

AISTハミングデータベース: 歌声研究用音楽データベース

後藤 真孝 西村 拓一

産業技術総合研究所

{m.goto@, taku@ni.}aist.go.jp

あらまし 本稿では、歌声、ハミングを対象とした研究の発展のために制作した歌声研究用音楽データベース「AISTハミングデータベース」(AIST-HDB)の制作方針と構成について述べる。RWC研究用音楽データベース(RWC-MDB)の普及に伴い、研究用に楽曲データが使用できるようになったが、初めて聴く楽曲を多数の被験者がうろ覚えの状態であつた大規模な音声データベースは存在しなかつた。そこでAIST-HDBでは、RWC-MDBから50曲100箇所を抜粋し、それを1回聴いた後と5回聴いた後に、思い出しながら歌う音声を収録した。100人が100箇所をハミング(歌詞なし歌唱)、歌唱、歌詞朗読した音声を含む多様な音声を、125.9時間分収録できた。

AIST Humming Database: Music Database for Singing Research

Masataka Goto Takuichi Nishimura

National Institute of Advanced Industrial Science and Technology (AIST)

Information Technology Research Institute, AIST, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

Abstract This paper describes the design policy and specifications of the *AIST Humming Database* (AIST-HDB), a music database that was built for singing and humming research. With the RWC Music Database (RWC-MDB), which has already been widely used, researchers can use copyright-cleared musical pieces for research. There has, however, been no large sound database that includes singing voices by many subjects who are not familiar with test pieces. The AIST-HDB was therefore built to record voices when each of 100 subjects sings 100 excerpts from 50 songs of the RWC-MDB after listening to each excerpt both once and five times. We made 125.9-hour recordings of various humming (singing without lyrics), singing, and reading voices.

1 はじめに

研究目的で共通に利用できる世界初の著作権処理済み音楽データベース(DB)「RWC (Real World Computing) 研究用音楽データベース」(以下、RWC-MDB)^{1)~4)}の登場以来、音楽情報処理分野においても、研究用DBの重要性が広く認知されるようになってきた。音声や画像等の他の研究分野では、古くから多様な共通DBが構築され、研究の進展に大きく貢献していた。一方、音楽情報処理の分野では、従来は共通楽曲DBが存在していない状況であったが、RWC-MDBの配布とその普及に伴い⁵⁾、各種手法の比較・評価や、DBに基づく学習、研究成果の対外発表等で、DBが活用される機会が増えてきた。RWC-MDBは、楽曲DBとして、ポピュラー、クラシック、ジャズ等の多様な音楽ジャンルの楽曲315曲、楽器音DBとし

て、ピアノやバイオリン等の約50種150個体の楽器の演奏音を収録している。そして、そのすべての楽曲に、コンパクトディスク(CD)に収録された音響信号、楽曲を可能な範囲で再現する標準MIDIファイル(SMF)、歌詞のテキストファイルが用意されている。

しかし、ハミングや歌唱を対象とした研究を進展させる上で、歌声に関するデータ収集はまだ不十分であった。従来の代表的な取り組みとして、日本語歌唱音声DB「日本語を歌・唄・謡う」⁵⁾が挙げられる。これは、邦楽(日本の伝統芸能)のほとんどすべてのジャンルと洋楽の、各ジャンルにおける最高クラスの演者72名が、同一の歌詞をそれぞれの典型的歌唱法で発声した音声を収録した試料集である。多様な「うたい分け」を比較研究することができる貴重なDBといえる。一方、メロディーをロずさむハミングに関しては、MAMI (Musical Audio Mining)プロジェクト⁶⁾によって、72人の被験者がハミングや歌唱で楽曲を検索する際の音声が入録された。各被験者が、30曲の有名楽曲の中から知っている10曲を

¹⁾ RWC-MDBを利用した具体的な文献リストや配布件数は、WWWサーチエンジン等で、「RWC研究用音楽データベースの利用報告」を含むページを検索すると閲覧できる。

表 1: 100人の被験者の構成 (歌唱力は自己申告)

母国語	日本語 75人				英語 25人			
性別	男性 37人		女性 38人		男性 14人		女性 11人	
年齢	20代 30人	30代 15人	40代 15人	50代 15人	20代 14人	30代 9人	40代 2人	50代 0人
音大	音大出身 6人		音大以外 69人		音大出身 3人		音大以外 22人	
歌唱力	上手 25人		普通 38人		下手 12人		上手 9人 普通 10人 下手 6人	

表 2: 100人の被験者に対するアンケート結果

母国語	日本語 75人			
音楽教育	学校教育以外で受けたことがある 43人		学校教育以外で受けたことがない 32人	
絶対音感	保持 5人		非保持 70人	
歌うことが好きか	非常に好き 12人	好き 34人	普通 24人	どちらかといえば嫌い 2人 嫌い 3人
歌に自信があるか	ある 2人	まあまあある 14人	普通 38人	そんなにない 13人 ない 8人

母国語	英語 25人			
音楽教育	学校教育以外で受けたことがある 13人		学校教育以外で受けたことがない 12人	
絶対音感	保持 1人		非保持 24人	
歌うことが好きか	非常に好き 10人	好き 6人	普通 2人	どちらかといえば嫌い 4人 嫌い 3人
歌に自信があるか	ある 5人	まあまあある 3人	普通 7人	そんなにない 4人 ない 6人

選び、タイトルとアーティスト名だけを見ながら自由に歌った音声収録された。さらに、知らないもしくは思い出せない楽曲を聴いた後に、ハミングや歌唱、口笛でメロディーを再現する音声も収録された。また、別のハミング検索プロジェクト⁷⁾でも、100人の被験者が平均約25曲をハミングした音声収録された。各被験者は、22曲の有名楽曲の中から20曲程度を、MIDIファイルの演奏を聴いた後にハミングし、さらに、それ以外に自由に選んだ曲を5~6曲ハミングした。しかし、上記のような取り組みはあったものの、初めて聴くポピュラー音楽の楽曲の同一箇所を、多数の被験者がうろ覚えの状態でもハミング・歌唱した大規模な音声DBは構築されていなかった。

そこで我々は、ハミング・歌唱を対象とした研究を一層進展させることを目的とした歌声研究用音楽DB「AISTハミングデータベース」(AISTは産業技術総合研究所の英語略称)を構築した。これは、RWC-MDB 50曲の抜粋を、100人の被験者がハミング、歌唱、歌詞朗読した音声を収録したものである。以下、被験者、対象楽曲、収録手順等の詳細を述べ、収録した計125.9時間分の音声データの全体構成を紹介する。

2 AISTハミングデータベース

本DBでは、初めて聴くポピュラー音楽を、うろ覚えの状態でも歌った音声を収録することを目指す。こうした歌声を収録するには様々な方法が考えられ、歌唱者の技量、対象楽曲、音楽ジャンル、収録内容等を検討する必要がある。本DBでは、ある歌唱者が多様な楽曲をどう歌うか、ある楽曲を多様な歌唱者がどう歌うかという調査ができることを重視し、特別な歌唱訓練を受けていない一般的な歌唱者を中心に100人集

め、RWC-MDBの楽曲から抜粋した100箇所の音響信号を初めて聴いた後に歌う音声を収録する。これにより、ある一人が100箇所を、ある1箇所を100人がどう歌うかがわかる。既に普及しているRWC-MDBと対応付けられる点で意義があるだけでなく、歌われるメロディーの作曲に関する著作権が処理済みであるため、研究目的で自由に使える利点もある。本DBではポピュラー音楽の範囲を広く捉え、J-POP、ロック、ヘビーメタル、ファンク、フォーク等の楽曲を混ぜて用いることとする。そして、ハミング(歌詞なし歌唱)と歌唱に加え、同じ歌詞を読み上げた音声も収録することで、歌声と朗読音声の比較を可能にする。

2.1 被験者

100人の被験者の構成を表1に示す。これらの被験者は、以下のように、可能な範囲でバリエーションが豊かとなるように集めた。まず、日本語と英語の歌詞を、それぞれを母国語とする被験者が歌唱した音声を収録するために、日本語を母国語とする75人の日本人と英語を母国語とする25人の外国人で構成した。男女比がほぼ等しくなるようにした結果、男性51名、女性49名となった。年齢は、20代、30代、40代、50代が2:1:1:1の比率を目指したところ、それぞれ44人、24人、17人、15人となった。また、今回は特別な音楽的訓練を受けていない人の歌声を主な対象と考え、音大出身者は9人と全体の約1割に抑えるようにした。歌唱力は、収録前の自己申告によれば、上手34人、普通48人、下手18人であった。

収録前に各被験者が記入したアンケートの回答の内訳を表2に示す。楽器練習等の何らかの音楽教育を学校教育以外で受けたことがある被験者は56人と

5割強であったが、絶対音感保持者は6人と少なかった。歌うことが「(どちらかといえば)嫌い」な被験者は12人、歌に自信が「(そんなに)ない」被験者は31人と、必ずしも歌が好きで自信のある被験者ばかりではなく、一般的な被験者が集められていた。

なお、すべての被験者は、収録音声に関する権利を放棄し、音声の研究目的で公開(ただし個人名は非公開)されることを事前に了承している。

2.2 対象楽曲

RWC-MDBの全315曲の中から、新規楽曲と有名楽曲の二つのカテゴリに分けて計73曲を選曲した。

2.2.1 新規楽曲 50曲 100サンプル

被験者が歌う対象の楽曲として、日本語歌詞と英語歌詞の曲数が同数となるように、「RWC研究用音楽データベース:ポピュラー音楽(RWC-MDB-P-2001)」(全100曲)^{1),2),4)}から35曲(日本語歌詞25曲と英語歌詞10曲)、「RWC研究用音楽データベース:音楽ジャンル(RWC-MDB-G-2001)」(全100曲)^{1),3),4)}から15曲(英語歌詞15曲)を抜粋した。具体的な楽曲番号を表3に示す。これら50曲はすべて、RWC-MDBのために新規に作曲、作詞、編曲、演奏、歌唱、録音されたものである(以下、**新規楽曲**と呼ぶ)。したがって、被験者はこれらを日常生活で聴いたことがなく、初めて聴くポピュラー音楽の楽曲を、どのように歌うかを調べることができる。これらは、多様な曲調、テンポの楽曲を含み、歌唱者の性別が偏らないように選んだ。RWC-MDB-P-2001からの抜粋では、男性ボーカル(単独)17曲、女性ボーカル13曲、男性グループ(複数)1曲、女性グループ4曲であり、RWC-MDB-G-2001からの抜粋では、男性ボーカル6曲、女性ボーカル9曲である。

そして、これら50曲の各曲について、

- **Aメロ**(楽曲中の歌唱の出だし部分)
- **サビ**(楽曲中で一番代表的な盛り上がる主題の部分)

の2箇所(2サンプル)を切り出した。その全100サンプルの切り出し区間長の分布を図1に示す。この切り出し区間は、歌いやすいように、その部分の歌詞の始まりから区切りの良いところまでとし、平均10.2秒であった。

2.2.2 有名楽曲 23曲 23サンプル

以上は、すべての被験者が歌う必須の楽曲であるが、予定より早く収録が終了した被験者は、上記の新規楽曲50曲とは別に、有名な楽曲23曲の一部を時間

表3: 使用した新規楽曲50曲の楽曲番号

	RWC-MDB 楽曲番号
日本語歌詞25曲	RWC-MDB-P-2001 No. 1, 5, 8, 10, 13, 18, 22, 23, 24, 25, 27, 28, 33, 36, 37, 39, 41, 43, 46, 59, 66, 69, 71, 78, 79
英語歌詞25曲	RWC-MDB-P-2001 No. 82, 83, 87, 90, 91, 92, 93, 95, 97, 100 RWC-MDB-G-2001 No. 2, 4, 6, 7, 10, 13, 17, 22, 24, 25, 27, 39, 41, 67, 69

表4: 使用した有名楽曲23曲の楽曲番号

	RWC-MDB 楽曲番号
日本語歌詞5曲	RWC-MDB-R-2001 No. 1, 2, 3, 4, 5
英語歌詞10曲	RWC-MDB-R-2001 No. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
歌なし5曲	RWC-MDB-G-2001 No. 54, 55, 56, 57, 60
イタリア語歌詞3曲	RWC-MDB-G-2001 No. 88, 89, 90

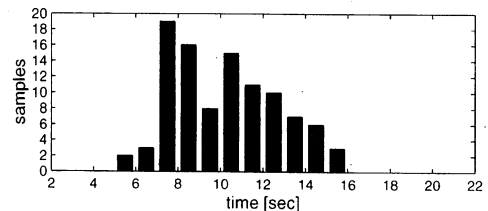


図1: RWC-MDBの50曲から切り出した100サンプルの長さの分布

が許す限り歌うこととした。「RWC研究用音楽データベース:著作権切れ音楽(RWC-MDB-R-2001)」(全15曲)^{1),2),4)}から15曲(日本語歌詞5曲と英語歌詞10曲)、RWC-MDB-G-2001から8曲(歌なし5曲とイタリア語歌詞3曲)を抜粋した。具体的な楽曲番号を表4に示す。これら23曲は、著作財産権の保護期間が終了した有名既存曲(原則として著作者の死後50年以上を経過した楽曲)をRWC-MDBのために新規に演奏、歌唱、録音したものである(以下、**有名楽曲**と呼ぶ)。したがって、被験者はこれらを日常生活で聴いたことがある可能性が高く、そうした曲をどのように歌うかを調べることができる。これらは代表的な1箇所(1サンプル)だけ切り出し、意図的に長いものを混ぜるようにしたところ、平均15.1秒であった。

2.3 計画した収録内容

以下の三種類の音声を収録する。

- **ハミング**(歌詞なし歌唱)
チャーチャーチャー、ターターター、ラーラーラーのような任意の音で、原曲の歌詞は付けずにメロ

ディーをロズさんだ音声収録する。本稿での「ハミング」という用語は、「ハミング」の本来の定義である唇を閉じたまま声を鼻に抜かしてメロディーを歌う歌唱よりも、広い意味で用いている。

- **歌唱** (歌詞つき歌唱)

原曲の歌詞でメロディーを歌った音声収録する。

- **歌詞朗読**

メロディーの音高を無視して、原曲の歌詞を見ながら読み上げた音声収録する。

ただし、歌唱と歌詞朗読は、以下のように母国語と歌詞の言語が一致しているもののみを対象とする。

- **日本語を母国語とする75人**

日本語歌詞25曲は、ハミング、歌唱、歌詞朗読を収録し、英語歌詞25曲は、ハミングのみ収録する。

- **英語を母国語とする25人**

日本語歌詞25曲は、ハミングのみ収録し、英語歌詞25曲は、ハミング、歌唱、歌詞朗読を収録する。

ハミングと歌唱に関しては、一度しか聴いたことがないフレーズを歌ったときに、うろ覚えでどの程度歌えるか、何度も聴いた後に歌ったときにどの程度歌えるかを調べるために、各曲から切り出された2サンプル(Aメロとサビ)のそれぞれにつき、以下の二種類の楽曲聴取回数で収録する。

- 1回聴いた後に収録
- 5回聴いた後に収録(上記の後、さらに4回聴いて収録)

以上のように、新規楽曲に関しては被験者一人当たり、

- ハミング 200サンプル(50曲×2サンプル×2種類)
- 歌唱 100サンプル(25曲×2サンプル×2種類)
- 歌詞朗読 50サンプル(25曲×2サンプル)

の計350サンプルを収録し、被験者100人で計35000サンプルとなるよう計画した。有名楽曲に関しては、被験者が歌うことが可能な曲に限定して歌唱と歌詞朗読を収録した。

2.4 収録環境・機材

防音室(カワイ音響システム ナサールSVU 遮音性能D-30(30dB/500Hz))内で、ピンマイク(audio-technica AT899)とスタンドマイク(SONY F-780)の二本のマイクで同時に収録した。これらのマイクは、XLRバランス入力でマイクアンプ兼A/D変換器(M-AUDIO DUO)の左と右にそれぞれ接続し、そのデジタルオーディオ(S/PDIF)出力をパーソナルコンピュータ(Intel Pentium 4 3.2 GHz CPU、Windows

XP)のサウンドカード(Cobra AW850 Deluxe)に同軸ケーブルで入力した。マイクの信号は、市販の録音用ソフトウェア(Sonic Foundry Sound Forge 4.5)によって、サンプリング周波数44.1kHz、16ビットリニア量子化、ステレオ(2チャンネル)で収録中は録音し続け、後で音声区間の切り出しをおこなった。

防音室内には、マイク以外に、椅子、机、その机の上にスピーカ、パソコンのディスプレイを設置した。各被験者は椅子に座り、スピーカから流れる音楽と音声指示を聴きながら、ディスプレイ上の画面指示を見て発声する。歌唱の妨げにならないように、ヘッドホンは使用せずにスピーカからすべての音を流した。画面には、何曲目の収録か、楽曲中の場所(Aメロ、サビ)、歌い方(ハミング、歌唱、歌詞朗読)、楽曲聴取回数を常に表示し、被験者が収録状況を把握できるようにした。ただし、被験者が集中できるように、楽曲聴取中と歌唱中は表示内容を変更しなかった。

2.5 収録手順

各被験者が、新規楽曲50曲の原曲の音響信号(伴奏付き歌唱)を順に聴きながら歌った音声収録した。その曲順は、被験者ごとに異なるようにした。各曲での収録手順を以下に示す。

- 歌詞が母国語の場合(50曲中の25曲)
 1. Aメロを1回聴いて、その後に鳴るビーブ音(歌い始めの合図音)の直後にハミングをする。それが終わると歌詞が画面に表示されるので、それを見ながら歌詞付きで歌う。なお、ビーブ音と歌詞は毎回提示されるが、以下の記述では省略する。
 2. Aメロをさらに4回(計5回)聴いた後に、ハミングをし、歌詞付きで歌う。
 3. サビを1回聴いた後に、ハミングをし、歌詞付きで歌う。
 4. サビをさらに4回聴いた後に、ハミングをし、歌詞付きで歌う。
 5. 画面に表示されたAメロとサビの歌詞を見ながら、歌わずに普通に読み上げる。
- 歌詞が母国語でない場合(50曲中の25曲)
 1. Aメロを1回聴いた後に、ハミングをする。
 2. Aメロをさらに4回聴いた後に、ハミングをする。
 3. サビを1回聴いた後に、ハミングをする。
 4. サビをさらに4回聴いた後に、ハミングをする。

ただし、1回聴いただけでは全く歌えない場合、もしくは、ほとんど歌えなかったり、歌以外の音が入って

表 5: 収録結果 (ファイル数, サイズはマイク一本分の集計値で, 実際にはマイク二本分あるため二倍の値となる)

		ファイル数		平均時間		合計時間	ファイルサイズ
		正常収録	失敗収録	正常収録	失敗収録		
新規楽曲	ハミング	20455	2921	11.7 秒	10.9 秒	75時間 24分 48.8 秒	22.3 Gbytes
	歌唱	9997	413	12.0 秒	12.4 秒	34時間 49分 22.3 秒	10.3 Gbytes
	歌詞朗読	4999	1	7.0 秒	7.9 秒	9時間 41分 19.4 秒	2.9 Gbytes
	小計	38786		平均 11.1 秒		119時間 55分 30.5 秒	35.5 Gbytes
有名楽曲	ハミング	846	66	15.5 秒	15.9 秒	3時間 55分 37.8 秒	1.2 Gbytes
	歌唱	369	8	15.2 秒	13.3 秒	1時間 35分 16.9 秒	0.5 Gbytes
	歌詞朗読	184		8.5 秒		0時間 26分 7.2 秒	0.1 Gbytes
	小計	1473		平均 14.5 秒		5時間 57分 1.9 秒	1.8 Gbytes
総計		40259		11.3 秒		125時間 52分 32.4 秒	37.2 Gbytes

しまったりして失敗した場合には, 7回までを上限として, 再度, 原曲の音響信号を聴くことができることとした. A メロを n 回聴いて初めて歌えた場合, その回数を記録して第一回目の音声収録をした上で, 次は A メロをさらに $(5 - n)$ 回聴いて計 5 回聴いた後に, 第二回目の音声収録をした (以下, この二回の音声収録を正常収録と呼ぶ). ただし, n が 5 から 6 だったときは, 計 $(n + 1)$ 回聴いて第二回目の音声収録をした. 第一回目の音声収録では歌詞の誤りは不問としたが, 第二回目では失敗として再収録した. また, 上記のそれぞれの失敗した場合の収録音も貴重な研究対象であるので保存しておいた (以下, 失敗収録と呼ぶ). 被験者は, 各曲の収録が終わる度に, その曲の収録を振り返ってアンケートに答えた.

なお, 上記の収録に先だって, 収録全体の構成と手順を被験者に説明した後に, 対象楽曲とは異なるサンプル曲を流して音楽再生の音量調整をし, 被験者が好きな歌を歌ったときの音声でマイクの音量調整をした. 被験者には以下の点を指示した.

- 移調して歌いやすい音高で歌ってよい.
- ハミング (歌詞なし歌唱) の場合には, 歌詞以外のどんな音で歌ってもよく, 歌い方を曲の途中で変えてもよいし, 曲ごとに変えてもよい. 例えば, チャーチャーチャー, ターターター, ラーラーラーのような任意の音で歌ってよい (ただし, 歌い方が偏る影響を減らすために, これを指示するときの「チャー」「ター」「ラー」の 3 種類の例示の順は, 被験者ごとにランダムに変えた).
- 歌詞つき歌唱の場合には, 正しい歌詞を歌っている限り, 原曲の歌手の歌い方を真似てもよいし真似なくてもよい.
- 適切なマイク収録のために, 歌唱中の姿勢は途中で変えず, 拍子を取ったり音を立てたりしない. 歌うとき以外は姿勢を変えたり発声してもよい.
- 歌唱中に目を閉じるかどうかは自由とする.

予定より早く新規楽曲の収録が終了した場合には, 有名楽曲 23 曲の中からランダムな順番で選曲して, 時

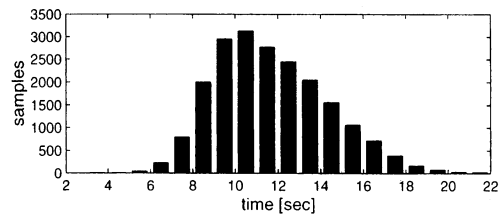


図 2: 新規楽曲のハミングで正常収録できた 20455 ファイルの長さの分布

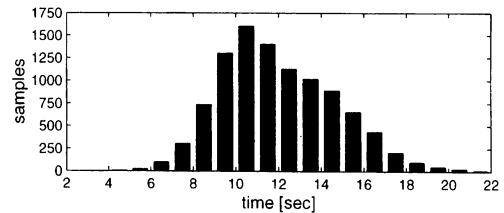


図 3: 新規楽曲の歌唱で正常収録できた 9997 ファイルの長さの分布

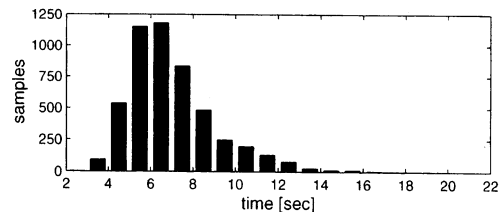


図 4: 新規楽曲の歌詞朗読で正常収録できた 4999 ファイルの長さの分布

間が許す限り同様の収録をおこなった.

2.6 収録結果

切り出した各サンプルは, ピンマイクとスタンドマイクのファイルは別に分けて, サンプリング周波数 44.1kHz, 16ビットリニア量子化, モノラル (1チャンネル) の RIFF WAVE フォーマットの音響ファイルで収録した. 収録したファイル数, 各ファイルの時間長の平均, 収録時間の合計, ファイルサイズを表 5 に示す. マイク一本分の集計値で, ファイル数

表 6: 新規楽曲のハミング, 歌唱, 歌詞朗読の計画したサンプル数のうち正常収録された回数

正常収録回数	0	1	2	3	4	5	6	7	計	総ファイル数
ハミングのサンプル数	7	8	9564	372	44	3	2	0	10000	20455
歌唱のサンプル数	2	2	4993	3	0	0	0	0	5000	9997
歌詞朗読のサンプル数	1	4999							5000	4999

表 7: 新規楽曲のハミング, 歌唱で何回目に正常収録できたか

楽曲聴取回数	1	2	3	4	5	6	7	計
ファイル数	11124	2854	947	247	14309	827	144	30452

の総計は 40259 ファイルであり, ファイルサイズの総計は 37.2 Gbytes, その総時間は約 125.9 時間となった(したがって, マイク二本分で 74.4 Gbytesであった)。新規楽曲における全収録の中の失敗収録の割合は, ファイル数で 8.60%, 合計時間とファイルサイズで 8.58% であった。一方, 有名楽曲における全収録の中の失敗収録の割合は, ファイル数で 5.02%, 合計時間とファイルサイズで 5.39% と, 新規楽曲の場合よりも少なかった。

被験者一人当りの拘束時間は総計 7.5 時間(休憩 1.5 時間分を含む)で, 喉への負担を考慮し, 2 日間もしくは 3 日間に渡って収録することとした。各曲収録後のアンケートの集計結果を図 5 に示す。

2.3 節で述べたように, 新規楽曲は被験者 100 人で計 35000 サンプル(15000 サンプル × 2 回分 + 5000 サンプル × 1 回分)収録する計画であったが, 正常収録で総計 35451 ファイル収録できた。計画より多いのは, 正常収録の場合でも余分に収録したことがあったためだが, その一方で, 上限 7 回の楽曲聴取回数では正常収録が 2 回できなかった事例もあった。そこで, 実際に何回正常収録されたかを調べた結果を表 6 に示す。ハミング 10000 サンプル, 歌唱 5000 サンプル中, 計画通り 2 回分収録できなかったのはそれぞれ 15 サンプル, 4 サンプルだった。歌詞朗読 5000 サンプル中, 計画通り 1 回分収録できなかったのは 1 サンプルだった(他に日本人が母国語でない英語歌詞を朗読した計画外収録が 1 サンプルあった)。また, 新規楽曲のハミング, 歌唱で何回目に正常収録できたかを調べた結果を表 7 に示す。上記に加えて, 有名楽曲も正常収録で総計 1399 サンプル収録できた^{☆2}。一人当りの有名楽曲の平均収録曲数は 4.2 曲であった。

3 おわりに

本稿では, 125.9 時間分のハミング(歌詞なし歌唱), 歌唱, 歌詞朗読を収録した大規模な歌声研究用音楽 DB「AIST ハミングデータベース」を構築したことを報告した。本 DB は, ハミング検索だけでなく, 歌

^{☆2} イタリア語の歌唱, 歌詞朗読は, 母国語話者に限定しなかった。

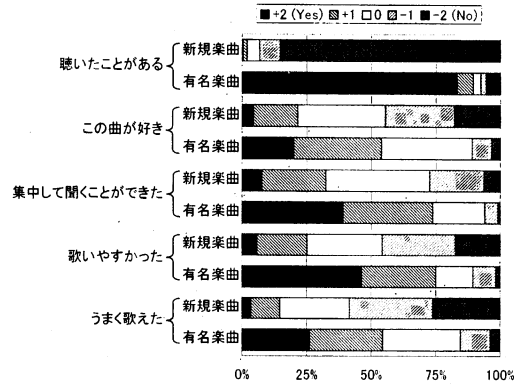


図 5: 各曲収録後のアンケートの集計結果 (100 人分)

声の歌詞認識, 歌手識別, 歌唱分析, メロディーの記憶力の分析, 歌声と朗読音声の識別等の多様な研究で有用である。今後は様々な分析やラベル付け(アノテーション)を実施していく予定である。

謝辞

本収録では, RWC 研究用音楽データベース(ポピュラー音楽 RWC-MDB-P-2001 と音楽ジャンル RWC-MDB-G-2001)を使用した。同 DB の企画, 構築に関わった関係各位に感謝する。また, 本 AIST ハミングデータベースの音声の収録を担当して頂いた株式会社アイアール・アルト(山本修氏, 金子満美子氏)に感謝する。

参考文献

- 後藤真孝, 橋口博樹, 西村拓一, 岡隆一: RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース, 情報学論, 45, 3, 728-738 (2004).
- Goto, M., Hashiguchi, H., Nishimura, T. and Oka, R.: RWC Music Database: Popular, Classical, and Jazz Music Databases, *Proc. of ISMIR 2002*, 287-288 (2002).
- Goto, M., Hashiguchi, H., Nishimura, T. and Oka, R.: RWC Music Database: Music Genre Database and Musical Instrument Sound Database, *Proc. of ISMIR 2003*, 229-230 (2003).
- Goto, M.: Development of the RWC Music Database, *Proc. of ICA 2004*, 1-553-556 (2004).
- 中山一郎: 日本語を歌・唄・謡う — 共通の歌詞をうたい分けた音声試料の紹介 —, 情処研報 音楽情報科学 2001-MUS-39-1, 1-4 (2001).
- Lesaffre, M. et al.: The MAMI Query-By-Voice Experiment: Collecting and annotating vocal queries for music information retrieval, *Proc. of ISMIR 2003*, 65-71 (2003).
- Unal, E. et al.: Creating Data Resources for Designing User-centric Frontends for Query by Humming Systems, *Proc. of MIR 2003*, 116-121 (2003).