことが確認された。しかし、タイトルを任意に付与した場合と提案システムを利用した場合では、タイトルを任意に付与した場合の方が有効に整理できることも確認された。

7.2 今後の課題

7.2.1 大量のデータに基づく本手法の検証

本研究で行った実験では、特に音声ログの検索に関する実験において、問題に回答する時間に個人差が見受けられた。そのため、提案システムと、タイトルとタグを手動入力するシステムの比較においても、問題の回答時間に個人差が影響

していると考えられる。被験者の数を増やすことにより、個人差の影響を極力吸収して実験を行い、どのような違いが結果に表れるか検証する必要があると考えられる。また、問題設定によって個人差が生じた可能性があるため、できるだけ恣意性を排除した問題を用意することにより、個人差を吸収する方法が考えられる。

また、長期間のデータ収集を行った上で音声ログの検索に関する実験を行い、さらに大量の音声ログが溜まった状態で、タイトルとタグが整理に役立つかを検証する必要がある。長期的なデータ収集においては、本研究の音声記録実験で行ったカテゴリー選択や、一日9件以上の記録などの必須条件を被験者に課さずに、より自然な状況で記録を行ってもらい、そのデータを分析する必要があると考えられる。

7.2.2 タイトルおよびタグの付与方法

音声の利用においては、手動で入力したタイトルが役立つことが分かった。タイトルの手動入力とノートタイトルの選択を組み合わせたタイトルの付与方法など、よりユーザーの意図に沿ったタイトルを付与できる仕組みを実現しつつ、さらに記録時のコストを抑える方法について実現する必要がある。

また、研究ノートから抽出したタグには、そもそも音声ログに付与したいタグがなかったという意見が被験者から得られた。この問題を解決するためにはまず、ユーザーが研究ノートに記述しているテキストの分量を調査する必要がある。その上で、必要となるタグが研究ノートに含まれている場合、そのタグを高精度に抽出するような方法の実現を目指す予定である。

また、音声ログへの付与に適したタグが抽出されていた場合であっても、タグがたくさん表示されているために、必要なタグを探すのが困難だったという意見が得られた。この問題を解決するために、タグの表示方法の改善を図る必要がある。例えば、抽出されたタグの並べ方を五十音順や使用頻度順に簡単に並べ替えたり、推薦タグとユーザータグを並べて同時に閲覧できるような仕組みなどが考えられる。

音声ログに付与するタグとして、本実験では必ず、アイデア思考、事実、TODOのいずれかのカテゴリーを選択してもらった。これらのカテゴリーが、短時間の音声ログを記録するうえで必要十分であるかを分析する必要がある。本実験のアンケートからは、TODOに対応する事実の記録は必要ないという意見があったが、事実にはTODOに対応しない内容の音声も存在するため、不必要なカテゴリーであるとは言えない。アイデア思考は他のカテゴリーに比べて長時間記録される傾向にあったため、録音時間からカテゴリーを判断することのできる音声ログがあるが、1つずつの音声ログを比較すると、アイデア思考であっても短時間の記録となっている場合があったため、必ずしも録音時間がカテゴリーの推定に役立つわけではない。また、必ずこれら3つのカテゴリーのうち1つを選択してもらう

ような実験設定としたため、それ以外の内容の音声をもっとも記録しようとしなかったり、わざわざこれらのカテゴリーに合致する内容を記録するという場面が見受けられた。カテゴリーがこれら 3 つで適切かどうかについては、より多数の被験者に自由なタグ付けが可能な音声記録を行ってもらい、発話内容や付与されたタグを分析する必要があると考えられる。

7.2.3 音声認識の利用

本研究では、音声ログの整理には音声認識はあまり役に立たないという前提から、システムの開発・実験を行った。実際、5.2 節で述べたような、適切なシーンでの音声ログのリマインダーなどを実現する際には、音声認識によるタグはあまり役に立たないことが考えられる。しかし、アンケートの結果から、ユーザーが単に音声ログを検索する場合、音声認識の結果も有用な手掛かりとなり得ることが分かった。そこで、提案システムを利用したタイトルとタグの付与だけでなく、同時に音声認識を行うことで、より多くの情報が付与され、利用の容易な音声ログが作成できると考えられる。

7.2.4 音声ログの検索

4.2 節の実験で利用したようなブラウザでの音声ログ一覧や、iRec の音声ログ一覧から必要な音声ログを探す場合においては、音声ログの表示方法や検索方法として、より効果的な方法を検討する必要がある。

本研究で作成したブラウザ用の閲覧インタフェースでは、音声ログのタイトル、タグ、録音日時、録音時間を表示した。しかし、カレンダーが表示され、日付ごとに音声ログがまとまって表示されれば、ユーザーは日付の記憶を頼りに容易に音声ログにアクセスすることができると考えられる。カテゴリーやタグごとに音声ログの文字や背景の色を変更すれば、視覚的にも探しやすくなるし、表示する情報についても吟味が必要である。今後は、音声記録時に位置情報を取得することにより、利用時の手掛かりとすることを考えている。

本研究では、ブラウザ用の閲覧インタフェースの検索機能として、カテゴリーの絞り込みとハイライト検索機能を提供した。今後、継続的な運用によって音声ログの数が膨大になった場合、カテゴリーの絞り込みが有効に作用するかどうかや、ハイライト検索で見落としが発生しないかどうかについて、分析を行う必要がある。また、録音日時による検索機能がどの程度有効であるのかについても分析する予定である。

7.2.5 ミーティングなどの音声記録

本研究で対象としなかった長時間の音声ログについても、整理を行う仕組みを実現する必要がある。提案システムでは、短時間の音声ログを分類・整理するために、タイトルとタグを付与した。しかし、大量の音声ログの中から目的の音声ログを探し出すためにタイトルとタグが役に立っても、その音声ログが長時間に及ぶ記録の場合、内容を正確に思い出すことは難しい。アンケート結果においても、アイデア思考カテゴリーの音声ログの内容について正確に思い出すことは難しかったという意見があった。5分に満たないアイデア思考であってもこのような意見が出るということは、5分以上の長時間の音声ログの内容を思い出すことはさらに難しいと考えられる。

研究活動における長時間の音声ログには、ミーティングなどが考えられる。ミーティングはおよそ30分から数時間程度行われる長時間の記録である。この内容を思い出すためには、やはり音声ログを聴き返す必要があると考えられる。そして、必要な箇所や重要な箇所を聴き返すためには、音声ログのセグメンテーションと、構造化が必要であると考えられる。

謝辞

本研究を進めるにあたり、指導教員である長尾確教授には、研究に対する心構えなどの基本的な考え方から、研究に関する数多くの貴重な御意見などを賜り、大変御世話になりました。心より御礼申し上げます。

松原茂樹准教授には、研究の要点を教えて頂き、研究の考え方や論文執筆に関することなど、幅広く御指導をいただき、大変御世話になりました。心より御礼申し上げます。

大平茂輝助教には、研究活動においていつも様々なバックアップをして頂き、実装から論文執筆に至るまで、多くの御意見や助言を頂き、大変御世話になりました。心より御礼申し上げます。

石戸谷顕太郎さんには、研究のことからプログラミングのことまで、的確なアドバイスやご指導を頂き、大変お世話になりました。ありがとうございました。

高橋勲さんには、プログラミングに関して多くのアドバイスを頂きました。困っている時に何度も助けていただき、感謝しております。ありがとうございました。

渡辺賢さん、棚瀬達央さんには、ゼミ等で貴重なご意見を頂いただけでなく、研究活動全般における様々な面で御世話になりました。ありがとうございました。

尾崎宏樹さん、矢田幸大さんには、研究活動において数々のご意見を頂き、また研究室生活の様々な場面で御世話になりました。ありがとうございました。

長尾研究室の秘書である鈴木美苗さんには、研究室生活において、数多くのサポートをしていただきました。ありがとうございました。

最後に、陰ながら支え、応援してくれた家族に、心より感謝いたします。

参考文献

- [1] V. Bush. As we may think. *The Atlantic Monthly*, Vol. 176, pp. 101–108, 1945.
- [2] 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*, pp. 235–238, 2002.
- [3] Keigo Kitamura, Toshihiko Yamasaki, and Kiyoharu Aizawa. FoodLog: capture, analysis and retrieval of personal food images via web. In *Proceedings of the ACM Multimedia 2009 Workshop on Multimedia for Cooking and Eating Activities(CEA 2009)*, pp. 23–30, 2009.
- [4] M. Wechsler, E. Munteanu, and P. Schäuble. New techniques for open-vocabulary spoken document retrieval. In *Proceedings of the 21st Annual international ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 20–27, 1998.
- [5] S. Renals, D. Abberley, D. Kirby, and T. Robinson. Indexing and retrieval of broadcast news. *Speech Communication*, Vol. 32, No. 1-2, pp. 5–20, 2000.
- [6] 緒方淳, 後藤真孝, 江渡浩一郎. Podcastle: ポッドキャストをテキストで検索, 閲覧, 編集できるソーシャルアノテーションシステム. 日本ソフトウェア科学会第14回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2006) 論文集, pp. 53–58, 2006.
- [7] 高橋勲, 石戸谷顕太郎, 大平茂輝, 長尾確. 議事録への自動アノテーションによる議論理解支援. 情報処理学会研究報告, 2012-ICS-165-4, 2012.
- [8] S. Vemuri, C. Schmandt, W. Bender, S. Tellex, and B. Lassey. An Audio-Based Personal Memory Aid. In Nigel Davies, Elizabeth Mynatt, and Itiro Siio, editors, *UbiComp 2004: Ubiquitous Computing*, Vol. 3205 of *Lecture Notes in Computer Science*, pp. 400–417. Springer Berlin / Heidelberg, 2004.

- [9] G.R. Hayes, S.N. Patel, K.N. Truong, G. Iachello, J.A. Kientz, R. Farmer, and G.D. Abowd. The Personal Audio Loop: Designing a Ubiquitous Audio-Based Memory Aid. In Stephen Brewster and Mark Dunlop, editors, *Mobile Human-Computer Interaction 2004*, Vol. 3160 of *Lecture Notes in Computer Science*, pp. 168–179. Springer Berlin / Heidelberg, 2004.
- [10] 河村竜幸, 中西英之, 石黒浩. 骨伝導マイクとICレコーダを用いた記憶補助装置の構築. 情報処理学会研究報告, pp. 35–42, 2006.
- [11] 中蔵聡哉, 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. 人工知能学会第22回全国大会論文集, pp. 3C3–07, 2008.
- [12] 前田勇希, 島津明. 音素認識に基づく音声全文検索. 言語・音声理解と対話処理研究会, Vol. 33, pp. 1–6, 2001.
- [13] J.G. Fiscus. A post-processing system to yield reduced word error rates: Recognizer output voting error reduction (rover). In *Proceedings of IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 347–354, 1997.
- [14] B.A. Wong, T. Starner, and R.M. McGuire. Towards conversational speech recognition for a wearable computer based appointment scheduling agent. 2002.
- [15] T.E. Starner. Attention, memory, and wearable interfaces. *Pervasive Computing*, Vol. 1, No. 4, pp. 88–91, 2002.
- [16] T.E. Starner. The role of speech input in wearable computing. *Pervasive Computing*, Vol. 1, No. 3, pp. 89–93, 2002.
- [17] L.D. Wilcox, B.N. Schilit, and N. Sawhney. Dynamite: a dynamically organized ink and audio notebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 186–193, 1997.
- [18] L.J. Stifelman. *The audio notebook*. PhD thesis, 1997.
- [19] K. Ishitoya, S. Ohira, and K. Nagao. Acquisition of contextual information among meetings and its application to meeting recollection support. *International Conference on Knowledge Discovery and Information Retrieval 2011*, 2011.

付録

1. 静岡大学や九州大学の三次元環境に関する論文を読んだのはいつですか
2. 卒論 3 章の構成を、経路生成、センサーデータの重ね合わせ、三次元地図のポリゴン化、と変更したのはいつですか
3. ノードの探し方を変えて、確実にループするように変更したのはいつですか
4. これまでに記録したアイデア思考の中で、最も大事な話題は何ですか？
5. 2012/1/13~2012/1/23 の期間の TODO で、現在も達成できていないことはありますか？
あればその内容を答えて下さい。(複数可ですが、一つずつ口頭でお答えください)
6. 2012/1/1~2012/1/11 の期間の TODO で、現在も達成できていないことはありますか？
あればその内容を答えて下さい。(複数可ですが、一つずつ口頭でお答えください)
7. 経路の評価方法についてまとめたのはいつですか
8. 卒論 3 章を書き終えたのはいつですか

図 7.1: 問題例

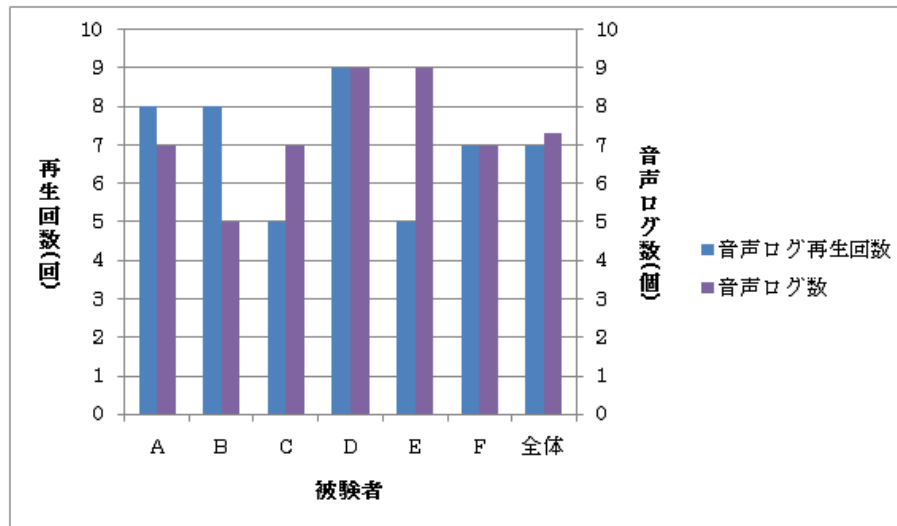


図 7.2: 音声ログの再生回数と音声ログ数 (アイデア思考)

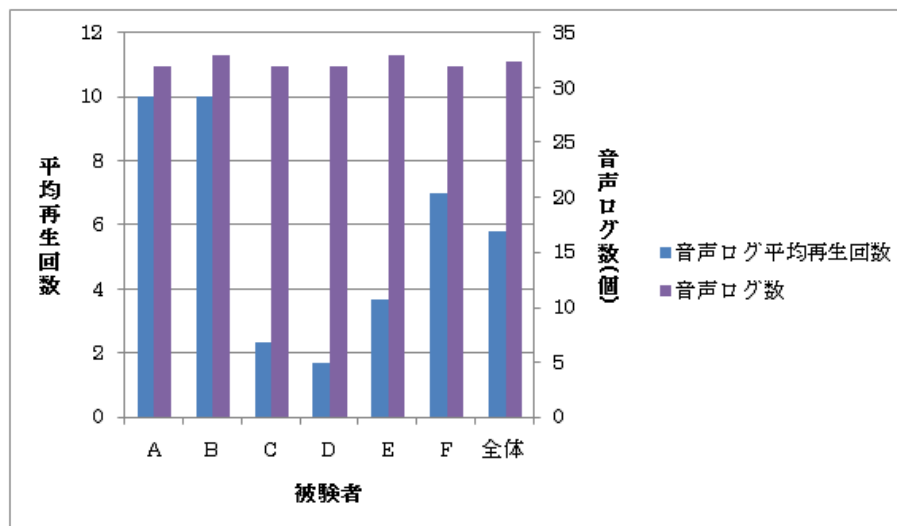


図 7.3: 音声ログの平均再生回数と音声ログ数 (事実)

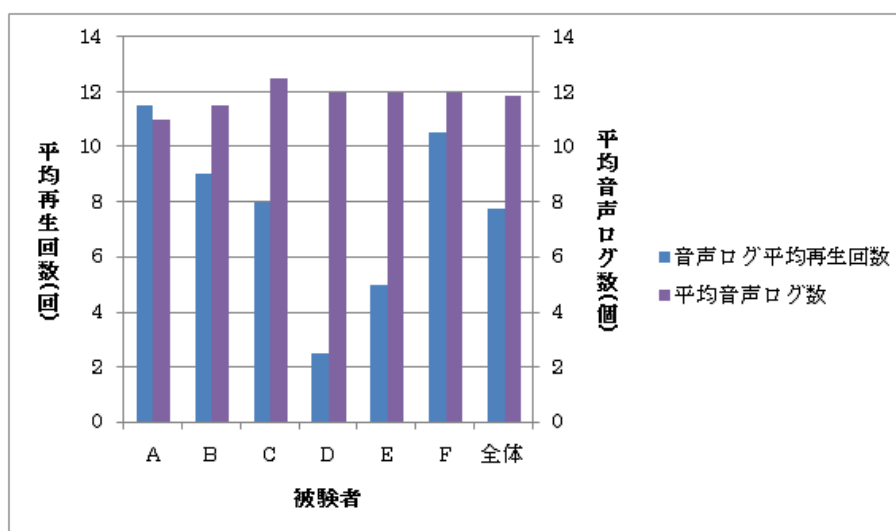


図 7.4: 音声ログの平均再生回数と平均音声ログ数 (TODO)