

平成29年度

戦略的基盤技術高度化・連携支援事業

戦略的基盤技術高度化支援事業

「話者の音声特性を改善する難聴者聴こえ支援会話システムの開発」

研究開発成果等報告書

平成30年3月

担当局 関東経済産業局

補助事業者 ユニバーサル・サウンドデザイン株式会社

目 次

第1章 研究開発の概要.....	- 3 -
1-1 研究開発の背景・研究目的及び目標.....	- 3 -
1-2 研究体制.....	- 5 -
1-3 成果概要.....	- 5 -
1-3-1 研究開発の内容と技術的目標.....	- 5 -
1-3-2 目標の達成度と研究成果.....	- 6 -
1-4 当該研究開発の連絡窓口	- 7 -
第2章 本論－（１）話者音声改善技術の開発 -音声明瞭度の改善-	- 8 -
2-1 概要.....	- 8 -
2-2 音声特性モデルの作成.....	- 9 -
2-2-1 音声の収録.....	- 9 -
2-2-2 音声分析方法.....	- 10 -
2-2-3 話者音声特性の比較と理想音声モデルの作成.....	- 10 -
2-3 補聴信号処理方法の開発.....	- 15 -
2-4 開発した信号処理の評価.....	- 17 -
2-4-1 語音弁別検査による明瞭度評価.....	- 17 -
2-4-2 語音弁別検査アプリ開発による明瞭度評価の効率化.....	- 17 -
2-4-3 語音弁別検査の条件と明瞭度の評価結果.....	- 19 -
第3章 本論－（２）話者音声認識技術の開発 -ユーザビリティの改善-	- 23 -
3-1 概要.....	- 23 -
3-2 短時間での学習が可能な話者音声認識技術の開発（学習用文章の短文化）	- 23 -
3-3 話者認識アプリの実装と話者認識結果の検証.....	- 25 -
3-4 操作用タブレットアプリの開発によるユーザビリティ改善.....	- 26 -
最終章 全体総括.....	- 27 -
4-1 研究履行状況について.....	- 27 -
4-1-1 当初の事業計画履行状況.....	- 27 -
4-1-2 目標達成度.....	- 27 -
4-2 事業化の展開と今後の課題.....	- 27 -
4-2-1 事業化（１）集音器タイプ補聴機器 comuoon pocket.....	- 27 -
4-2-2 事業化（２）医療機器対応卓上会話支援機器と認知症リハビリテーションプログラム事業.....	- 29 -
4-2-3 事業化（３）複数人対応 会話支援システム（会議システム）	- 30 -

第1章 研究開発の概要

1-1 研究開発の背景・研究目的及び目標

【中小ものづくり高度化法及び中小ものづくり高度化指針】

(二) 情報処理に係る技術に関する事項

1 情報処理に係る技術において達成すべき高度化目標

(3) 川下分野横断的な共通の事項

①川下製造業者等の共通の課題及びニーズ

カ. 製品・サービスのユーザビリティ向上

従来、補聴のための福祉機器としては、聞き手（難聴者自身）が個別に身に着けるものが殆どであり、その補聴効果も十分ではない。そのため補聴器等の利用率は2割以下と多くの難聴者が聴こえない状態で生活をしており、コミュニケーション障害による引きこもり、病院や薬局での不十分なインフォームドコンセントが問題となっている。そこで、話し手側から発生する音声を聞き手が聞き取りやすい音声へと変換するための「聴こえ支援スピーカ」を着想し、本事業着手前に聴こえ支援スピーカの研究開発及び製品化（comuoonシリーズ）を行った。本事業ではこの聴こえ支援スピーカの更なる高性能化及びユーザビリティ向上に取り組んだ。

難聴者にとって、聞き取りやすい話者の声と聞き取りにくい話者の声の違いは健聴者より顕著であり、複数人での会話時に聞き取りにくい話者がいると、補聴器などをしていても会話の内容を把握することが難しいことがある。従来の聴取補助機器は主に難聴者の聴力特性（オーディオグラムなど）のみを考慮した補聴処理であり、話者の声の特性は考慮されていない。前述の聴こえ支援スピーカ comuoon においても、話者により聞き取りにくい場合があることが確認されており、特に高度難聴者以上の聴こえの改善が不足していた。そこで本事業では、話者の音声特性を把握して、より聞き取りやすい音声に変換する話者音声改善信号処理技術を開発し、聴こえの改善を行った（研究要素1）。聞き取りやすい音声特性を「理想音声モデル」として設定し、話者の音声特性を理想音声モデルに近づける信号処理方法を開発し、当初の目標であった高度難聴者の明瞭度50%以上（日常会話が成り立つ聴こえ）を達成した。

研究要素1で開発した信号処理技術は、話者ごとに信号処理の設定を変更しなければならないため、複数の話者と会話する場合、ユーザーが手動で操作しなければいけないとユーザビリティの低下を招く。そこで、話者の音声特性を自動認識する機能を搭載することにした。ただし、従来の話者認識では長時間の事前学習が必要である問題があったため、短時間で学習が可能な話者認識技術の開発および事前学習を支援するタブレットアプリの開発を行った（研究要素2）。事前学習では学習用文章を音読して学習させるため、学習用文章の短文化による時間短縮を図った。日常会話における短音節の出現頻度を考慮した音読時間1分以内の学習用文章を作成することで、短時間の学習でも93%以上の高い精度で話者を認識することが可能であることを確認し、当初目標である音読時間1分以内の学習用文章の作成を完了した。また、音声学習作業をスムーズに実行するための操作用タブレットアプリを開発し、ユーザビリティの改善を行った。

上記の研究要素1、2の研究開発促進と、開発したシステムの効果を確認するため、難聴者による音声評価技術の開発・評価と聞き取りやすい音声の物理的要因の解析をおこなった（研究要素3）。音声評価技術の開発では、補聴器のフィッティング効果を確認するために使用される語音弁別検査を参考にして音声明瞭度を評価する方法を開発するとともに、タブレットにて効率よく検査及び自動分析が可能なアプリを開発した。明瞭度の評価により聴こえ改善の効果を定量的に確認すると共に、改善不足である語音が明確に分析可能となった。また明瞭度評価の結果と、読み上げ文章および単音節の詳細な周波数分析結果により、聞き取りやすい音声と聞き取りにくい音声の物理特性の違いを明確にし、研究要素1の理想音声モデルおよび話者音声改善信号処理の開発に役立てた。これらの難聴者向け音声評価技術の開発により、研究要素1、2の開発を促進し、全体の目標達成に貢献した。

上記のように、話者音声の明瞭度改善およびユーザビリティの改善に取り組み、話者音声・聴力補正による高度難聴者までの聴こえを改善する補聴技術の開発を達成した。

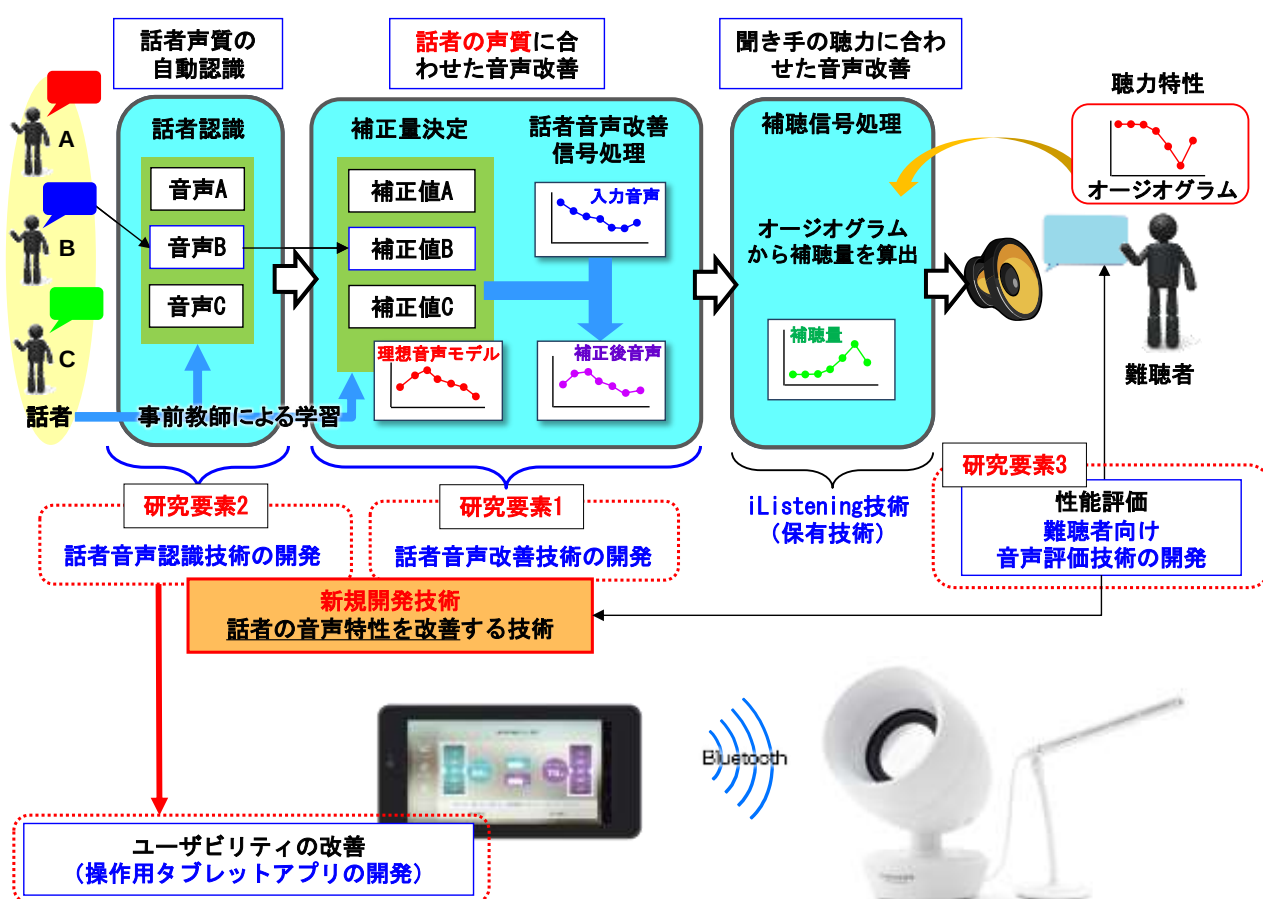


図 1-1 開発するシステムと研究要素

1-2 研究体制

本事業の研究開発は、図 1-2 に示す事業体制にて実施した。

研究開発共同体

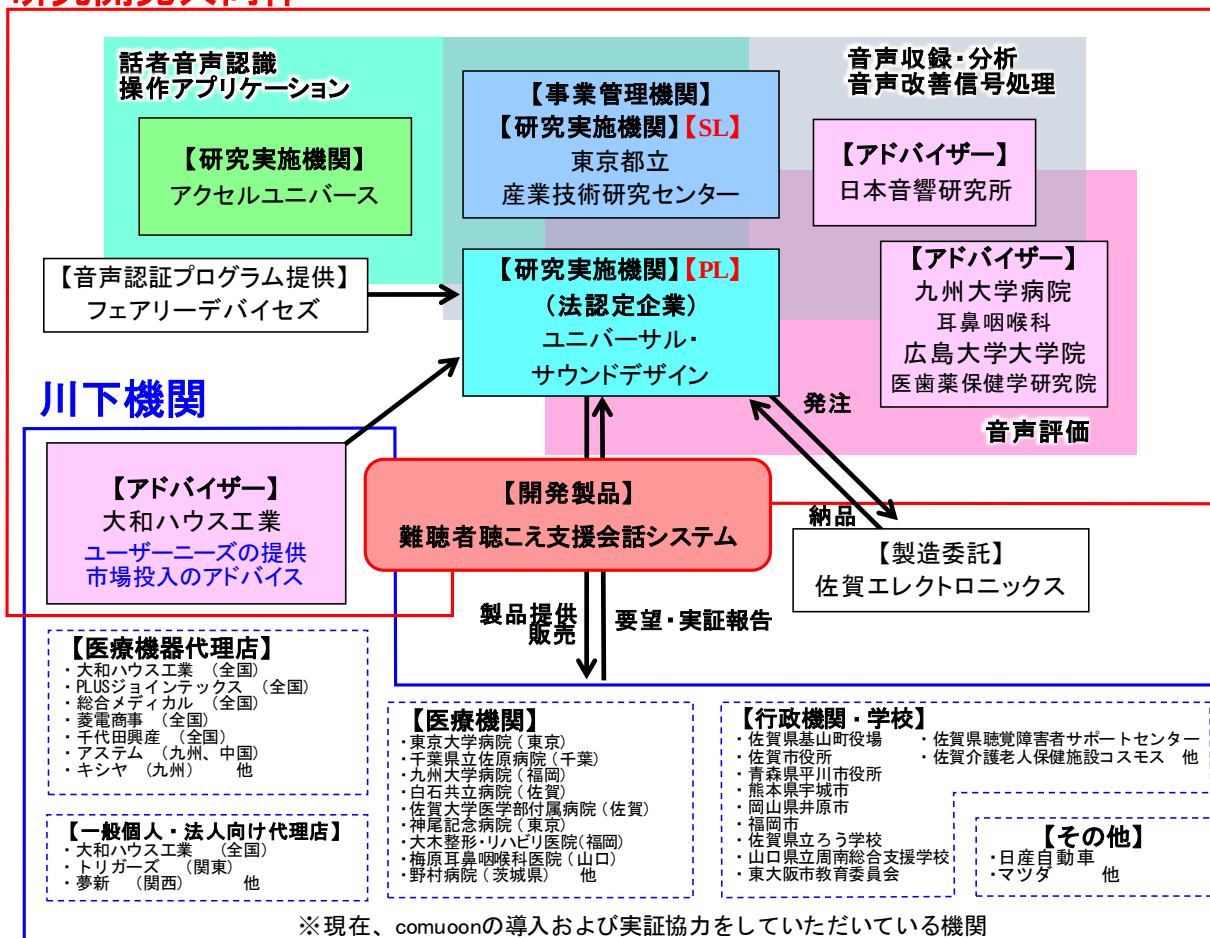


図 1-2 事業体制図

1-3 成果概要

1-3-1 研究開発の内容と技術的目標

【研究要素 1】 話者音声改善技術の開発

- ・話者の音声に合わせた音声改善信号処理手法の開発を次の【1-1】，【1-2】の2点により実現する
- ・達成目標として、【3-1】の語音弁別検査による高度難聴者の音声明瞭度 50%以上を目指す

【1-1】 音声特性モデルの作成

- ・話者音声を分析し、聴き取りやすい音声特性のモデル（理想音声モデル）を作成

【1-2】 補聴信号処理方法の開発

- ・話者の音声特性を【1-1】理想音声モデルに変換するための補聴信号処理方法を開発・実装

【研究要素 2】 話者音声認識技術の開発

- ・話者声質を自動認識する機能を追加することによるユーザビリティ改善

【2-1】 短時間での学習が可能な教師あり学習による話者音声認識技術の開発

- ・話者一人あたりの事前学習時間が1分以内の話者音声認識技術を開発

【2-2】 操作用タブレットアプリの開発によるユーザビリティ改善

- ・話者認識に必要な事前学習をタブレットにて容易かつ適正に実行できる操作用アプリを開発

【研究要素3】 難聴者向け音声評価技術の開発

- ・医学分野、工学分野の両面から音声の評価および聴こえやすさに起因する物理要因の解析を行い、研究要素1,2の信号処理開発にフィードバックする

【3-1】 難聴者による音声評価（医学分野）

- ・開発システムの音声明瞭度の評価が可能、難聴者による語音弁別検査を設計・実施
- ・明瞭度評価結果と【3-2】の解析結果により聴こえやすさの物理要因を分析

【3-2】 音声の物理特性解析（工学分野）

- ・音声の物理特性解析（語音の周波数分析）により聴こえやすさの物理要因の指標を見出す

1-3-2 目標の達成度と研究成果

研究課題	研究目標	研究目標に対する研究達成度と成果	
		達成率	成果
研究要素1 話者の音声改善技術の開発 【1-1】 音声特性モデルの作成 【1-2】 補聴信号処理方法の開発	・高度難聴者の音声明瞭度 50%以上	100%	・聞き取りやすい理想音声モデルおよび補聴信号処理方法を開発・実装 ・【3-1】の明瞭度評価により、複数の高度、中等度難聴者の音声明瞭度 50%以上を達成。補聴器を併用した場合でも改善効果があることを確認 ・80dB 程度の聴力低下があると明瞭度の改善が難しいことを確認（効果範囲の確認）
研究要素2 話者音声認識技術の開発 【2-1】 短時間で学習が可能な話者音声認識技術の開発 【2-2】 操作用タブレットアプリ開発	・学習用文章音読時間 1 分内 ・操作アプリの開発	90%	・学習用文章 音読時間 1 分以内達成 ・認識率・認識速度の高度化、事業化促進のためにベースとなる音声認証プログラムを変更、Android アプリとして実装済み ・93%以上の精度で認識可能であることを確認 ・高精度・高速な話者認識の実用化には、音声認証プログラム開発元との共同開発が必要 ・操作アプリ「iListening」を開発・実装済み
研究要素3 難聴者向け音声評価技術の開発 【3-1】 難聴者による音声評価 【3-2】 音声の物理特性解析	・データ蓄積アプリの改良 ・難聴者による明瞭度評価実施 ・聴こえやすさの物理要因分析	100%	・音声明瞭度評価のためのアプリを改良し、開発の効率化を行った ・【1-2】補聴信号処理法を評価可能な明瞭度評価試験を設計・実施し、信号処理の効果を確認 ・語音ごとの周波数分析により聞き取りやすい音声と聞き取りにくい音声の物理特性の違いを明確化し、信号処理法の開発に活かした

1-4 当該研究開発の連絡窓口

事業管理機関

名称： 地方独立行政法人東京都立産業技術研究センター

住所： 〒135-0064 東京都江東区青海 2-4-10

連絡窓口： 開発本部 開発企画室

Tel: 03-5530-2528

Fax: 03-5530-2458

第2章 本論一（1）話者音声改善技術の開発 -音声明瞭度の改善-

2-1 概要

声が小さくて聞き取りにくい人、モゴモゴとした話し方で聞き取りにくい人、アナウンサーのように明瞭な話し方をする人など音声の特徴は様々である。また音声をマイクを通して再生すると、その特徴(声の印象や明瞭度)の差が大きくなること、難聴者は健聴者よりもその差を大きく知覚することが確認されている。特に音声明瞭度に関しては健聴者が聴取した場合は明瞭度に違いがなくても、難聴者が聴取すると音声が認識しやすい人、認識しにくい人がはっきりと分かれる場合がある。つまり従来の難聴者自身の聴力補正に加え、話者の音声特性を改善することにより、難聴者の聴こえが大きく改善できる。そこで、聞き手の聴力特性に合わせて音声を改善する従来の補聴信号処理に加えて、話者の音声を聞き取りやすい音声に変換して音声を改善するシステムを考案した(図2-1)。本事業では、話者音声を改善する方法として、聞き取りやすい明瞭な音声特性(理想音声モデル)にあわせて話者音声を補正する信号処理手法の開発を行った(研究要素1)。また、信号処理法を開発するための「聞き取りやすい音声特性」の解析と、開発したシステムを定量的に評価する評価方法の開発を研究要素3「難聴者向け音声評価技術の開発」として実施した。

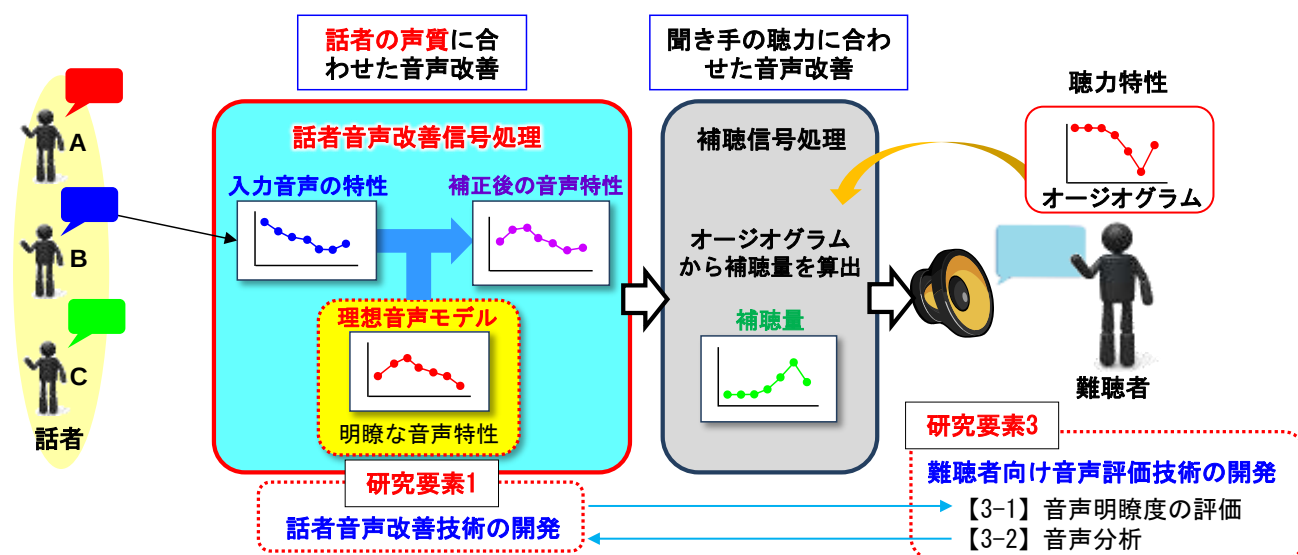


図2-1 開発する補聴システムの概要と研究要素

2-2 音声特性モデルの作成

難聴者へのヒアリングにより、聞き取りやすい音声の代表としてアナウンサーが挙げられた。そこで、アナウンサーと一般人との音声を収録して比較を行った。また一部の難聴者と対話する機会が多い人には、難聴者と話すことを意識した聞き取りやすい発音と、聞き取りにくそうなもごもごとした発音をしてもらい、それぞれの音声の比較を行った。これらの話者の音声特性を分析することで、聞き取りやすい音声と聞き取りにくい音声の物理的な違いを調べ、聞き取りやすい音声の特性を「理想音声モデル」とした。

2-2-1 音声の収録

話者による音声特徴を分析するためには、分析する音声音源に日常で使用される様々な語音が含まれていなければならない。そこで、下記の2種類の音源を収録した。

- 1) 日常会話で使用される語音がバランスよく含まれた文章を音読した音源
- 2) 単音節ごとに個別に発音した音源

1) については、研究要素2で開発した、日常会話における単音節の出現頻度を考慮して作成した学習用文章（ニュース原稿）を利用した（詳細は 3-2 で後述）。この文章には、日常会話で使用する単音節がバランスよく含まれているため、話者の総合的な音声特徴が分析可能である。ただし、語音ごとに細かく分析することは難しく、特に語音ごとに周波数が大きく異なる第2フォルマントや子音部分の特性が分かり難い欠点がある。そこで、2) の単音節ごとの個別に発音した音源による詳細な分析も合わせて行った。2) の単音節は、補聴器適合検査の指針(2010)に記載されている語音弁別検査で使用する語表の単音節を参考にし、清音44語と57-S語表に含まれる濁音（ガ、ゴ、ジ、ズ、ダ、デ、ド、バ）の計52語とした。

音声の収録時、被験者は半無響室内に設置した椅子に着席して原稿を読み上げた。口元正面 20cm～30cm の点にマイクロホン(SENNHEISER MKH40-P48)を設置し、オーディオインターフェース(RME Babyface Pro)を通して PC にて 24bit、48kHz サンプルングで録音した。収録システムの構成を図 2-2-1 に示す。表 2-2-1 に示す被験者 33 名の収録を行った。

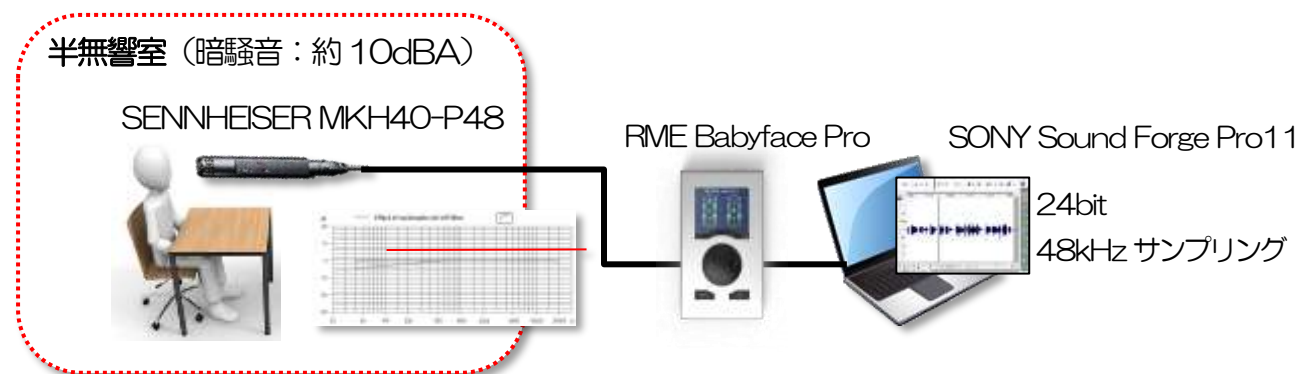


図 2-2-1 収録システムの構成

表 2-2-1 収録した被験者の内訳

	アナウンサー	一般	合計
男性	5	14	19
女性	7	7	14
合計	12	21	33

2-2-2 音声分析方法

音声の分析方法は目的に応じて様々あり、詳細な特徴量の分析にはケプストラム分析や LPC 分析などが用いられるのが一般的である。ただしこれらの分析は信号処理量が多いため、実用化時に実装予定の DSP やタブレットでは計算できない、または時間がかかりすぎることで懸念された。また、後述する補聴信号処理でも、複雑なフィルターや非線形信号処理方法を採用した場合、歪の発生や時間遅れにより明瞭度が低下してしまう恐れがあるため、線形イコライザーでの信号処理ができることが望ましい。そこで今回は、音声特徴としてフォルマント周波数範囲が確認可能であること、分析に必要な信号処理量が少ないこと、特徴量を 1/3 オクターブバンド線形イコライザーで信号処理しやすい状態で分析可能であることのすべてを満たす 1/3 オクターブバンド分析を採用した。ニュース原稿音源は全文の平均エネルギーとして分析を行い、単音節音源は単音節ごとに分析を行った。

2-2-3 話者音声特性の比較と理想音声モデルの作成

ニュース原稿音源のアナウンサーと一般人の音声分析結果の一例を図 2-2-2 に示す。図 2-2-2 (a) は男性アナウンサーの 3 名、図 2-2-2 (b) は男性一般人の 3 名を示したものである。一般人の多くは基本周波数(F_0)のエネルギーが最も大きく右肩下がりになっているのに対し、アナウンサーの多くは数百 Hz 辺りが山となる点に違いが見て取れる。200Hz~1kHz 辺りの帯域は、母音の認識に重要な第 1 フォルマント(F_1)の成分が集中する帯域であり、アナウンサーは F_0 と比べて F_1 のエネルギーが相対的に大きいことにより明瞭度が高くなっていると考えられる。ただし、アナウンサーの中にも、アナウンサー B のように一般人に近い特徴となっている人も見られた。そこで、難聴者にこれらの音源を視聴してもらい、聞き取りやすい音源を順位付けしてもらったところ、下記のような結果が得られた

【聴こえやすさの評価結果（聞き取りやすさの順）】

アナウンサー A、C > アナウンサー B > 一般人

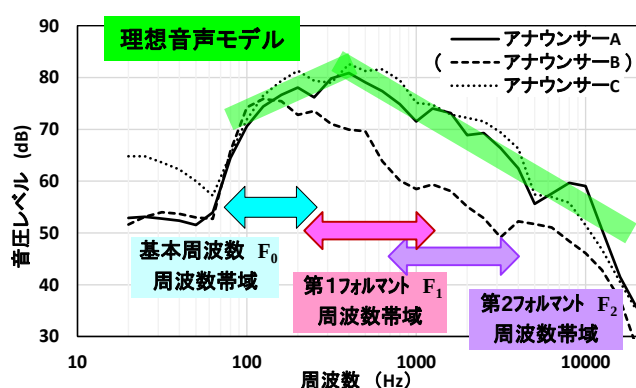
この結果より、アナウンサー A、C のように F_0 と比べて F_1 のエネルギーが相対的に大きい音声特徴が聞き取りやすい音声であると考えられる。

次に男性一般人のうち、難聴者と対話する機会が多い人に、難聴者と話すことを意識したはっきりとした発音（聞き取りやすい発音）をしてもらった場合ともごもごした発音をしてもらった場合の 2 名の結果を図 2-2-3 に示す。どちらの被験者も、もごもご発音時には F_0 に比べて F_1 のエネルギーが小さかったが、はっきり発音時には F_0 と F_1 のエネルギーがほぼ等しくなっており、図 2-2-2 のアナウンサー A、C の音声特徴に近くなっていることがわかる。この結果より、同一人物であってもアナウンサーと同様な音声特徴であれば明瞭度が高くなることが分かった。

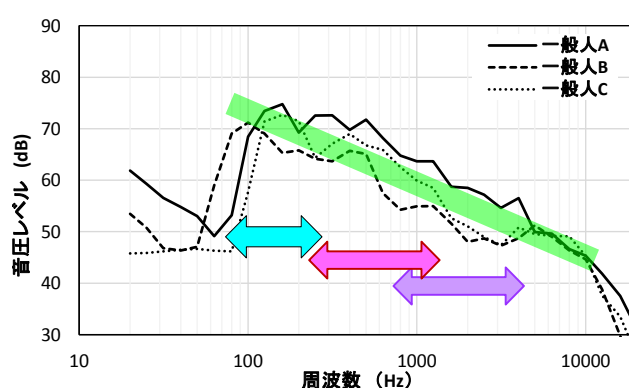
次に単音節ごとの音声分析を行った結果を示す。ニュース原稿音声は様々な語音が含まれており、前述の分析結果は音声全体を平均化処理しているため、語音ごとに周波数が大きく異なる第 2 フォルマント(F_2)や子音部分の特性が分かり難い。そこで単音節ごとに収録した音声を語音ごとに分析し、 F_2 や子音部分の分析を行った。アナウンサー A および一般人 A のもごもご発音、はっきり発音時の分析結果をあ段、い段、う段、

え段、およびお段に分けて図2-2-4～図2-2-8に示す。すべての結果において、アナウンサー、一般人のはっきり発音、一般人のもごもご発音の順で F_2 のエネルギーが大きいことがわかる。また、語音により F_2 のピーク周波数だけでなく F_2 周波数帯域全体のエネルギーが大きい。この周波数帯域は、無声子音を認識するために重要な情報が含まれている周波数帯域でもあるため、 F_2 だけでなく子音についてのエネルギーも大きくなることにより、明瞭度が高くなっていると考えられる。上記分析結果は男性話者について示したが、女性話者についてもほぼ同様な結果が得られた。

以上の結果より、 F_0 と比べて F_1 および F_2 周波数帯域のエネルギーが相対的に大きい音声聞き取りやすい音声の特徴であることが分かった。そこで、この音声特徴を満たし、難聴者による試聴結果も一番良かったアナウンサーA、Cの音声特徴（図2-2-2（a）緑線）を理想モデルとして設定することとした。



(a) アナウンサーの音声特徴



(b) 一般人の音声特徴

図2-2-2 アナウンサーと一般人の音声特徴の比較

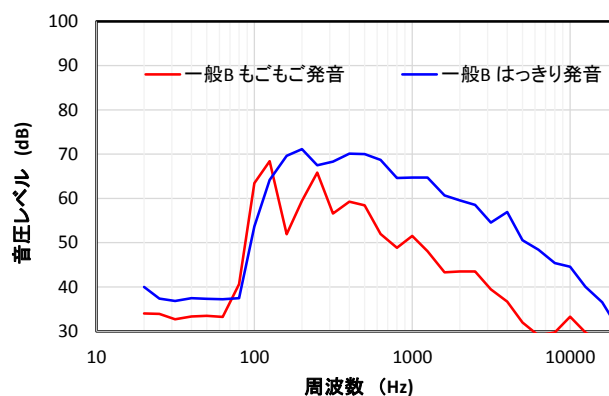
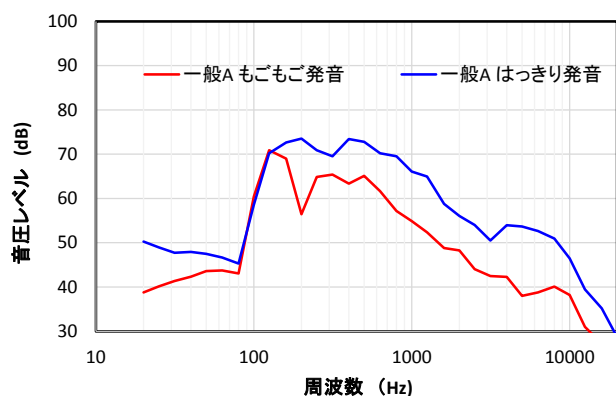


図2-2-3 一般人のはっきり発音ともごもご発音の特徴比較

あ段の周波数分析結果

〔あ, か, さ, た, な, は, ま
, や, ら, わ, が, だ, ば〕

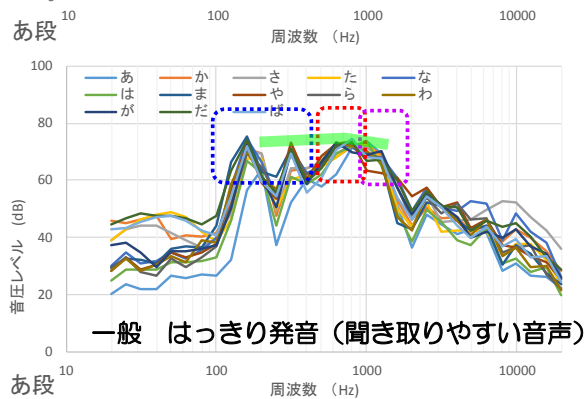
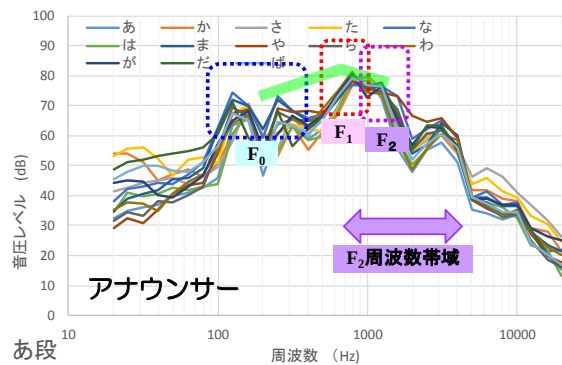
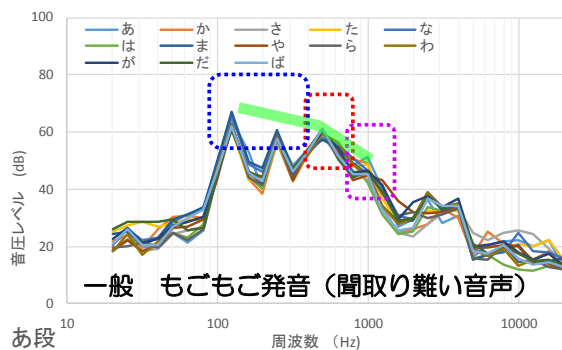


図 2-2-4 あ段の音声分析結果

い段の周波数分析結果

〔い, き, し, ち, に, ひ, み
, り, じ〕

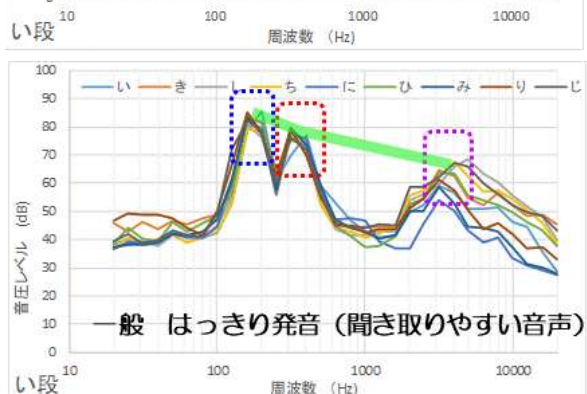
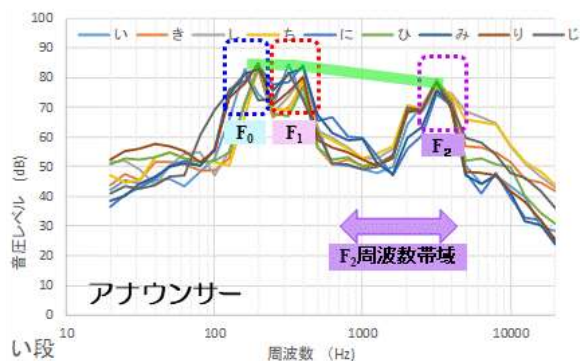
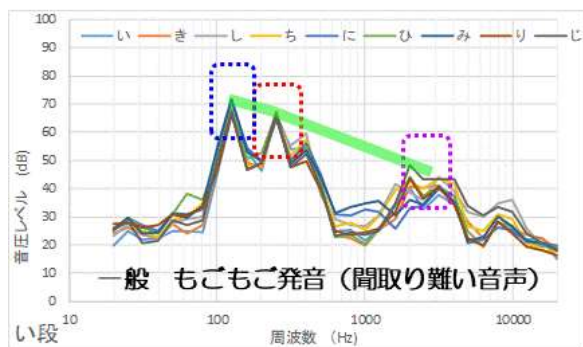


図 2-2-5 い段の音声分析結果

う段の周波数分析結果

〔う, く, す, つ, ぬ, ふ, む〕
〔, ゆ, る, ず〕

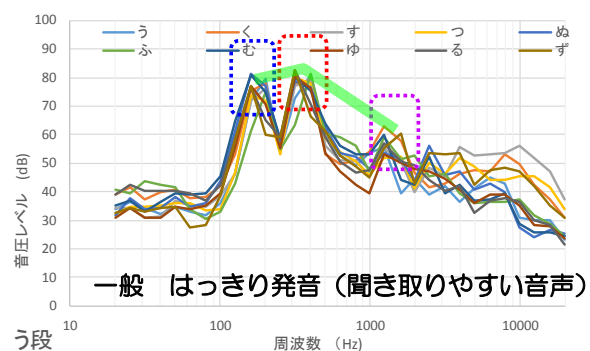
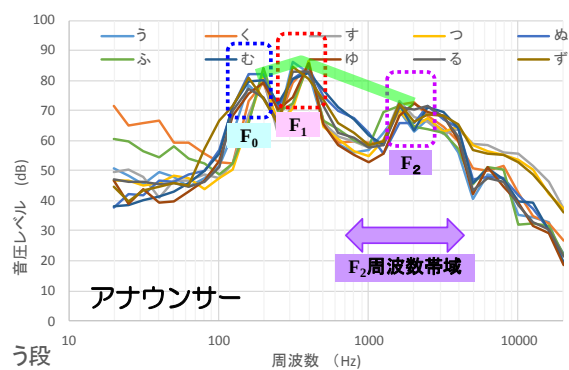
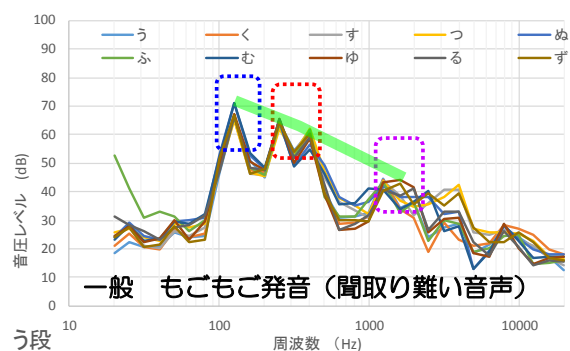


図 2-2-6 う段の音声分析結果

え段の周波数分析結果

〔え, け, せ, て, ね, へ, め〕
〔, れ, で〕

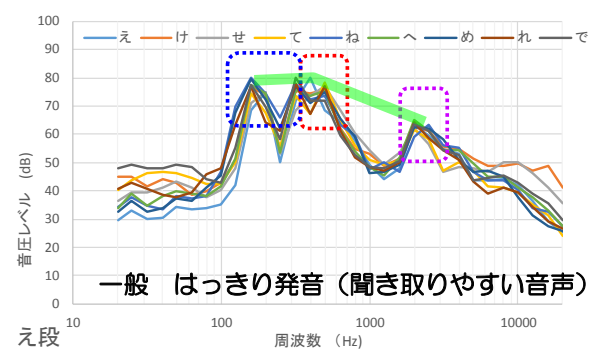
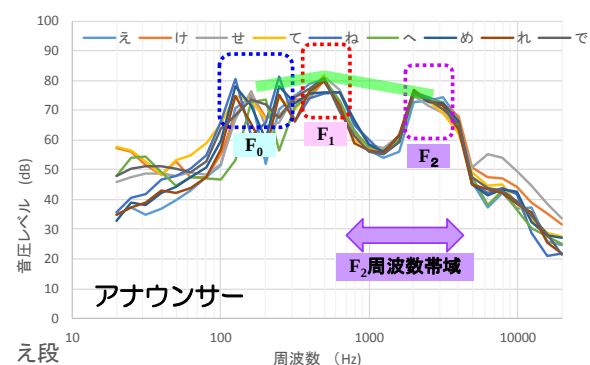
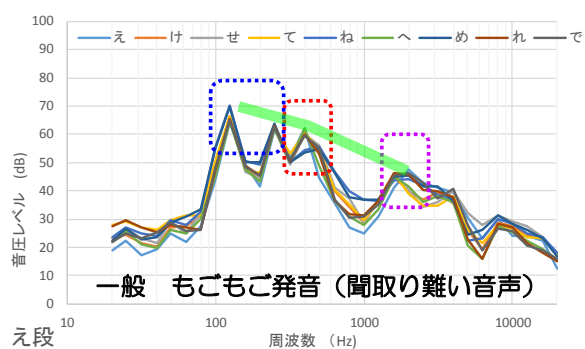


図 2-2-7 え段の音声分析結果

お段の周波数分析結果 〔 お, こ, そ, と, の, ほ, も , よ, ろ, ご, ど 〕

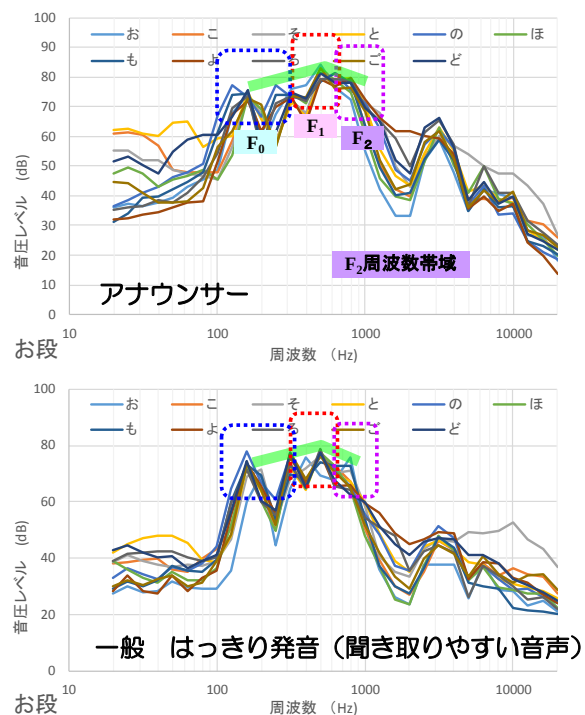
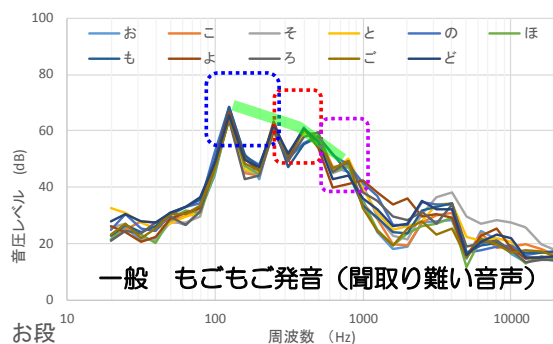


図 2-2-8 お段の音声分析結果

2-3 補聴信号処理方法の開発

2-2 で開発した理想音声モデルに話者の音声を近づけることで音声明瞭度を改善する信号処理方法について開発を行った。信号処理を行うにあたり注意しなければいけない点として、話者の個人識別は可能なまま（個人を識別する情報を保ったまま）明瞭度を改善する必要がある。話者の個人識別は、時間的な要素である話し方（話速、間など）と声帯から口までの生体的特徴が表れる周波数的な要素（声の高さ（ピッチ）など）により識別される。今回開発した信号処理方法は、時間的変動やピッチにかかわる補正は行わず、各周波数帯域の音圧を調整して理想音声モデルに話者の音声特性を近づける方法とした。そのため、話者の個人を識別する情報は保たれたまま明瞭度を改善可能である。

2-2 で開発した理想音声モデルは、 F_0 、 F_1 、 F_2 周波数帯域のエネルギー（音圧レベル）のバランスがポイントであるため、それぞれの帯域ごとのエネルギー（平均音圧レベル）を算出し、話者と理想音声モデルとの差から補正値を算出すれば補正が可能である。ただしこの算出方法では音声全体の情報が必要となるため、1/3 オクターブバンドの中心周波数で 100Hz～6300Hz の 19 バンドの音声情報が必要となる。搭載する DSP によっては計算するバンド数が多すぎるため、明瞭度に一番重要な F_0 と F_1 の音圧レベル差に着目した計算方法により、計算に必要な音声情報の周波数バンド数を約半分の 8 バンドで補正値を算出する方法を考案した。まず F_0 と F_1 の周波数帯域の平均音圧レベルからの F_0 と F_1 の音圧レベル差を算出し、理想音声モデルと話者音声特性の差から F_0 から F_1 帯域までの各周波数帯域（1/3 オクターブバンド）の補正値を算出する。次に F_2 帯域以上については、増幅上限周波数（男女により異なる）と増幅率上限値（使用するハードウェアのシステムにより異なる）を設定して各周波数帯域の補正値を算出する。各周波数帯域の補正は線形イコライザーによって行う。詳細な補正値の算出手順は次のとおりである。

【補正値の算出手順】

- 1) 表 2-3-1 に示す F_0 周波数帯域（5 バンド）、 F_1 周波数帯域（4 バンド）の平均音圧レベル（パワー平均）を話者および理想音声モデルごとに算出
 - ・話者 F_0 平均： L_{tF_0} 、話者 F_1 平均値： L_{tF_1}
 - ・理想モデル F_0 平均： L_{iF_0} 、理想モデル F_1 平均値： L_{iF_1}
- 2) F_0 周波数帯域平均値に対する F_1 周波数帯域平均値のレベル差 ΔL を算出
 - ・話者レベル差： $\Delta L_t = L_{tF_1} - L_{tF_0}$
 - ・理想モデルレベル差： $\Delta L_i = L_{iF_1} - L_{iF_0}$
- 3) 話者のレベル差と理想音声モデルのレベル差の差分 ΔL_{dt} を算出
$$\Delta L_{dt} = \Delta L_i - \Delta L_t$$
- 4) ΔL_{dt} を F_0 周波数帯域の中心周波数と F_1 周波数帯域の中心周波数間の差分とみなし、その傾きから表 2-3-1 の各周波数帯域の補正値を算出
- 5) 表 2-3-1 の最大周波数帯域以降は、増幅上限周波数（男性：2500Hz、女性：5000Hz）で増幅率の上限値となるような傾きで各周波数帯域の補正値を算出
（増幅率の上限値は、音割れやキンキンとした不自然さが発生しない程度の値を設定する）

6) 増幅上限周波数以降の周波数帯域は増幅率の上限値とする（または減衰させてもよい）

表 2-3-1 計算に使用する周波数バンド

	1/3オクターブバンドの中心周波数(Hz)								
	F ₀ 周波数帯域					F ₁ 周波数帯域			
男性	100	125	160	200	250	250	315	400	500
女性	200	250	315	400	500	500	630	800	1000

上記の算出手順であれば、1)でF₀とF₁が存在する帯域の平均音圧レベルをパワー平均で求めているため、話者のピッチ（声の高さ）によらずF₀とF₁の関係量を抽出できる。また周波数変調などのピッチ補正は行わずに音圧レベルの補正を行っているため、話者の周波数的な個人情報は保たれる。この信号処理方法にて補正を行った前後の話者音声特性と理想音声モデルを図 2-3-1 に示す。音声特性が異なる4名の結果を示しているが、すべての話者の特性が理想音声モデルに近づいていることがわかる。

以上の信号処理方法を話者音声改善信号処理方法として開発を行った。この信号処理方法は、スピーカシステムとの通信を行うことが可能なタブレットアプリにて補正値の算出を行い、各周波数帯域の補正値をスピーカシステムのDSPに転送して、DSP内の線形イコライザーにて信号処理する方法で実装を行った。すでに開発済みの聞き手の聴力特性に合わせた補聴信号処理方法と合わせて信号処理することで（図 2-1）、難聴者の明瞭度改善が期待できる。

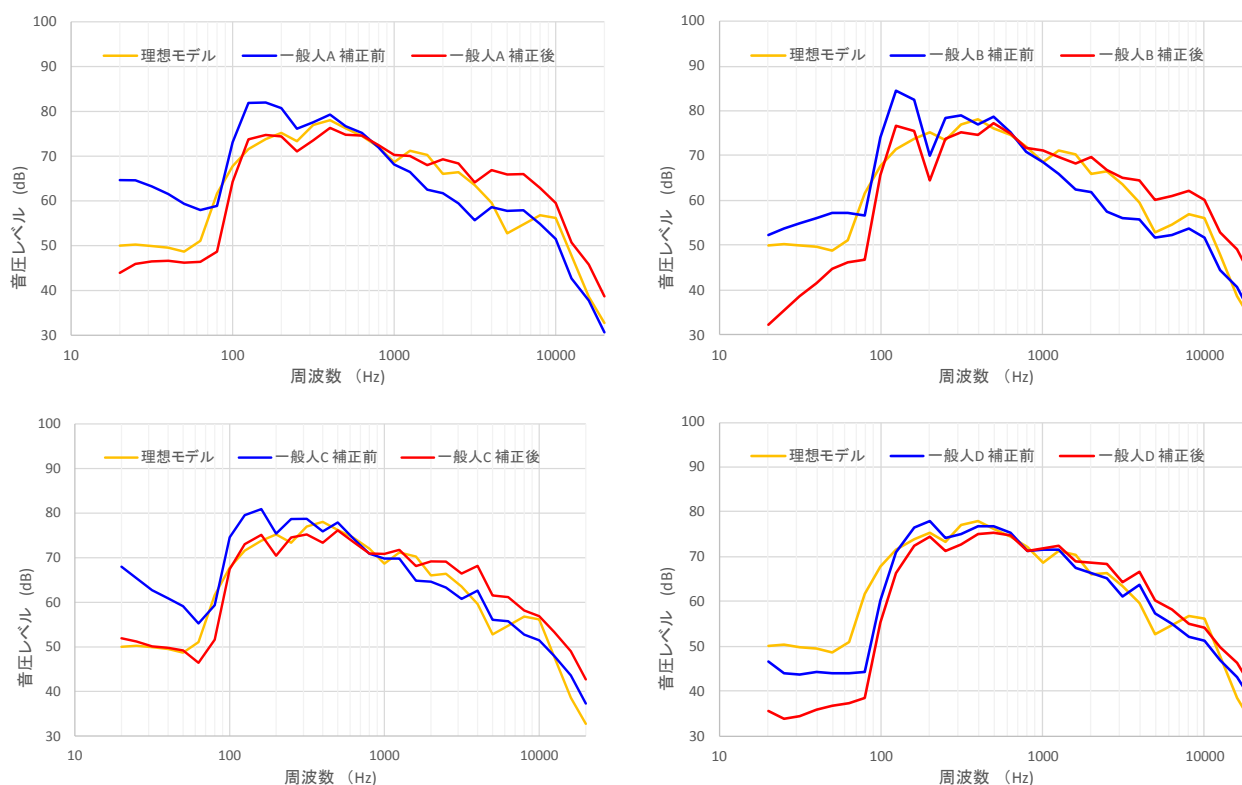


図 2-3-1 補正前後の音声特性と理想音声モデル

2-4 開発した信号処理の評価

2-4-1 語音弁別検査による明瞭度評価

スピーカシステムによる音声情報の明瞭度を評価する方法には明瞭度指数 STI や AI などがあるが、これらは健聴者が音声のある音響空間で聴取した場合を対象とした評価方法であり、難聴者に対する評価はできない。そこで難聴者による評価を行うため、日本聴覚医学会の補聴器適合検査の指針(2010)に記載されている、補聴器のフィッティング確認の明瞭度評価方法を参考にスピーカの明瞭度評価方法を考案した。

前述の指針では、補聴器のフィッティング効果を確認するため、補聴器を装着した状態でスピーカからの音声を聴取して明瞭度の評価を行う。この時使用するスピーカは、周波数特性を平坦に校正したものが用いられるが、このスピーカを開発したスピーカシステムに置き換えて、スピーカを含めた音声再生装置を聴取補助機器と位置付けて聴取試験を実施することにより、明瞭度の評価を実施することとした。事前にサウンドレベルメータなどでスピーカシステムの音量を校正し、指針と同様に試験を実施すれば、補聴器と同様にスピーカシステムの明瞭度を評価可能である。

具体的な明瞭度評価の方法は指針と同じ語音弁別検査による語音明瞭度の評価とした。語音弁別検査は、決められた単音節を聞き取って回答し、その正答率が語音明瞭度となる。正答率の算出を母音のみまたは子音のみに着目して算出することにより、母音明瞭度、子音明瞭度を算出することも可能であり、これらの明瞭度を評価することで、どのような語音が聞き取りにくいかを評価することも可能である。検査に用いられる語音は全 50 語の 57-S 語表または全 20 語の 67-S 語表として決められており、これらの語表を使用した時の語音明瞭度（正答率）が 50%以上であれば日常会話が可能な聴こえであると評価できる。被験者の聴力特性により最適な音量が異なるため、提示音量を変えた検査を複数回実施する必要があることから、被験者の負担を軽減するために全 20 語の 67-S 語表を使用して明瞭度の評価を行うこととした。

本事業の達成目標は、上記明瞭度評価方法による高度難聴者までの語音明瞭度が 50%以上となることを目標値としている。

2-4-2 語音弁別検査アプリ開発による明瞭度評価の効率化

前述の明瞭度評価方法は、聞き取った語音を回答用紙に記載し、検査終了後に検査実施者が正誤を手作業で採点することで明瞭度を算出する。母音明瞭度や子音明瞭度を評価する場合は、母音と子音とを分けて算出する必要があるため、評価に時間がかかるとともに採点間違いの可能性も高くなる。通常この明瞭度算出作業には採点のダブルチェックも含めて 20 分以上を要するため、多くの検査を実施する場合には効率が悪い。また、算出した明瞭度を参考にして次に検査する音量や補正処理パラメータ等を決定するため、被験者の待ち時間も長くなり、被験者の負担にもなる。そこで本事業では、効率よく語音弁別検査を実施するため、検査用タブレットアプリ開発による効率化に取り組んだ。

タブレットアプリにて語音の再生と回答の入力を行い、検査終了後には自動で語音明瞭度、母音明瞭度、子音明瞭度を算出する。高齢者でも入力が容易なように手書き入力を採用するなど、被験者の負担が少なくなるような設計を行った。また検査前の音量校正、検査条件の入力、検査結果の管理をアプリ内で実施できる機能を搭載し、語音弁別検査の一連の作業をこのアプリで実施できるようにした。開発した語音弁別検査

アプリを使用した検査手順を図2-4-1に示す。この音声弁別検査アプリにより、採点時間は1分程度に短縮し、音声弁別検査の大幅な時間短縮と被験者の負担軽減が可能となった。本アプリは今後の医療機関、高齢者関連施設等での明瞭度の検証や、簡易聴力検査にも活用可能である。

音声弁別検査アプリによる明瞭度評価手順

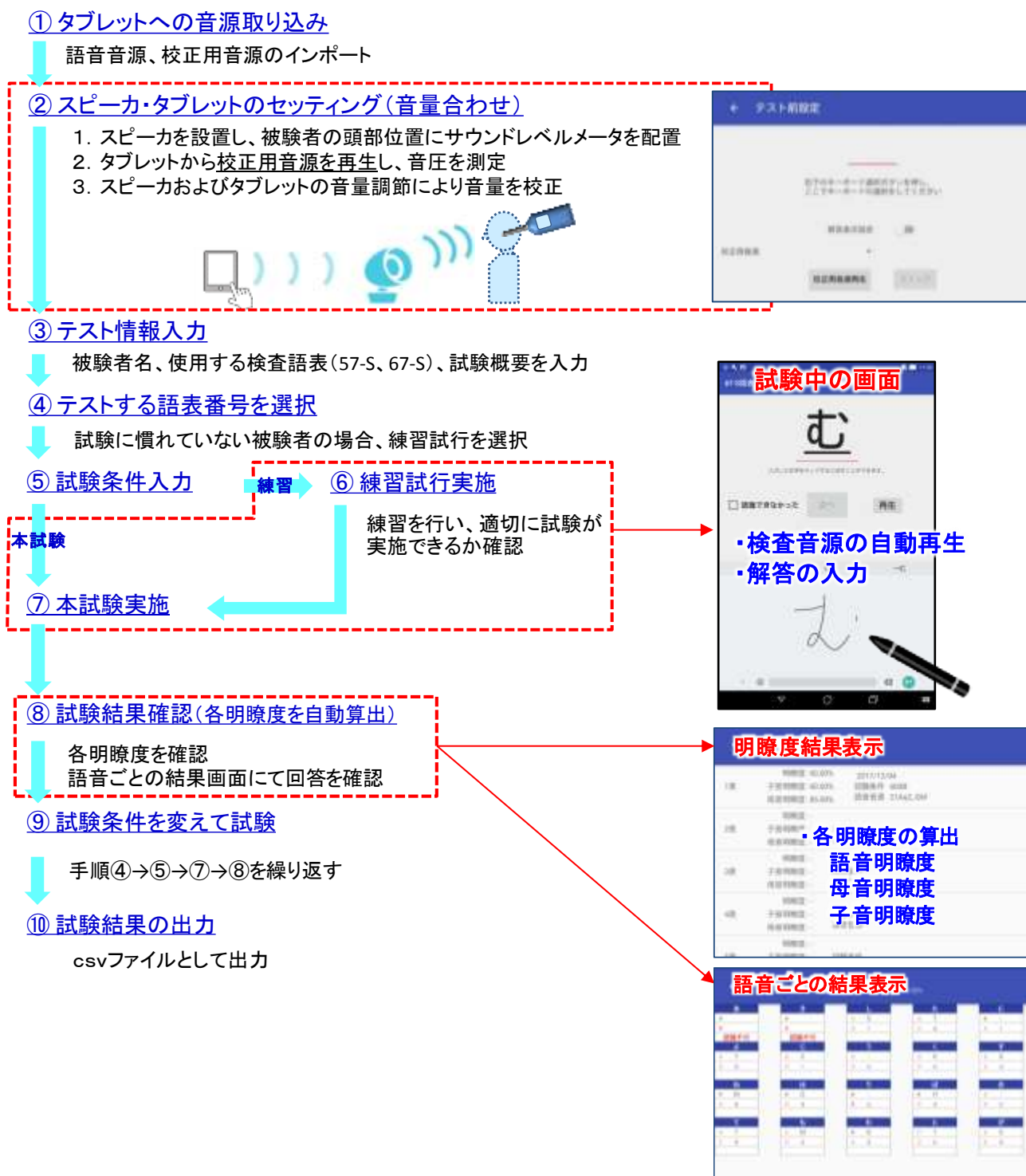


図2-4-1 開発した音声弁別検査アプリによる明瞭度評価手順

2-4-3 語音弁別検査の条件と明瞭度の評価結果

2-3 で開発した補聴信号処理の効果を確認するため、前述の語音弁別検査による音声明瞭度の評価を実施した。事前に 2-2-1 で収録したニュース原稿音声を難聴者に試聴評価してもらい、聞き取りにくい話者であると評価された一般話者 A の音声について信号処理前後の明瞭度を評価した。67-S 語表として聴取する語音の音声音源は、2-2-1 で収録した一般話者 A の単音節ごとの音声音源を利用し、この音源の信号処理前と処理後の 2 種類の音源を用意した。ただし、音声の音量は単音節ごと、信号処理前後で異なってしまう。そこで日本聴覚医学会から語音弁別検査用に提供されている検査用音源 CD に収録されている 67-S 語表の単音節ごとの音量を分析し、その音量にそれぞれの音源の音量を合わせて検査音源とすることで、処理前後の音源の音量は同一とした。検査環境は、通常の語音弁別検査では暗騒音や残響の低い防音室などが使用されるが、本開発システムは会議室などの騒音や残響が存在する空間での使用を想定しているため、一般的な会議室にて検査を実施した。検査環境の詳細は図 2-4-2 に示す。

被験者は高度難聴者を中心とした 24 名（軽度：2 名、中等度：5 名、高度：13 名、重度：4 名）とした。被験者により最適な音量や補聴器および人口内耳の使用の有無が異なるため、被験者に合わせて音量および話者音声改善信号処理・聴力補正信号処理の機能を切り替えた。

語音明瞭度の評価を行い、明瞭度が目標値の 50%以上となった人数を表 2-4-1 に、割合として表示したものを図 2-4-3 に示す。補聴器や人工内耳を装着していない裸耳では軽度・中等度難聴者はすべての被験者で目標値をクリアし、高度難聴者も 7 割以上が目標値をクリアした。補聴器や人口内耳を装着した状態では、中等度・高度難聴者はすべての被験者で目標値をクリアし、重度難聴者であっても 7 割以上が目標値をクリアした。本結果には補聴信号処理前でも明瞭度が 50%以上であった被験者も含まれているが、そのような被験者でも、補聴信号処理後には明瞭度の改善がみられている。以上の結果より、裸耳だけでなく補聴器や人口内耳を併用した場合でも明瞭度が 50%以上となることが確認でき、本事業の目標値である明瞭度 50%以上を達成した。次に、代表的な被験者の詳細な結果について示す。まず、改善効果の見られた 2 名の高度難聴者の事例を図 2-4-3、図 2-4-4 に示す。図 2-4-3 の被験者は裸耳ではほぼ音声か認識できないため、日常生活では両耳に補聴器を装着している被験者である。裸耳では信号処理無しはもちろんのこと、話者音声改善信号処理による話者補正のみでは明瞭度は 50%以下であり、明瞭度の改善は見られなかった。これは話者音声の改善はされていても、難聴者自身の聴力が低いいため明瞭度の改善が見られないと考えられる。そこで、難聴者自身の聴力を補正する聴力補正信号処理と組み合わせた結果、音量 80dBHL では語音明瞭度は 85%（子音明瞭度：85%、母音明瞭度：100%）となり大幅な改善が見られた。この検査において、被験者は裸耳の状態で音声か聴取できることに驚いていた。また、補聴器装用時には信号処理なしでも明瞭度は 50%を超えているが、話者音声補正を行うことで 5~10 ポイントの明瞭度改善が見られ、被験者からも聞き取りやすくなったとの感想をいただいた。図 2-4-4 の被験者は大きな音量（70dBHL 程度）であれば聴取が可能であるが、補聴器を装用してもあまり明瞭度が改善しない被験者である。裸耳の場合には話者補正と聴力補正を組み合わせることによって 60dBHL での語音明瞭度が 40%から 80%へ大きく改善した。聴取できる音量が 70dBHL から 60dBHL に 10dB 低減されたことは、感覚的に約半分の音量で聴取が可能になったことを意味する。また補聴器装用時は、話者補正のみで語音度が 50%から

80%へ大きく改善した。この結果より、従来の補聴器のような聴力補正のみでは明瞭度の改善が難しい難聴者に対して、話者音声改善信号処理が有効であることが確認された。次に改善効果の低かった高度難聴者の事例を図2-4-5に示す。この被験者は特に高い周波数の聴力レベルが80dBと低いため、子音の聞き取りが難しい被験者である。裸耳では明瞭度の改善が見られず、補聴器装用時には話者補正により明瞭度が50%を超えたが、多くの子音が聞き取れておらず十分な改善とは言えない。高い周波数帯域の聴力補正が不足しているのではないかと考え、追加で聴力補正を行ったが、音が歪んでしまい明瞭度は逆に低下してしまった。他の聴力レベルが低い被験者についても同様の結果が得られている。検査に使用した試作のスピーカシステムでは、適正な補正が可能な増幅率が限られており、概ね聴力レベルが80dB以上の難聴者だと明瞭度の改善が不足してしまうことが確認された。

以上の結果より、2-1で開発した話者音声改善信号処理は、適切な聴力補正と組み合わせることにより明瞭度を改善することが可能であり、高度難聴者の明瞭度を50%以上に改善できることを確認した。また、補聴器や人工内耳を装用している場合、話者音声信号処理のみで明瞭度が改善することを確認した。課題としては、聴力レベルが約80dB以上の周波数帯域があると、聴力補正に適切な増幅率が得られず、明瞭度が低下してしまうことが分かった。改善方法としては、スピーカの設置位置をより耳の近くに設置して運用する方法や、イヤホン・ヘッドホンのように外耳道を密閉する形で装着する方法により、音量を確保する手法が考えられる。

試験環境：実使用環境で実施

測定場所：一般的な会議室
広さ：D4m × W7m × H2.4m
暗騒音：約40dBA

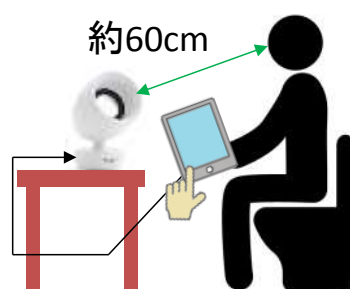


図2-4-2 検査環境の詳細

表2-4-1 明瞭度50%以上の人数

	裸耳		補聴器・人工内耳有り	
	被験者数	明瞭度 50%以上	被験者数	明瞭度 50%以上
軽度	2	2	—	—
中等度	4	4	2	2
高度	13	10	8	8
重度	—	—	4	3
合計	20	16	13	13

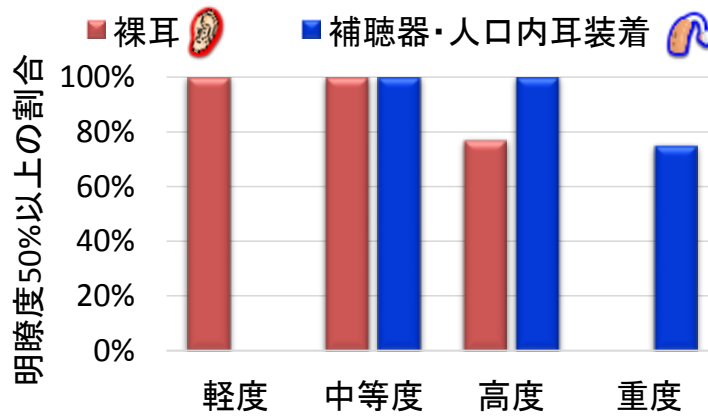


図2-4-2 明瞭度50%以上の割合

語音明瞭度の測定結果

高度難聴者

被検者N0	性別	年齢	補聴器	聴力レベル (4分法)	
				右	左
2	女	59	両耳	83dB	84dB
				高度	

補聴器	信号処理 無し	話者補正 のみ	聴力補正 のみ	話者補正+ 聴力補正
なし	×	△	□	○
あり	×	△	□	○

聴力補正パラメータ (Listening)

125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
0	5	5	9	9	5	5

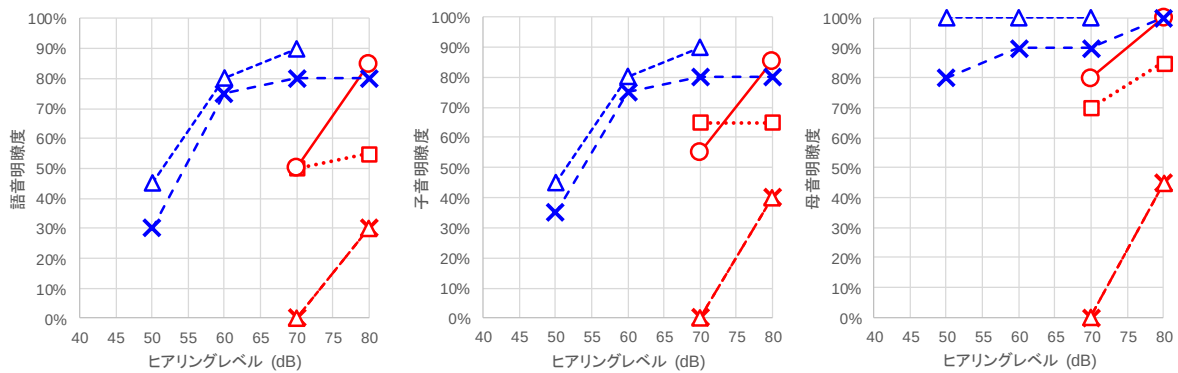
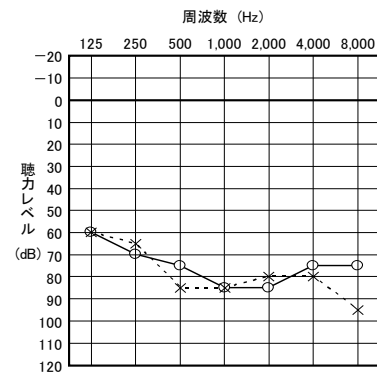


図2-4-3 高度難聴者の明瞭度測定結果1

語音明瞭度の測定結果

高度難聴者

被検者N0.	性別	年齢	補聴器	聴力レベル (4分法)		
5	女	47	両耳	右	78dB	高度
				左	78dB	

補聴器	信号処理 無し	話者補正 のみ	聴力補正 のみ	話者補正+ 聴力補正
なし	×	△	□	○
あり	×	△	□	○

聴力補正パラメータ (listening)						
125Hz	250Hz	500Hz	1kHz	2kHz	4kHz	8kHz
-5	0	9	9	9	5	5

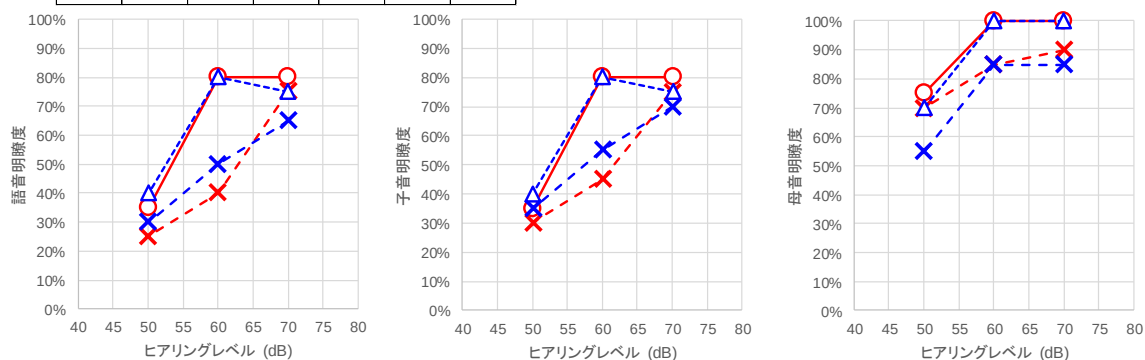


図 2-4-4 高度難聴者の明瞭度測定結果2

語音明瞭度の測定結果

高度難聴者

被検者N0.	性別	年齢	補聴器	聴力レベル (4分法)		
17	男	35	両耳	右	83dB	高度
				左	83dB	

補聴器	信号処理 無し	話者補正 のみ	聴力補正 のみ	話者補正+ 聴力補正
なし	×	△	□	○
あり	×	△	□	○

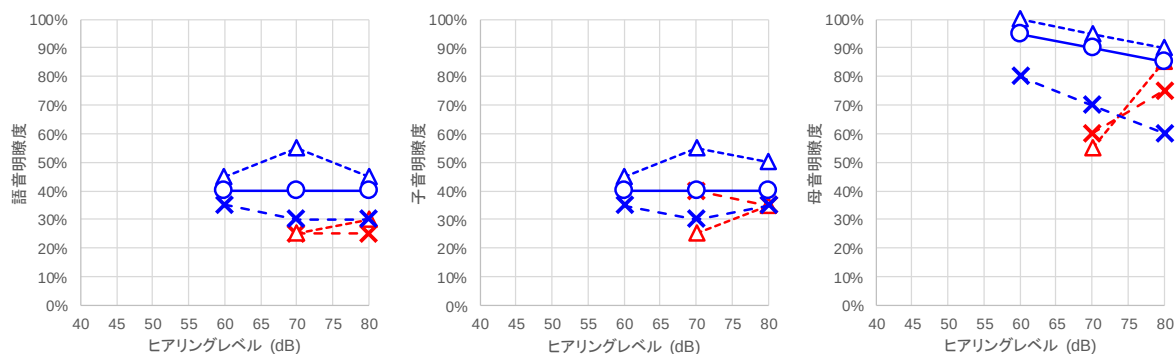


図 2-4-5 高度難聴者の明瞭度測定結果3

第3章 本論一（2）話者音声認識技術の開発 -ユーザビリティの改善-

3-1 概要

第2章で開発した補聴信号処理システムは、入力される話者が変わるたびに補聴信号処理の補正值を変更する必要がある。複数の話者がいる場合、これを手動で変更しなければならないとユーザビリティの低下を招く。そこで話者音声自動認識技術の開発を行った。話者音声認識には既存の音声認証技術を使用することで解決できるが、ユーザー（話者）による音声の事前学習が必要となる。この事前学習は文章音読作業となるが、数分～数十分の音読時間が必要となる。新しい話者が利用するたびにこの長時間の学習作業を行うことは不便であるため、1分以内の音読時間で事前学習が完了できる話者音声認識技術の開発を行った。また、既存のスピーカシステム本体の操作ボタンだけでは、事前学習を適切かつ容易に行うことは難しい。また、事前学習以外にも難聴者の聴力特性の入力などの操作が必要となるため、スピーカシステムと通信が可能なタブレットにて適切かつ容易に操作ができる操作用アプリの開発により、ユーザビリティの改善を行った。

3-2 短時間での学習が可能な話者音声認識技術の開発（学習用文章の短文化）

開発する話者認識技術は既存の音声認証技術を使用する。音声認証技術は本人認証を行うための音声パスワードなどに利用されている技術である。音声認証はにユーザーの音声データを学習データとして事前学習させることで話者を認識できるようになる。学習データとして必要な音声データの時間（文章読み上げ時間）は、認識させる人数、学習データの質（語音数、SN比・残響などの音響条件など）により変化する。現状の音声認証技術では、数人程度の認識に一人当たり数分間の音声読み上げ時間が必要となる場合があり、使用前に多くの学習時間が必要となる。これは使用者に大きな負担となることから、学習時間短縮のため、1分以内で学習が完了できる学習用文章の開発を行った。

前述の音声認証では、同じような声質の人が複数いると認証が困難になる。しかし本開発システムでは、同じような声質の人の補正量はほぼ同じとなるため区別する必要がない。「個人」を認識するための情報量（発声パターン）よりも「声質」のみを認識する情報量は少なくともよいと考えられるため、「声質」のみを学習すればよいのであれば学習用文章の短文化は可能であると考えた。ただし、認識時にはどのような語音が入力されるかわからないため、学習用文章中には日常会話で使用される多種の音素が含まれていることが望ましい。そこで、日常会話における単音節の出現頻度に合わせた文章を作成すれば、学習用文章の短文化と認識率の向上が期待できると考えた。日本語会話における単音節の出現頻度は先行文献（小寺ら, Audiology Japan 41, pp73-78, 1998）を参考にした。また、読みやすさを考慮してニュース原稿風の文章とした。上記条件を満たし、読み上げ時間が1分以内となるように作成した学習用文章を図3-2-1に示す。先行文献にある会話における単音節の出現頻度を表3-2-1に、作成した学習用文章の単音節の出現頻度を表3-2-2に示す。また、表3-2-1と表3-2-2の差を表3-2-3に示す。それぞれの表を比較すると、直音・拗音ともに頻度の差は5%未満であることが分かる。以上のように、会話における単音節の出現頻度を保ちつつ、読み上げ時間1分以内の学習用文章の作成を完了した。

4年前、世界で初めて小惑星の微粒子を地球に持ち帰った日本の探査機「はやぶさ」の後継機、「はやぶさ2」が3日、鹿児島県の種子島宇宙センターから打ち上げられました。「はやぶさ2」は午後3時10分ごろ、予定の起動に投入され、打ち上げは成功しました。

「はやぶさ2」を載せたH2Aロケット26号機は、午後1時22分、打ち上げの5秒前に1段目のメインエンジンに点火され、ごう音と共に発射台を離れました。

補助ロケットや1段目を切り離して上昇し、打ち上げからおよそ1時間47分後の午後3時10分ごろ、「はやぶさ2」を予定の軌道に投入し、打ち上げは成功しました。

図3-2-1 作成した学習用文章

表3-2-1 先行文献の会話における単音節の出現頻度(%) (19,988 モーラ)

直音		k	s	t	n	h	m	y	r	w	g	z	d	b	p	計
a	2.2	3.6	1.6	3.4	4.5	0.8	1.5	1.1	1.7	2.8	1.7	0.1	2.2	0.6	0.2	27.8
i	7.6	1.8	3.3	0.8	2.6	0.5	0.6		1		0.09	0.9		0.1	0.07	19.2
u	4.2	1.7	1.3	0.9	0.03	0.5	0.4	0.2	2.1		0.1	0.1		0.4	0.04	11.9
e	1.3	1.5	0.9	3.6	0.6	0.1	0.7		1.3		0.1	0.1	1.7	0.2	0.01	12.0
o	3.6	2.8	1	3.2	3.4	0.6	2.1	1.8	0.6		0.4	0.1	1.2	0.08	0.07	20.9
計	18.9	11.4	8.0	11.9	11.0	2.3	5.2	3.1	6.7	2.8	2.4	1.3	5.0	1.4	0.4	91.9

拗音	ky	sy	ty	ny	hy	my	ry	gy	j	by	py	計
a	0.06	0.20	0.30	0.00	0.03	0.00	0.07	0.01	0.60	0.00	0.00	1.2
u	0.02	0.09	0.02	0.03	0.00	0.00	0.02	0.00	0.10	0.00	0.01	0.3
o	0.03	0.03	0.20	0.04	0.03	0.01	0.08	0.03	0.30	0.03	0.01	0.7
計	0.1	0.3	0.5	0.1	0.1	0.0	0.2	0.0	0.9	0.0	0.0	2.1

表3-2-2 作成した学習用文章の単音節の出現頻度(%) (264 モーラ)

直音		k	s	t	n	h	m	y	r	w	g	z	d	b	p	計
a	1.6	2.9	2.9	2.9	1.3	2.9	1.9	1.9	0.8	1.9	1.3	0.0	0.8	0.0	0.0	23.3
i	4.8	2.1	2.9	3.2	3.8	0.0	0.3		0.8		0.0	1.9		0.3	0.3	20.4
u	7.8	1.3	0.5	2.1	0.0	0.5	0.0	0.0	0.5		0.0	0.0		1.3	0.5	14.7
e	1.6	1.3	2.7	1.9	1.1	0.0	1.1		0.8		1.3	0.0	0.5	0.0	0.0	12.3
o	2.9	0.8	0.3	2.1	3.8	0.5	0.8	1.3	1.3		2.9	0.0	0.5	0.0	0.0	17.4
計	18.8	8.6	9.4	12.3	9.9	4.0	4.0	3.2	4.3	1.9	5.6	1.9	1.9	1.6	0.8	88.2

拗音	ky	sy	ty	ny	hy	my	ry	gy	j	by	py	計
a	0.00	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.3
u	0.54	0.00	0.27	0.27	0.00	0.00	0.27	0.00	1.07	0.00	0.00	2.4
o	0.27	1.07	0.00	0.00	0.00	0.00	0.00	0.00	0.54	0.27	0.00	2.1
計	0.8	1.3	0.3	0.3	0.0	0.0	0.3	0.0	1.6	0.3	0.0	4.8

表3-2-3 先行文献と作成した学習用文章の単音節の出現頻度の差(%)

直音		k	s	t	n	h	m	y	r	w	g	z	d	b	p	計
a	-0.6	-0.7	1.3	-0.5	-3.2	2.1	0.4	0.8	-0.9	-0.9	-0.4	-0.1	-1.4	-0.6	-0.2	-4.5
i	-2.8	0.3	-0.4	2.4	1.2	-0.5	-0.3		-0.2		-0.1	1.0		0.2	0.2	1.2
u	3.6	-0.4	-0.8	1.2	0.0	0.0	-0.4	-0.2	-1.6		-0.1	-0.1		0.9	0.5	2.8
e	0.3	-0.2	1.8	-1.7	0.5	-0.1	0.4		-0.5		1.2	-0.1	-1.2	-0.2	0.0	0.3
o	-0.7	-2.0	-0.7	-1.1	0.4	-0.1	-1.3	-0.5	0.7		2.5	-0.1	-0.7	-0.1	-0.1	-3.5
計	-0.1	-2.8	1.4	0.4	-1.1	1.7	-1.2	0.1	-2.4	-0.9	3.2	0.6	-3.1	0.2	0.4	-3.7

拗音	ky	sy	ty	ny	hy	my	ry	gy	j	by	py	計
a	-0.1	0.1	-0.3	0.0	0.0	0.0	-0.1	0.0	-0.6	0.0	0.0	-0.9
u	0.5	-0.1	0.2	0.2	0.0	0.0	0.2	0.0	1.0	0.0	0.0	2.1
o	0.2	1.0	-0.2	0.0	0.0	0.0	-0.1	0.0	0.2	0.2	0.0	1.4
計	0.7	1.0	-0.2	0.2	-0.1	0.0	0.1	0.0	0.7	0.2	0.0	2.7

3-3 話者認識アプリの実装と話者認識結果の検証

3-2で作成した短時間学習用文章の検証を行うため、話者認識アプリを実装して話者認識率の検証を行った。ベースとなる音声認証プログラムには、端末の処理能力に依存しないクラウド型であり、自由度が高くあいまいな識別（個人識別ではなく声質の識別）にも対応可能なプログラム「mimi SRS（フェアリーデバイス（株））」を採用し、Android アプリとして話者認識アプリの実装を行った。実装した話者認識アプリの学習及び認識手順を図3-3-1に示す。

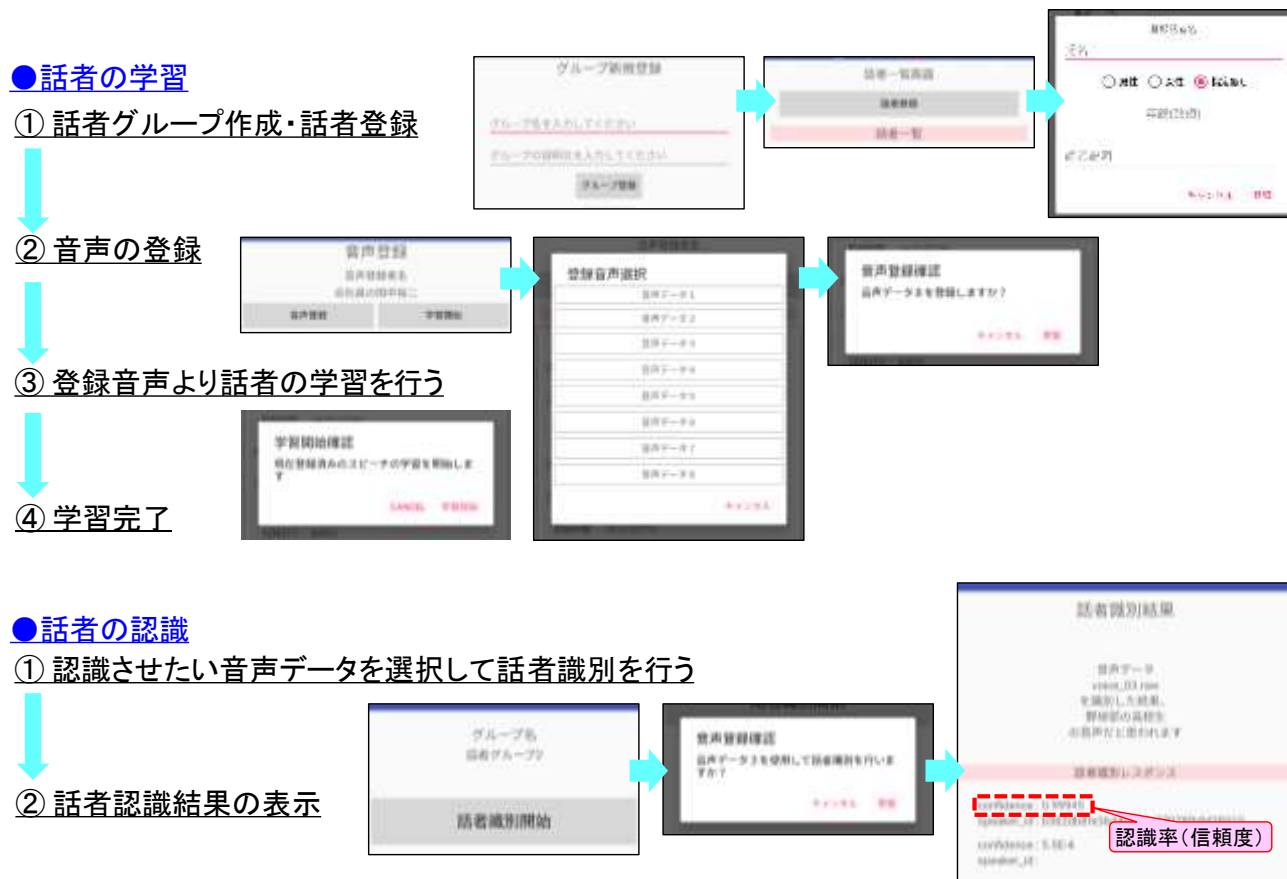


図3-3-1 実装した話者認識アプリの学習及び認識手順

3-2で作成した学習用文章を音読した音声を使用し、上記話者認識アプリにて話者の学習と認識率の検証をおこなった。男女15名の音声を学習させた後、学習させた音声とは別の内容の音声（話者ごとに5パターン）を用いて話者の認識を行った結果を表3-3-1に示す。すべての話者において1分以内の学習で93%以上の精度で話者を認識することが可能であることを確認し、目標を達成した。ただし実用化時の課題として、認証時間が平均で約25秒と遅いことが判明した。この認識時間の改善についてはベース音声認証プログラム開発元によるカスタマイズにより短縮が可能であることを確認している。今回利用したベース音声認証プログラム「mimi SRS」は現在トライアル状態であり、プログラム側の学習条件の最適化や、ノイズキャンセラー・エコーキャンセラー機能の追加により、認識速度と精度の改善が期待できる。また実用化のためには、会議用マイクなどの入力デバイスの開発も必要であり、音声認証プログラム開発元との共同研究開発を検討している。

表 3-3-1 話者の認識結果

話者	認識率						
	音声①	音声②	音声③	音声④	音声⑤	平均	最低
A (男)	99.2%	99.8%	99.8%	99.8%	99.8%	99.7%	99.2%
B (男)	95.4%	94.5%	97.3%	98.3%	94.6%	96.0%	94.5%
C (男)	99.8%	99.8%	99.7%	99.9%	99.9%	99.8%	99.7%
D (男)	99.3%	98.6%	98.2%	96.5%	98.9%	98.3%	96.5%
E (男)	98.6%	97.3%	98.9%	99.5%	98.8%	98.6%	97.3%
R (男)	98.8%	98.5%	99.1%	99.2%	98.5%	98.8%	98.5%
G (男)	99.6%	99.4%	99.6%	99.8%	99.7%	99.6%	99.4%
H (男)	99.7%	99.6%	99.6%	99.8%	99.8%	99.7%	99.6%
I (男)	99.7%	99.8%	99.8%	99.7%	99.7%	99.7%	99.7%
J (女)	97.5%	94.3%	97.6%	97.1%	97.9%	96.9%	94.3%
K (女)	98.4%	97.9%	96.0%	98.3%	97.4%	97.6%	96.0%
L (女)	93.5%	94.0%	97.4%	98.8%	98.6%	96.4%	93.5%
M (女)	99.1%	98.9%	99.1%	99.1%	99.6%	99.2%	98.9%
N (女)	99.1%	99.6%	99.5%	99.6%	99.4%	99.4%	99.1%
O (女)	99.5%	99.5%	99.5%	99.6%	99.3%	99.5%	99.3%

3-4 操作用タブレットアプリの開発によるユーザビリティ改善

3-3 で実装した話者認識アプリの音声学習は、ユーザーが学習用文章を読み上げることによって行われ、学習後にシステムが学習を適正に終了できたかを確認する必要がある。これらの作業を簡易かつ適正に実施するための音声学習ナビゲートアプリの開発を行った。これまでに話者音声特性の分析、難聴者の聴力特性（オーディオグラム）をタブレットにて入力するユーザインタフェース、入力した値をスピーカシステムに転送・設定する機能を持ったアプリ「iListening アプリ」を開発しているため、これに機能を追加する形で開発を行った。また、システム調整者がより詳細な音質調整を可能とするための話者音声特性の分析・表示、手動音質調整が可能なイコライザ調整 UI を改良し、これらの機能を含めたものを新 iListening アプリとして実装し、ユーザビリティの改善を行った。実装したアプリのスクリーンショットを図 3-4-1 に示す。



スピーカ本体(comuoon)とBluetoothで通信し、コントロールする機能を実装済み

図 3-4-1 実装した iListening アプリ

最終章 全体総括

4-1 研究履行状況について

4-1-1 当初の事業計画履行状況

1-3 成果概要に示したように、研究内容については概ね計画通り進捗した。本事業の研究開発において、多くの時間と労力がかかるとみられた難聴者による評価試験において、明瞭度評価のための語音弁別検査アプリや操作用タブレットアプリを開発して難聴者の明瞭度評価試験に使用することで、効率的・効果的に検証を実施することができた。ここで開発したアプリについては、アドバイザーから「本事業の研究開発用途のみならず、製品化することで幅広い事業で活用可能」とのアドバイスをいただいた。今後もアプリのブラッシュアップを行い、事業化のための有効なツールとして活用していく。話者音声認識技術の開発では、効率的な事業化のために適切なベースプログラムの選別を行い、ベースプログラム開発企業とも相談して今後もカスタマイズ可能なトライアル段階のプログラムを試用して研究を進めた。その結果、当初予定していたプログラムライセンス使用費用の削減にもつながった。その他、アドバイザーや川下機関ユーザーから適宜意見や情報を収集することで、市場ニーズに合った技術開発と事業化の計画を進めた。事業化の詳細については4-2に記載する。

4-1-2 目標達成度

話者音声改善技術の開発では、難聴者の被験者試験により高度難聴者の明瞭度が50%以上であることを確認し、目標を達成した。話者音声認識技術の開発では、1分以内の学習用文章の開発と話者認識プログラムを実装し、93%以上の高い精度で話者の認識が可能であることを確認し、目標を達成した。ただし、実用化のためにはより高速・高精度であることが必要であり、そのためには音声認証プログラム開発元企業と共同でシステム開発が必要であることが判明した。現在、音声認証プログラム開発元企業との検討を行っている。また、開発した話者音声認証技術に合わせた操作用アプリの開発により、ユーザビリティの改善を達成した。研究項目ごとの目標および達成度は1-3 成果概要を参照されたい。

4-2 事業化の展開と今後の課題

今回得られた研究成果および事業化における課題と現在の社会ニーズをもとに、当初の事業化計画（複数人対応 会話支援システム）を含めた3件の事業化に向けた取り組みを開始した。

4-2-1 事業化(1) 集音器タイプ補聴機器 comuoon pocket

聴こえ支援対話システム comuoon の既存顧客でも最も要望が高かった「外出先でも利用したい」というニーズに応えるために開発を行った製品であり、2-3 で開発を行った話者音声改善信号処理をボイスエンハンサープログラムとして搭載を予定している。具体的には、ボイスエンハンサープログラムのモードに「はっきりボイス（男性）」と「はっきりボイス（女性）」のモードを設定し、男性話者の時は男性のモードをセレクトすると音声が明瞭になる機能である。製品のイメージ図を図4-2-1に示す。

技術的な課題としては、開発した話者音声改善信号処理は、事前に話者の音声分析して補正値を算出する必要があり、毎回分析を行うのは不便であること、事前収録・分析が可能な機能を搭載しなければならない点

挙げられる。そこで、事前の収録・分析が必要がない補聴システムを検討している。課題の解決手法としては、多くの話者（特に聞き取りづらい話者）の音声改善可能なプリセット補正値を作る方法を検討している。2-2による音声分析結果により、話者の音声特性、補正周波数と補正量、および明瞭度の関係が把握できたため、最適ではなくても多くの話者に対して明瞭度の改善が期待できる補正値を検討することが可能である。それらの補正値で複数人の音声を信号処理した音源を用意し、被験者試験により幅広い話者の明瞭度が改善する補正値を見出すことで最適なプリセット補正値を設定することが可能である。ただし、男性と女性（声が低い人と高い人）では信号処理する周波数帯域が異なるため、男性と女性の2つのモードを設定することを予定している。

ターゲットの設定および市場については次のように見込んでいる。平成30年度の難聴者数は1500万人と言われておりこの数字は、自覚していない難聴者を含んでいる。また調査によると軽度・中等度が難聴者の89%をしめているが、補聴器の装用率は7%となっていることから軽度・中等度の難聴者をターゲットとする。5年後の目標の市場規模を15%とし、製品販売価格を100,000円と想定すると、約2,000億円の規模となる。その市場の35%を本研究による製品で獲得したいと考えている。しかしながら補聴器の浸透率を見てもわかるように補聴器店以外での製品の展開が必要だと考えており、眼鏡店や携帯ショップで気軽に貸出しができるスキームを構築中である。パートナーは日本のメガネ卸の大きなシェアを持つ内田屋との提携が決まっている。事業化スケジュールを図4-2-2に示す。すでにサンプル出荷は行っており、本年10月に製品の発表を行い11月より予約販売を行う予定で動いている。

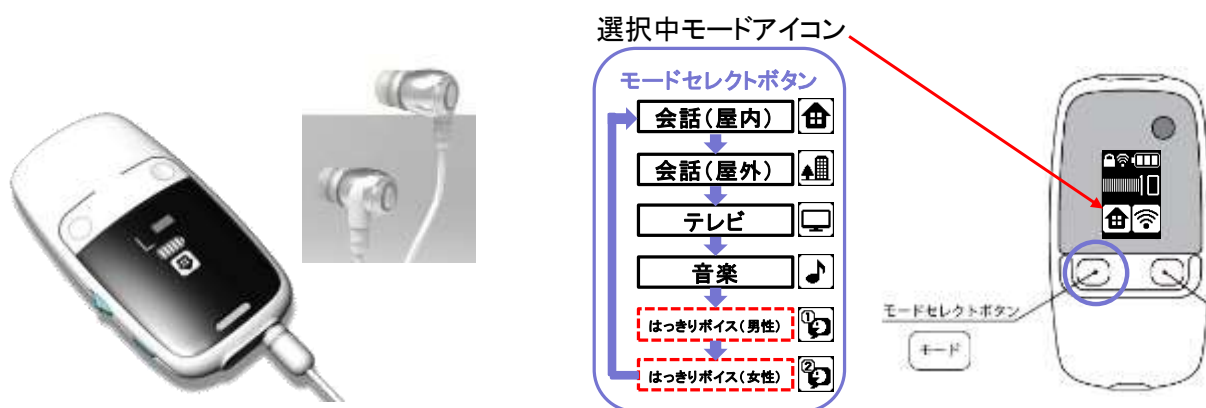


図4-2-1 聴こえ支援会話システム 集音器タイプの製品化イメージ

事業年度	H29年度	H30年度	H31年度	H32年度	H33年度	H34年度
サンプル出荷						
追加研究						
設備投資						
製品の生産						
製品の販売						

図4-2-2 聴こえ支援会話システム 集音器タイプの事業化スケジュール

4-2-2 事業化(2) 医療機器対応卓上会話支援機器と認知症リハビリテーションプログラム事業

これまでの comuoon では対応できなかった高度難聴者へ対応するための製品であり、2-1 で開発した話者音声改善信号処理と難聴者の聴力特性に合わせた補聴信号処理を組み合わせた機能を搭載する。これらの機能には難聴者の聴力特性の入力や話者音声特性の分析および補正量の算出が必要であるため、3-4で開発したタブレットアプリ「iListening」を利用する。また、これに合わせて認知症対応型共同生活介護施設との共同研究を進めており、本研究で開発を行った語音弁別検査アプリを活用した「認知症リハビリテーションプログラム」の開発を進めていく。製品のイメージ図を図 4-2-3 に示す。

ターゲットの設定および市場については次のように見込んでいる。平成 30 年度の難聴者数は 1500 万人と言われておりこの数字は、自覚していない難聴者を含んでいる。また調査によると高齢難聴者は補聴器の装用をためらうことから現在補聴器以外の支援機器の必要性が生じている。本製品は補聴器装用をためらう高齢者とのコミュニケーションを行う必要がある医療機関や福祉施設に有用であり、ターゲットとする。認知症対応型共同生活介護施設が 13,000 あり、5 年後の採用目標を 30% とし、製品販売価格を 9,800 円/月額と想定すると、約 4.5 億円の規模となる。事業化スケジュールを図 4-2-4 に示す。難聴者の聴力特性に合わせて補聴信号処理を行う補聴機器は「医療機器」となるため、事業化にあたり医療機器対応のための検証が必要となる。また語音弁別検査アプリをベースとした聴力管理システムを開発し、これを活用した認知症リハビリテーションプログラムの開発を進める。平成 31 年度にサンプルを出荷し、対象施設で検証を行い 32 年度に製品化を行う予定である。



図 4-2-3 医療機器対応卓上会話支援機器の製品化イメージ

事業年度	H29年度	H30年度	H31年度	H32年度	H33年度	H34年度
サンプル出荷				→		
追加研究			→			
設備投資		・医療機器対応のための検証 ・聴力管理システムの開発など				
製品の生産				☆7月製品発表	→	
製品の販売				☆10月販売開始	→	

図 4-2-4 医療機器対応卓上会話支援機器および認知症リハビリテーションプログラム事業化スケジュール

4-2-3 事業化(3) 複数人対応 会話支援システム（会議システム）

本事業での開発目標である難聴者が苦手とする複数人での会話コミュニケーションを可能とする製品であり、話者音声改善・話者自動識別を搭載した複数人対応 会話支援システムとなる（製品化イメージは図 1-1）。

技術的な課題としては、より高速かつ高精度な話者認識を行う必要がある。この課題解決には、認証プログラムの学習条件の最適化や、ノイズキャンセラー・エコーキャンセラー機能の追加することで、認識速度と精度の改善が期待できる。また実用化のためには、会議用マイクなどの入力デバイスの開発も必要であり、フェアリーデバイスズ(株)との共同研究開発を検討している。事業化スケジュールを図 4-2-5 に示す。

ターゲットと市場規模は、会議で使用したいというニーズが多い法人ユーザーと、高齢者同士のコミュニケーションの場である高齢者施設などを対象とし、市場規模は 4-2-2 の事業化案 2 と同等と見込んでいる。

事業年度	H29年度	H30年度	H31年度	H32年度	H33年度	H34年度
サンプル出荷						
追加研究		話者認証システムの実用化 Fairy Devices社参画				
設備投資						
製品の生産			会議用マイク試作開発	☆7月製品発表		
製品の販売				☆10月販売開始		

図 4-2-5 複数人対応 会話支援システム（会議システム）の事業化スケジュール

以上