

研究ノートに基づく音声ログの整理とその利用

川 西 康 介^{†1} 大 平 茂 輝^{†2}
松 原 茂 樹^{†3} 長 尾 確^{†3}

研究者は、一般的に、研究活動における経過や成果をテキストで記録する。しかし、テキストによる入力が困難な状況も多く、音声でも記録できることが望まれる。音声による記録（以下、音声ログ）は、入力は容易であるものの、探すことが難しく利用性に乏しいという問題がある。そこで本稿では、テキストによる研究活動の記録（以下、研究ノート）を用いて音声ログを整理するシステムを提案する。本システムでは、研究ノートから自動抽出したタイトルとタグからユーザーが選択する。ユーザーは音声ログの記録のたびに、新規にタイトルやタグを作成する必要がなく、音声ログを整理するための労力を軽減できる。評価実験により、提案システムによる整理法の有効性と整理された音声ログの有用性を確認した。

Annotation of Voice Logs based on Laboratory Notebooks and its Utilization

KOUSUKE KAWANISHI,^{†1} SHIGEKI OHIRA,^{†2}
SHIGEKI MATSUBARA^{†3} and KATASHI NAGAO^{†3}

Generally, researchers keep logs of research activities in laboratory notebooks. However, depending on situations, it might be difficult to keep research logs by text. On the other hand, voice logs are easy to keep but not easy to search. This paper proposes a system for annotating voice logs using laboratory notebooks. In the system, the user selects the title and tags for a voice log from those extracted from laboratory notebooks. Experimental results have shown the effectiveness of annotation using the system and the usability of annotated voice logs.

^{†1} 名古屋大学 工学部電気電子・情報工学科 School of Engineering, Nagoya University

^{†2} 名古屋大学 情報基盤センター Information Technology Center, Nagoya University

^{†3} 名古屋大学 大学院情報科学研究科 Graduate School of Information Science, Nagoya University

1. はじめに

研究者は、研究活動の経過や成果などを、備忘録のような簡単なメモや議事録などとして記録し、以降の活動で活用している。これらの多くはテキストで残されることが多く、その理由は閲覧性が高く検索が容易であるためである。一方で、たとえば着席していない状況などでは、テキスト入力よりも音声による記録の方が容易であるが、音声による記録（以下、音声ログ）の場合、探し出して利用するのが難しいという問題点がある。

音声ログの利用を容易にするには、記録時に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¹⁾が存在するが、手動による入力コストは大きい。また、音声認識を用いて音声ログを文字化する方法があるが²⁾、認識精度の問題や、利用する際に重要なフレーズを省略して発話することも多く、必ずしも十分な整理を行うことはできない。

そこで本稿では、音声ログの利点である記録の容易さを損なわずに、研究活動における音声ログの整理を目的として、テキストによる研究活動の記録（以降、研究ノート）から自動抽出したタイトルとタグを、ユーザーが容易に音声ログに付与するシステムを提案する。

2. 研究活動における音声ログ

2.1 音声ログの整理に関する従来法

音声ログの録音機器としては、IC レコーダーが一般的であり、主に会議やインタビューなどで利用される。また、スマートフォンには録音機能があり、その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^{3),4)} や録音アプリ¹⁾ において、ユーザーは音声ログにタイトルやタグなどを付与し整理することができる。しかし、タイトルやタグを手動で入力する方法では、ユーザーの作業コストが増大し、音声入力の利点が損なわれる。

一方、音声認識を利用した整理方法もある。たとえば、音素表記によるマッチングを行う検索方法²⁾ や、認識誤りに強い検索方法⁵⁾ がある。しかし、現在の自由発話の音声認識の精度は依然として高くない。また、たとえ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。内容が整理されていない音声ログが記録されることも多く、検索に有効な言葉が音声ログには含まれていないと、それを検索して探しだすことが難しくなる。すなわち、音声認識によって音声ログが文字化されるだけでは、音声ログの整理として不十分である。

2.2 研究ノートを利用した音声ログの整理

研究者は一般的に、デジタルかアナログかに関わらず、研究ノートの記録を日常的に行っ

ている．研究ノートに記録される内容は，TODO のリスト，新たな知見や事実，活動の経過，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して次の活動にフィードバックしている．研究活動における音声ログは，研究ノートとしての記録が難しい場面において使用され，その内容は研究ノートと同様，TODO や簡単なメモなどである．このため，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られる．

また研究活動は，特定のテーマを継続的に取り組むという点に活動の特徴があり，内容が大きく変わることは稀である．すなわち，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予想される．

そこで本研究では，研究ノートを利用し，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方法を提案する．提案手法により，手動で入力する場合と比べて，音声ログに付与されるタイトルの自由度は低下するものの，豊富なタグの提示によって多様なタグに基づく整理を実現できると考えられる．

3. 音声ログの整理手法

3.1 システムの概要

提案システムの構成を図 1 に示す．システムはノートサーバー，録音アプリ iRec，ボイスサーバーから構成される．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管理し，研究ノートが更新されると本文を形態素解析してタグを抽出し，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アプリケーションである．事前にノートサーバーからタイトルやタグを取得しておき，録音が完了するとユーザーに提示してタイトルとタグの付与を支援する．ボイスサーバーは，iRec からアップロードされた音声ログと，付与されたタイトル・タグ，録音日時・録音時間などのメタデータを保存・管理する．ユーザーは，Web ブラウザでボイスサーバーにアクセスすることにより，音声ログを視聴することができる．

3.2 研究ノートからのタグ抽出

3.2.1 研究ノートの作成

提案システムで用いる研究ノートの構成要素は，タイトル，本文，作成日時，更新日時であり，これらの要素をデータ構造として保持してノートサーバーにアップロードする機能を有していれば，研究ノートを作成するツールの種類は問わない．本研究では，iSticky⁶⁾

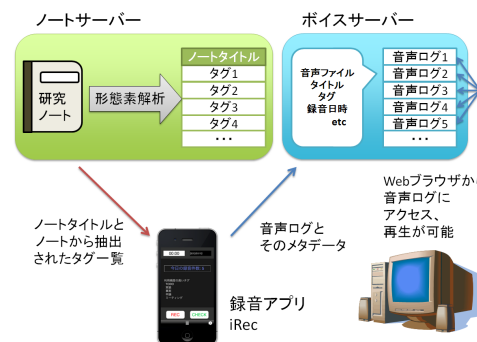


図 1 システムの構成

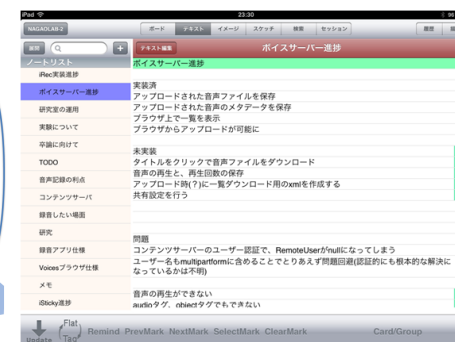


図 2 iSticky の研究ノート作成・閲覧インタフェース

とブラウザでのノート編集機能を利用する．iSticky 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．提案システムでは，iSticky の研究ノート作成機能のみを利用する．iSticky における研究ノートの作成・閲覧インタフェースを図 2 に示す．

3.2.2 タグの抽出方法

研究ノートの作成・更新に際し，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．本研究では，形態素解析器として lucene-gosen^{*1)} を利用する．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．タグの出現数もカウントしてデータベースに保存する．

3.3 音声ログへのタイトルとタグの付与

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は，簡単なメモから長時間に及び議論まで様々であり，それらが分類されていれば，音声ログの利用に有用である．本研究では，5 分以内の比較的短時間の音声ログを記録の対象とし，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声に出し，考えをまとめるために記録した音声ログ
- (2) 事実：行動や調査などにより分かったことなどを記録した音声ログ
- (3) TODO：実行し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

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

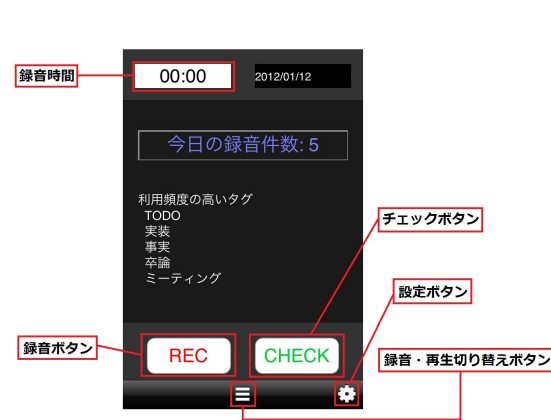


図 3 iRec の録音画面



図 4 タイトルの選択画面



図 5 カテゴリーの選択

3.3.2 録音アプリ iRec

iRec の録音画面を図 3 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として、タイトル、音声ログ用のタグ、およびその他のメタデータを取得する。iRec で録音ボタンをタップすると録音が始まり、再びタップすると録音を終了する。録音を終了すると、研究ノートのタイトル一覧が更新日時順に表示される (図 4)。

ユーザーが一覧の中からノートのタイトルを一つ選択すると、該当タイトルが音声ログに付与され、該当ノートから自動抽出されたタグの一覧が表示される (図 6)。タグ一覧の一番上には、アイデア思考、事実、TODO の 3 つのカテゴリの選択インターフェースが表示される (図 5)、ユーザーはいずれか 1 つのカテゴリを選択し、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を形態素解析することにより抽出されたタグの一覧であり、本文中の出現頻度が高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、新たに必要タグを作成すると、以後はユーザータグとして音声記録時に一覧表示され、容易に付与できるようになる。推薦タグ及びユーザータグは複数個選択でき、選択中のタグには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。タグの選択後に保存ボタンをタップすると、選択中のタグが、カテゴリを含めてす



図 6 タグの選択画面



図 7 iRec の再生画面

べ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 (図 7) には、記録された音声ログ一覧が表示される。一覧の中から任意の音声ログをタップすることにより音声ログを選択することができ、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。ボイスサーバーは、アップロードされた音声ファイル、タイトルとタグ、および録音日時などのメタデータを音声ログとして蓄積・管理する。

なお、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。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。

3.3.3 タイトルとタグの入力例

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 8 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。

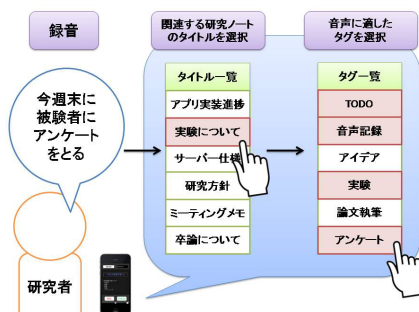


図 8 タイトルとタグの入力例

3.4 音声ログの視聴

Web ブラウザで音声ログを視聴するインタフェースを図 9 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する．前者は，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により一覧を絞り込むことができ，後者は，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を黄色くハイライト表示する．空白を入れることで，複数キーワードの指定もできる．各音声ログには，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．

4. 評価実験

提案システムを利用することにより，音声ログに対するタイトルとタグの付与を短時間でいえること，また，付与されたタイトルとタグ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．

4.1 音声ログ記録時のタイトル・タグ入力に関する実験

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有効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 16 日間である．被験者には，提案システムを用いてタイトルとタグを付与する方法と，タイトルを手動で入力し，さらに



図 9 音声ログの視聴インタフェース

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の 2 通りを一日ごとに交互に行ってもらった．被験者には，1 日 9 件以上の音声ログを記録してもらった．内訳は，アイデア思考を 1 件以上，事実を 4 件以上，TODO を 4 件以上である．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い，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を問い，回答および自由意見を録音・分析した．

4.1.2 実験結果と考察

タイトルとタグの平均入力時間を図 10 に示す．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の 60.8% の時間で入力を終えている．実験結果から，手動で入力する場合と比較してタイトルとタグの入力時間が短くなっており，提案システムが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，タイトルの手動入力に面倒という意見が 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方が考える手間が要らないというものであり，状況に応じて，手動入力に関連ノートの選択のどちらかを行える，あるいは両方とも行うことを許す仕組みを望む意見も得られた．そのため，手動入力と簡単な入力の両方を備えた仕組みを実現し，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．

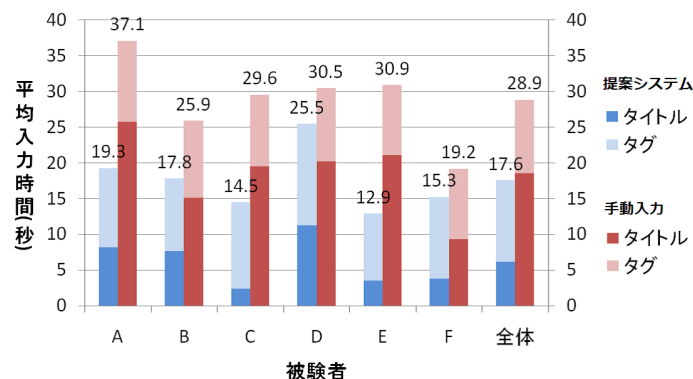


図 10 平均入力時間

また、提案システムを用いる場合、有限のノートの中から選択することによって、手動入力に比べてタイトルの多様性がないという根本的な問題が存在する。しかし、手動入力においても、同じようなタイトルを付けてしまうという意見が得られ、その理由として、記録時に手間をかけられない状況下においては、タイトルの自由な表現よりも短時間で入力を済ませたいという要求が強いことが分かった。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考えられる。

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という意見や、付与したいタグが存在しないなどの意見があった。その理由として、まず、提案システムによるタグの抽出方法が不十分であることが挙げられる。本研究では、選択したノートからのみタグを抽出したが、複数ノートからのタグの抽出や TF-IDF 法などによるタグの重み付け、ノートのクラスタリングによる選択ノートと関連性の高いノートを利用したタグの拡張などの工夫が考えられる。そのためには、ユーザーにとって必要なタグが何であり、どのようにして研究ノートから抽出すべきかについて、今回の実験で収集したデータをさらに細かく分析して検討する必要があると考えられる。

また、本研究では実験期間が短かったため、実験期間中に作成または更新された研究ノート数は平均 3.3 件と少なく、提案システムが有効なタグを抽出できなかった可能性がある。そのため、継続して運用を行っていく必要があると考える。

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、提案システムによって付与したタイトルとタグが、音声ログ利用の際に役立つかを検証することである。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 節の実験終了後、1 週間が経過した後に本実験を行った。被験者には、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ず、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た。

被験者を、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、それぞれのタイトルとタグの有効性を比較した。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。カテゴリーごとに問題を作成し、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 (表 1)。

表 1 グループ分けと問題設定

グループ	表示する音声ログ	カテゴリー別問題数		
		アイデア思考	事実	TODO
I (3 名)	提案システムで付与したタイトル・タグ	1	3	2
	タイトル・タグ非表示	-	2	-
II (3 名)	手動入力したタイトル・タグ	1	3	2
	タイトル・タグ非表示	-	2	-

アイデア思考に関しては、すべ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。事実に関しては、被験者ごとに問題を作成し、「○○を行ったのはいつですか」「○○が分かったのはいつですか」のように、一度だけ記録された事実を問う問題とした。TODO に関しては、「○月○日～○月○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？あればその内容を答えてください」という問題とした。なお、実験後に口頭アンケート調査を行い、「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、回答および自由意見を録音した。また、紙面アンケート調査では、事実の 3 問について、「役に立ったタグ」「抽出してほしいタグ」「検索したキーワード」について質問した。

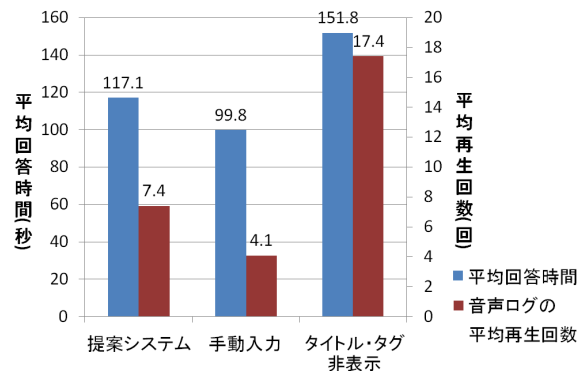


図 11 平均回答時間の比較

4.2.2 実験結果と考察

平均回答時間を比較した図 11 から、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た場合に比べて 1.17 倍の時間が掛かり、回答中の音声ログの再生回数も約 1.8 倍と多いことが分かる。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、視聴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、タグ付与することによって、タイトルとタグを入力しない場合に比べて回答時間を 30 秒以上短縮し、音声ログの再生回数も半数以下で済んでいる。タイトルを手動入力する場合と比べても、回答時間の差はそれほど大きくない。

4.2.3 考 察

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。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。しかし、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得られた。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、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。

タグの有効性

役に立ったタグの数は、グループ I では 9 個、グループ II では 4 個であり、提案システムを用いた場合にタグが重視されていることが分かる。さらにタグと発話を比較したところ、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、発話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、抽出してほしいタグについては、発話内容には含まれないタグが 4 個であるのに対し、発話内容に含まれるタグが 15 個であった。また、検索したキーワードも、発話内容に含まれるキーワードが 8 個に対し、発話内容に含まれないキーワードはわずか 1 個であった。以上のことから、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。

5. まとめと今後の課題

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。提案システムの有効性を検証するために被験者実験を行い、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、タイトルとタグの付与が音声ログの整理に役立つことを確認した。今後の課題としては、まず、継続的な運用による詳細な分析、タグ抽出手法・提示方法の改善、音声認識技術の併用が挙げられる。また、本研究では、短時間の音声ログを対象としたが、ミーティングなどの長時間の音声ログを整理する仕組みや、音声ログ以外の写真やビデオへの本手法の適用についても検討したいと考えている。

参 考 文 献

- 1) Apple: ボイスメモ. <http://www.apple.com/jp/iphone/built-in-apps/more-built-in-apps.html>.
- 2) Wechsler, M., Munteanu, E. and Schäuble, P.: New techniques for open-vocabulary spoken document retrieval, *Proc. of the 21st Annual international ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 20-27 (1998).
- 3) Olympus: VoiceTrek DM-4. <http://olympus-imaging.jp/product/audio/dm4/>.
- 4) SONY: ICD-BX312. <http://www.sony.jp/ic-recorder/products/ICD-BX312/>.
- 5) Renals, S., Abberley, D., Kirby, D. and Robinson, T.: Indexing and retrieval of broadcast news, *Speech Communication*, Vol.32, No.1-2, pp.5-20 (2000).
- 6) 高橋勲, 石戸谷頭太朗, 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).