



(2018、2019 年度 日本財団助成事業)

**「電話利用における音声認識ソフトの調査」**  
**～文字付き電話の可能性と課題～**

**報告書**

2020 年 3 月

**一般社団法人 全日本難聴者・中途失聴者団体連合会**

## はじめに

本事業は一般社団法人全日本難聴者・中途失聴者団体連合会が2018、2019年度に日本財団の助成を受けて実施したものである。本報告書の内容はすべて一般社団法人全日本難聴者・中途失聴者団体連合会に責任と著作権がある（禁無断転載）。

# 目次

はじめに

第1章	エグゼクティブサマリー	1
第2章	事業背景・概要と体制・スケジュール	4
1.	聴覚障害者と電話利用	4
2.	電話リレーサービスと日本財団のモデルプロジェクト	6
3.	難聴者の特性と文字付き電話	7
4.	文字付き電話への音声認識ソフトの活用	9
5.	事業スケジュールと実施体制	10
第3章	開発アプリケーションと実証実験の概要	11
1.	文字付き電話アプリケーションの概要	11
2.	開発中に判明した課題と対応	13
3.	実証実験概要と参加モニターの属性	14
第4章	利用者属性とアンケートデータ分析	20
1.	評価設計	20
2.	利用状況	25
3.	発信時の実用性に関する評価結果	30
4.	考察	37
第5章	実証実験後の活動と今後の方向性	38
1.	全難聴福祉大会での中間報告と実証実験後のアンケート	38
2.	電話リレーサービス制度化の動きと文字付き電話	42
3.	総括	42
資料	電話利用における音声認識ソフトの調査 会議開催状況	44

## 第1章 エグゼクティブサマリー

当事業は、聞こえにくさを持つ当事者等（以下、当事者）の電話を支援する文字付き電話での音声認識ソフトの有用性について、実証実験を通して評価することを目的としておこなわれた。

文字付き電話とは、当事者が相手先と通常電話と同様に音声で直接会話やりとりをおこない、同時に遠隔地にいる通訳オペレーター（以下、オペレーター）または音声認識ソフトが電話音声を文字に変換し利用者の端末に表示する仕組みである。

図表1 文字付き電話のイメージ



現在日本財団が2013年よりモデルプロジェクトとして電話リレーサービスの手話リレー、文字リレーをおこなっているが、手話リレー、文字リレーは手話やタイピングスキルが必要とされる。文字付き電話は手話もタイピングも不要で、相手先の電話音声を聞きながら文字を画面でみることができるものである。この特長から聞きづらくなった高齢者等も利用しやすく、大きなニーズが想定される。

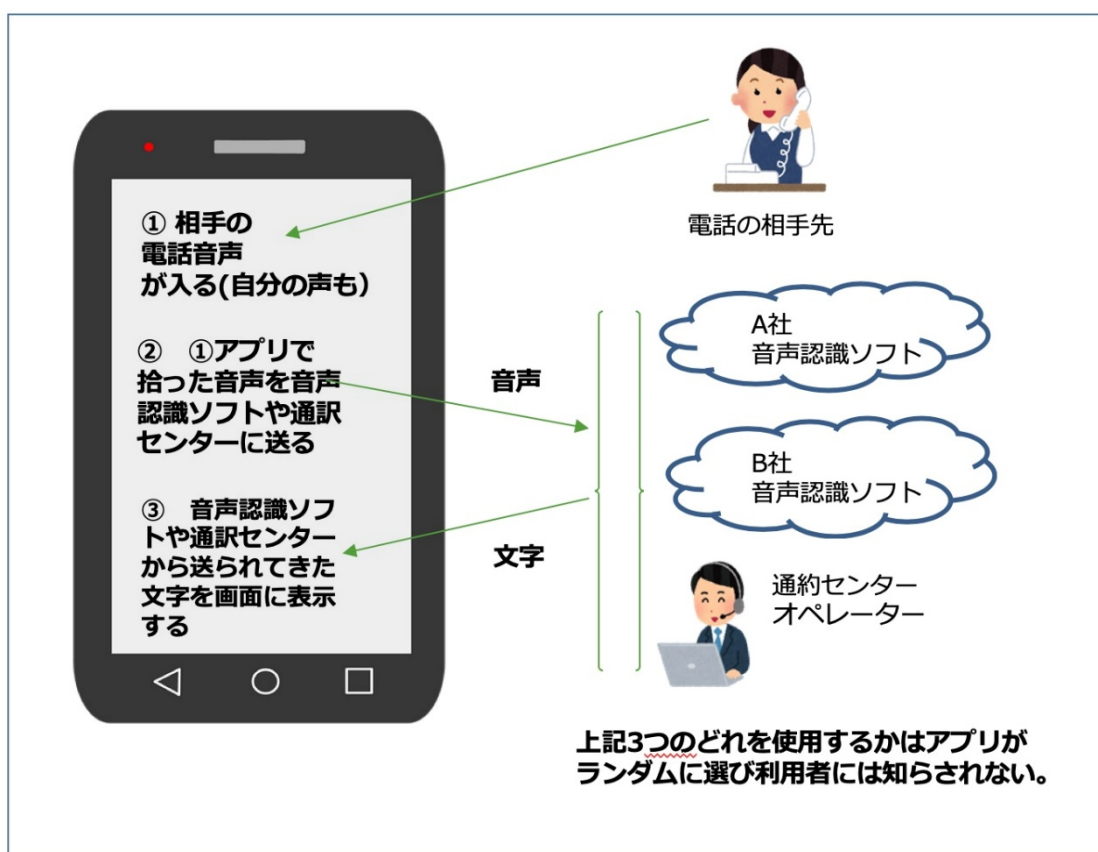
電話音声を文字化する場合、遠隔の通訳オペレーターが電話音声を聞いて入力することが通常考えられるが、オペレーターの代わりに最近進歩が著しい音声認識ソフトを活用し文字化する方法も考えられる。そのため文字付き電話では、音声認識ソフトがあればオペレーターは不要ではないかという意見もよく聞かれるようになってきた。

しかし音声認識ソフトは不安定な通信環境や不特定な相手先発音の滑舌などに認識精度が左右されやすい。そのような環境下で音声認識ソフトの電話利用が実用になりうるのかを調査した事例はない。そこで当事業では、文字付き電話を実施するとき現時点の音声認識ソフトの活用がどれだけ現実的（実用的）なのかを分析・評価することを目的とし実証実験をおこなった。

今回の実証実験では、調査環境を整えた音声認識率の数値比較（ラボ型）ではなく、社会的な実証実験としてモニターが普段の生活のなかで文字付き電話を使い、利用者としての実用性を評価する方法（フィールド型）を採用している。

実証実験用に開発した文字付き電話アプリケーションは、電話音声をスマートフォン内部でとらえその音声をクラウド上にある音声認識ソフト（2種）や通訳センターに送信、文字に変換したうえで端末に表示する。文字化の際に音声認識ソフト、オペレーター入力の選択はランダムにされ、何が選ばれたかは評価の客観性のために利用者には知らされない。利用者は電話のつど終了後に電話目的と4つの評価項目（選択型）のアンケートに答える。

図表 2 文字付き電話の仕組み

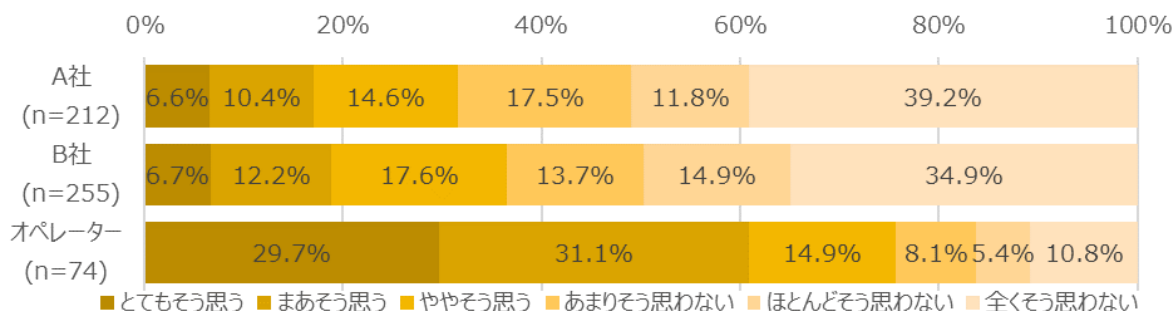


実証実験は、2018年の実験設計と環境整備（開発）、2019年4月からのモニター募集をふまえ、同年6月から10月まで行われた。実証実験中に、応募モニターのうち109人がのべ2,304回電話利用した。利用時得られたアンケートデータと登録された利用者属性データをもとに分析・考察をおこなった。

その結果、アンケートで主たる問いとした「音声認識ソフトの実用性の実感値」については、実用性の各項目（意味、ニュアンス、テンポ）すべてにおいて、音声認識

ソフトよりもオペレーターによる入力の優位性が確認された。さらに意味がある通話と仮定した1分以上の通話の満足度を分析したところ、「やや思う」までを含む肯定的評価は2種の音声認識ソフトでそれぞれ31.6%、36.5%である一方、オペレーター入力は75.7%と倍以上の満足度であった。

図表3 通話の満足度【1分以上】



何%の満足度で実用性があると判断できるかは一概に言えない。しかし今回の実験結果からは、オペレーター入力には実用性があると判断される一方、音声認識ソフトは（少なくとも今回の環境下では）音声認識ソフトのみで文字付き電話を実施した場合は利用者の十分な満足は得難いのではないかと考えられる。

しかし音声認識ソフトも肯定的評価は3割をこえており、電話の内容や相手先によっては有用であることも示されている。音声認識ソフトには表示スピードが速い、オペレーターに内容を知られないなどの長所もあり、利用者が文字付き電話で使う場面に応じてどちらも選択できる形態が最適であることを示唆している。

また今回は事後に音声認識ソフトの評価以外にも操作性などを含めたアンケートも実施したが、そのなかで「可能であれば今後も文字付き電話を利用したい」との回答はほぼ8割にもなり、当事者に適した文字付き電話に対する強いニーズが示されている。

このニーズの大きさにともなう文字付き電話での音声認識ソフトの活用については、今回の実証実験で明らかになった環境的制約を解消する研究・実証を今後も継続することが非常に重要と考える。

## 第2章 事業背景・概要と体制・スケジュール

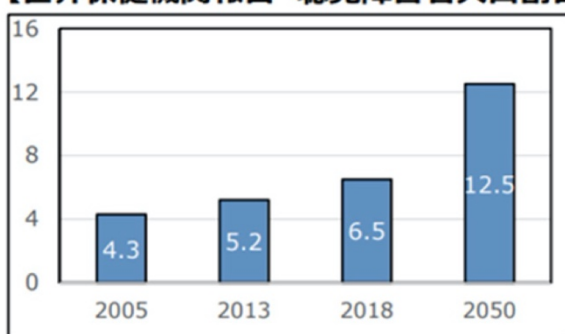
### 1. 聴覚障害者と電話利用

#### (1) 人口に占める聴覚障害者の割合

世界保健機関が報告したレポートによれば世界の人口に占める聴覚障害者の割合は2025年には12.5%になると予想されている。また日本補聴器工業会が実施したJapanTrak2018の調査によれば、聴覚に障害を感じる日本人の割合は11.3%でこれは1,400万人をこえる数値である。世界的に広がる高齢化社会がこれらの数値を押し上げている。また厚生労働省の2016年度調査によれば、障害者手帳保持者のうち聴覚・言語障害者は人口の0.3%、34万人とされる（※いずれも聴力の基準が異なることに注意が必要である）。

図表1 聴覚障害者の人口割合

#### 【世界保健機関報告 聴覚障害者人口割合 (%)】

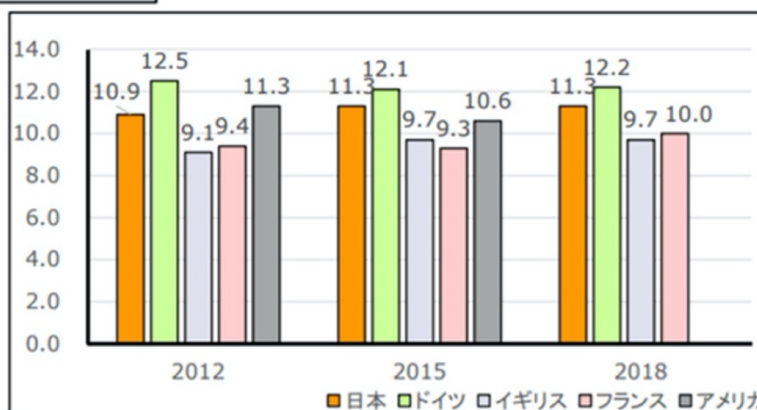


日本の身体障害者手帳（聴覚）保持者人口割合0.3%

（注）2018年世界保健機関報告は、2050年の推定として聴覚障害者の数を9億人以上としている。世界人口を70億人として人口比を算出した

#### 【JapanTrak報告 聴覚障害者人口割合 (%)】

（注）Japan Trakは日本補聴器工業会が実施。欧米データはEuroTrak調査による

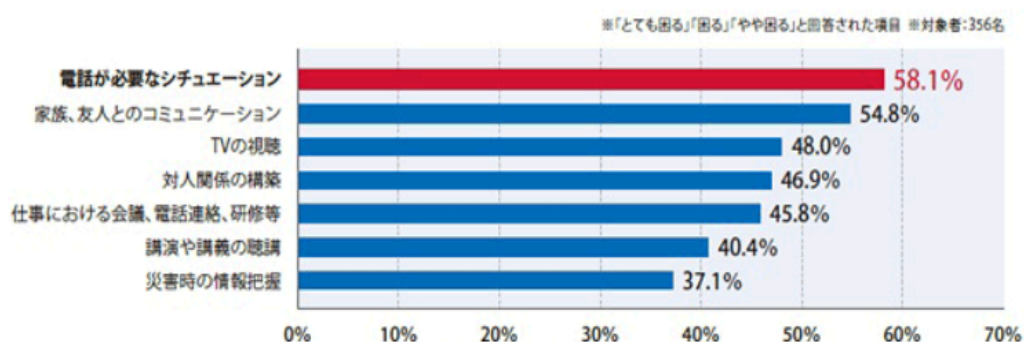


#### (2) 聴覚障害者と電話利用

聴覚障害者は毎日の生活においてどのような場面で困難や支障を感じているのだろうか？図表2は2016年にNTTドコモがおこなった調査での「聴覚障がいがあることで困ることは何ですか」という質問に対する回答サマリーであるが、第1位に「電話が必要なシチュエーション」があげられている。

図表 2 聴覚障害で困るシチュエーション

「聴覚障がいがあることで困ることは何ですか」という質問に対する回答



やむなく電話が必要なシチュエーション

- クレジットカード紛失や水もれなどトラブル発生時の問い合わせ
- 病院の予約、WEBやメールの窓口がない販売店、企業などへの連絡
- 代理電話を頼めない状況での仕事の電話

電話以外の代替手段がなく、必要性が高い状況

参考資料:NTTドコモ調べ(2016年)

電話が第1位となる理由の一つとして考えられるのは、聴覚障害者にとって相手が対面で目の前にいれば身ぶりや手ぶり、口話読取り、筆談といったさまざまな代替手段を使ってコミュニケーションを図ることも可能であるが、電話はこれらの手段がいっさい使えず情報が音声だけのために困難度が高いということがある。

聴覚障害者のうち難聴者に焦点をあててみると、難聴者には普段電話を使っていない人もいれば使っている人もいるが、それぞれが電話に対して違った悩みをかかえている。電話を使わない人には、「落とし物をして離れた場所で気がついたが電話連絡ができず、わざわざ窓口までいかねばならなかった」、「深夜に子どもが熱をだし、家に誰もいず途方にくれた」など、電話ができないことによる生活上の支障がある。一方難聴であっても電話を使う人の悩みは、「話がだいたいわかっても、確実に理解できていたかと心配である」、「固有名詞やアルファベットは何度聞きなおしてもわからない時がありお手上げになる」といった電話内容が十分把握できない、把握できていないのでは？という不安がある。

また職場などで電話が鳴り続けるときに、周りに誰もいずいたたまれない思いをすることや、家庭や職場で代理電話を依頼することに負い目を感じるなど、精神面での負担も大きい。最近はSNSの浸透で電話を使う場面は減っているが、職場や公共施設など電話が主な連絡手段であるところはまだまだ多くあり、聴覚障害者のQOLと社会参加を阻んでいる。



## 2. 電話リレーサービスと日本財団のモデルプロジェクト

### (1) 電話リレーサービスとは

前節で述べた、電話をうまく使えず生活に支障をきたしている聴覚障害者の社会参加を支援するためのサービスとして、電話リレーサービスがある。電話リレーサービスにはさまざまな形態があるが、現在主なものは手話もしくは文字を使うものである。

図表 3 電話リレーサービスの概要



具体的には手話 (ビデオ) リレーサービスは、聴覚障害者のもつスマートフォンやパソコンの画面に映る利用者の手話を、通訳センターにいるオペレーターが読み取ってその内容を相手先に通常の電話で伝え、さらに相手先からの返答を利用者に同じく手話で返すサービスである。

また、文字 (チャット) リレーサービスは、利用者がスマートフォンやパソコンに入力した文字 (テキスト) を通訳センターのオペレーターが読み、その内容を相手先に通常の電話で伝え、さらに相手先からの返答を利用者に同じくテキストで返すサービスである。

このようなサービスを利用することにより聴覚障害者もこれまで使えなかった電話を使える可能性が高まる。ただし、利用者は直接相手先と通話をしているわけではない (後述)。

### (2) 日本財団の電話リレーサービス・モデルプロジェクト

日本では過去に民間団体がこの電話リレーサービスをおこなった時期もあったが、その多くは資金的な問題で中止している。こうしたなか日本財団は東日本大震災をきっかけに、多くの国で実施されている公共インフラとしての電話リレーサービス化を目指し 2013 年秋に電話リレーサービス・モデルプロジェクトを開始した。

提供されるサービスは上記手話リレーと文字リレーであり 2019 年 10 月時点で約 10,000 人の登録、月間 28,000 回以上の利用がある。障害者手帳を持つ人は無料で利用できるが、緊急通報や夜間サービスはない。このモデルプロジェクトは現在も継続中であるが、国が 2021 年度より電話リレーサービスの公的サービス化を開始する方向で現在動いている。

### (3) 手話リレー、文字リレーのもつ課題

多くの利用者に使われ聴覚障害者の電話を支援している電話リレーサービスであるが、手話リレー、文字リレーには緊急通報や夜間サービスがないこと以外にもいくつかの課題がある。それは手話リレーの利用には手話のスキルが必要なこと、文字リレーでは会話についていけるある程度のタイピングスピードが求められることで、手話リレーは手話ができない人には使えない、文字リレーも高齢者などタイピングが得意でない人には使いづらい面がある。

さらに、手話リレーにせよ文字リレーにせよ、利用者は直接相手先と電話するわけではなく、相手先に話すのはあくまでオペレーターである。そのため、自分の意思が正確に伝わらずコミュニケーションがうまくいかないことも起こりうる。また相手との電話にでるオペレーターが不審に思われたり、電話先が銀行など金融機関の場合、(オペレーターが代わりに電話口にでるため) 本人確認が必要な場合は一部の銀行を除いて用件を達せない。このような課題が手話リレー、文字リレーにはある。

## 3. 難聴者の特性と文字付き電話

### (1) 難聴者の特性と文字付き電話

ここで視点を変えて聴覚障害者を俯瞰(ふかん)してみると、聴覚障害者には大きくわけてろう者と難聴者がある。両者の大きな違いの一つは前者が手話を主言語とするのに対して、後者は日本語を話すことといえる(もちろん、ろう者で日本語を話す人、難聴者で手話を主とする人もいる)。難聴者あるいは聞こえにくい高齢者の多くは、話すことは可能で、相手の言っている言葉さえわかれば、電話でも自分の声で相手先と直接コミュニケーションできるのである。

この特性を生かした電話リレーが「文字付き電話」であり、手話リレー、文字リレーとは異なるタイプの電話リレーサービスである。今回の事業ではこの文字付き電話に焦点をあてている。

図表 4 文字付き電話の概要



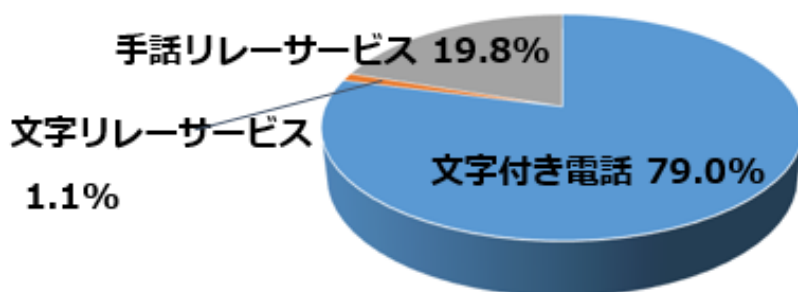
文字付き電話は、利用者は相手先と通常電話と同様に音声で直接会話し、オペレーターが相手先音声を聞いて文字に変換し、利用者の端末に表示する。利用者は手話もタイピングも必要とせず、相手先の声を聞きながら文字でもみることができる（ここが文字リレーとは大きく異なるところである）。オペレーターは両者の会話を間にたって仲介するというよりも両者の会話を裏で助ける黒子のイメージに近い。相手先に不審に思われることもなく、金融機関の本人確認も問題がない。

## (2) 文字付き電話の大きな潜在ニーズ

はじめに述べたように、日本には聴覚に障害があることを自覚している人が1,400万人以上いるが、このうち手話ができる人は推定6万人とされている。1,400万人以上いるうちの大多数をしめる高齢者と難聴者のほとんどは手話ができないが、自らの声で話すことは可能であり、この文字付き電話なら使うことができると考えられる。したがって高齢者を含む潜在的ニーズは大きく、もし文字付き電話のサービスが実現すれば、その福音は大きなものがある。

アメリカではすでに文字付き電話は公的にサービス提供されており、急速な利用の伸びにより、最近では全電話リレーサービス利用時間の8割近くを占めている。使い勝手のよさと高齢化社会の広がりが背景にある。

図表5 アメリカの電話リレーサービスの時間構成



## 4. 文字付き電話への音声認識ソフトの活用

### (1) 文字付き電話における二つの文字変換の方法

文字付き電話で音声文字化手段は通常オペレーターによる手入力だが、最近進歩が著しい音声認識ソフトを活用して音声を文字に変換する方法も考えられる。

図表 6 文字付き電話のオペレーター入力と音声認識ソフト利用



オペレーターが音声を聞いて文字入力する方法の長所は、多くの場合音声認識ソフトより正確な内容が表示されることであるが、オペレーターの人件費が必要となる。一方音声認識ソフトは表示スピードが速く、プライバシーが守れること、オペレーターの人件費が不要などのメリットがあるが、変換された内容の正確度は現時点でオペレーター入力に比べ低めである。

オペレーターによる入力はアメリカで Caption Telephony Service (CTS) として実施している。アメリカの CTS では基本はオペレーター入力、音声認識ソフトはオペレーターが相手先音声を聞いてそれをリスピーク（復唱）する形で使われている（相手先の音声を直接音声認識ソフトに変換はしていない）。オペレーターによる文字入力での文字付き電話は過去に日本財団でも提供されたことがあるが、現在はおこなわれていない。一方の音声認識ソフトを活用した文字付き電話は、2019年3月より NTT ドコモが「みえる電話」として NTT ドコモの契約者に無料で提供している。

### (2) 電話リレーサービスへの音声認識の活用と当事業の目的

最近の音声認識ソフトの技術の進歩は目覚ましく、みえる電話のようにこの技術を活用する文字付き電話も始まっている。音声認識ソフトを活用することで、コストをおさえた文字付き電話のサービス実現への期待も大きい。そのため最近、音声認識ソフトがあるからオペレーター入力はもう不要なのではないかという声もよく聞かれる。しかし電話の利用時には音声認識ソフトの他の利用時とは違う通信の状況や周りの騒音、相手先音声の明瞭度など、音声認識ソフトにとって厳しい環境にさらされる。これらが認識精度に影響を与える。

しかし現在の音声認識ソフトが電話での厳しい利用環境のなかでどこまで利用できるのか調査した事例はない。そこで当事業では、現段階で音声認識ソフトでの電話利用が実際のところどれだけ実用的なのかを調査することを目的に、実証実験をおこな

うこととした。その際、調査環境だけを整えた単純な音声認識率などの調査比較（ラボ型）ではなく、社会的な実証実験として、モニターに実生活のなかで文字付き電話を使ってもらい、利用者としての実用性を電話のつど評価してもらう方法（フィールド型）を採用した。あわせて、音声認識ソフトとオペレーター入力との比較もおこない分析・評価を行った。その結果を第4章で述べる。

## 5. 事業スケジュールと実施体制

当事業は2018、2019年度の2年間、以下のスケジュールで実施された。

2018年8月～ 実証実験設計と環境整備（開発）

2019年4月～ モニター募集

2019年6月～10月 実証実験実施（5ヵ月）

2019年11月 全国中途失聴者・難聴者福祉大会 in 滋賀 中間報告

2020年3月 報告書完成

事業の実施に当たっては、実証実験の設計、データの分析・評価に三菱UFJリサーチ&コンサルティング株式会社、実証実験の環境整備、モニター募集、実証実験運用に合同会社シーコミュの協力を得た。また、共用品推進機構の星川安之氏を委員長とする、「電話利用における音声認識ソフトの調査 事業委員会」（巻末名簿参照）を設け、事業内容に対するアドバイスを受けている。

## 第3章 開発アプリケーションと実証実験の概要

### 1. 文字付き電話アプリケーションの概要

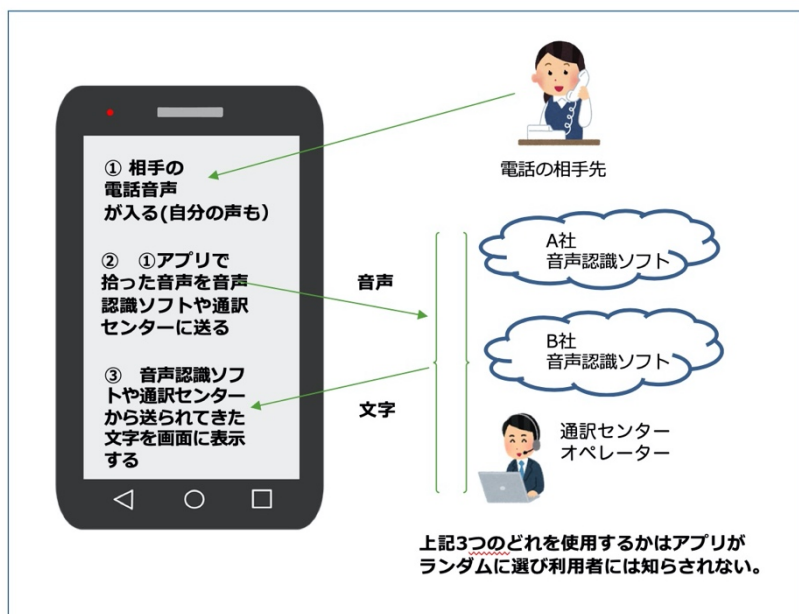
#### (1) 文字付き電話アプリケーションの主な仕様

第2章で述べたように、電話音声文字化の方法としてオペレーター入力と音声認識ソフトがあるが、実証実験のためにどちらでも電話音声を文字表示できるアプリケーションを開発した。主な仕様は以下の通りである。

- ・スマートフォンの電話音声を聞くことも文字でみることもできるアプリケーション
- ・Androidバージョン6以上のスマートフォンで利用可能
- ・契約電話会社を問わず利用可能
- ・新たに電話番号を取得する必要はなく今の携帯電話番号がそのまま使える
- ・かけ先に制限はない（ただし緊急通報 110、118、119 は文字表示されない）
- ・発信だけでなく着信も可能
- ・かけ先、利用者音声とも文字表示される
- ・音声認識ソフトは世界的規模の IT 企業 2 社（以下、A 社、B 社とする）のサービスを利用

#### (2) 文字付き電話アプリケーションでの音声認識の仕組み

図表 1 文字付き電話の仕組み



当アプリケーションは、電話音声をスマートフォン内部でとらえ、その音声をサーバー上にある音声認識ソフトや通訳センターに送り、音声認識ソフトもしくは通訳センターのオペレーターにより文字に変換、そのうえで端末に表示する。電話音声を文字化するにあたってはそのつどアプリケーションが音声認識（A

社またはB社）とオペレーター入力をランダムに選択し、利用者が選ぶことはできない。またこの時、利用者にバイアスを与えないため、利用者には3種のうちどれが選択されたかはわからないように設計している。音声認識は24時間使用可能だが、オペレーター入力は午前10時から午後6時までとした。

### (3) アプリケーション画面の流れと利用時アンケート

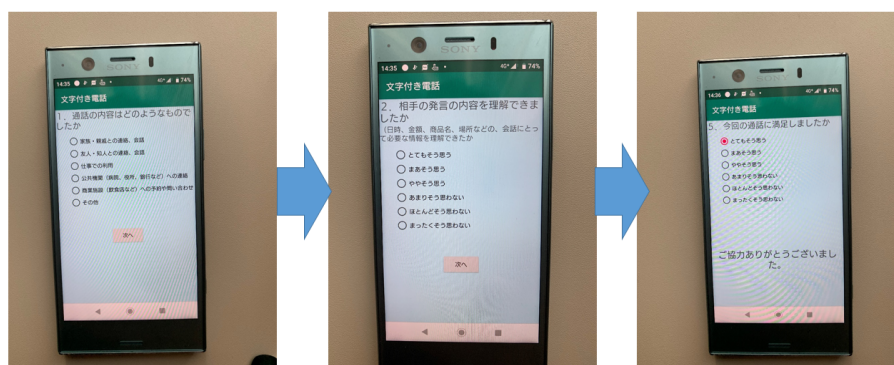
利用者のアプリケーション画面の流れ（発信の場合）は以下のとおりである。

図表 2 文字付き電話 画面の流れ



アプリケーションをインストールするとホーム画面にアイコンが表示される。

- ①通常電話のアイコンのかわりに文字付き電話アイコンをタップする
- ②ロゴ画面が表示される
- ③ダイヤル画面に変わる
- ④電話番号の入力、電話帳、通話履歴のいずれかから相手先を特定し発信すると、通話画面にかわり通話内容が文字表示される



- ⑤通話画面で終了ボタンを押す、あるいは相手が電話を切断するとアンケート画面に変わる。簡単な5つの質問（選択型）に答えるとアプリケーションが終了する。

## 2. 開発中に判明した課題と対応

### (1) Android のバージョンアップにより電話音声の取得が不可に

従来から iOS (iPhone など) はアプリケーションでの電話 API (Application Program Interface) を提供しておらず、スマートフォン本体のみで電話音声は取得できない。一方、Android はこれまで電話音声 API を提供していたことから今回は Android での開発をすすめていたが、その Android もバージョン 6 以降 API 提供を中止していたことが判明したため、iOS でも Android でもスマートフォン本体のみで電話音声を取得することが極めて困難となった。

### (2) 外付けアダプターでの対応

図表 4 文字付き電話の仕組み



この状況を打破するために代替策を検討した。具体的にはマイクから入る音であれば音声認識ソフトや通訳センターに転送することはできるので、イヤホンに流れる電話音声をマイクに戻す外部アダプターを製作し、その音声を音声認識ソフトや通訳センターに転送するデザインに変更した。あわせて音声を利用者に届け、利用者の音声をスマートフォンに入力するためのイヤホンマイクもつなぐ必要がある。この使用形態は利用者にわかりづらく、ケーブルを経由することで音質の劣化も起こるため実証実験にも影響を与えた。

その後実証実験期間中に、人工内耳の場合イヤホンマイクが使いづらいという意見もあり、電話音声をスピーカーフォンに出し、その音を本体マイクで拾う対応もとった。この方法ではアダプターは不要になるが、周りに人がいる時は使いづらいことや音質も劣化するため本質的な解決にはなっていない。



### 3. 実証実験概要と参加モニターの属性

#### (1) 実証実験モニター募集とモニター属性

①実証実験のモニター募集期間は2019年4月1日～9月30日である。応募資格は「発声ができ聞こえに不自由を感じているAndroid（バージョン6.0以上）スマートフォンをお持ちの方で、電話ごとのアンケートにお答えいただける方」とした。参加費用は無料であるが、電話料金、通信料金はモニターの負担である。全難聴ホームページでの告知、加盟団体での紹介、ポスターの作成などで勧誘をはかった。その結果、期間中のモニター申込数は283人であった（一部重複など含む）。

②登録属性として、都道府県、年齢、性別、世帯状況、職業、補聴器使用の有無、聴覚障害の程度、語音明瞭度、電話利用の有無を質問した。以下に概要を紹介する。

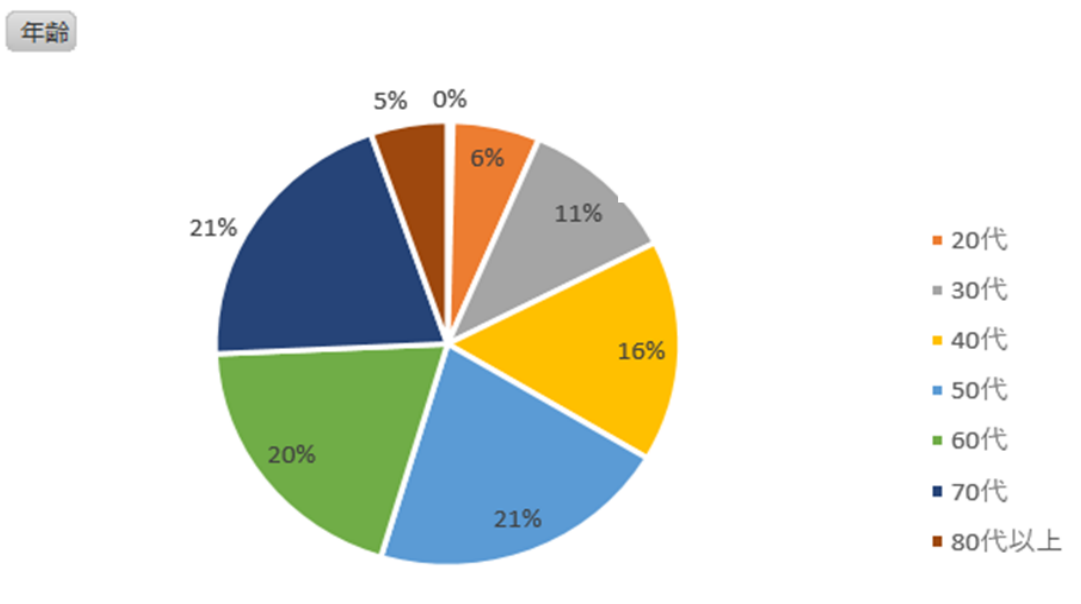
- ・41都道府県から申し込みがあり、東京、神奈川、兵庫が上位3位を占める。

図表5 登録モニターの都道府県別人数

都道府県	申込数								
東京都	79	岡山県	6	福井県	4	長野県	1	秋田県	1
神奈川県	33	熊本県	6	栃木県	3	青森県	1	合計	283
兵庫県	19	宮崎県	5	滋賀県	3	大分県	1		
愛知県	13	静岡県	4	群馬県	3	徳島県	1		
大阪府	11	富山県	4	京都府	3	島根県	1		
埼玉県	11	奈良県	4	香川県	2	岩手県	1		
千葉県	10	山梨県	4	茨城県	2	宮城県	1		
北海道	10	広島県	4	山口県	2	岐阜県	1		
福岡県	8	鹿児島県	4	愛媛県	2	福島県	1		
三重県	8	石川県	4	新潟県	1	鳥取県	1		

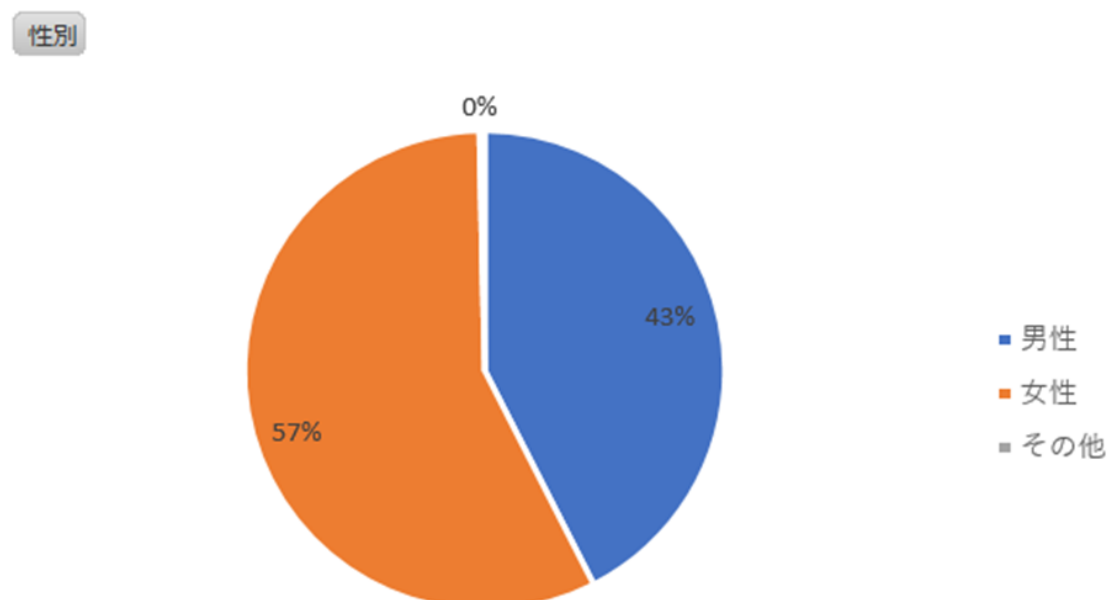
・年齢は10代（少数）から80代に幅広く分布し、50代、60代、70代がそれぞれ20%前後で拮抗している。

図表6 登録モニターの属性



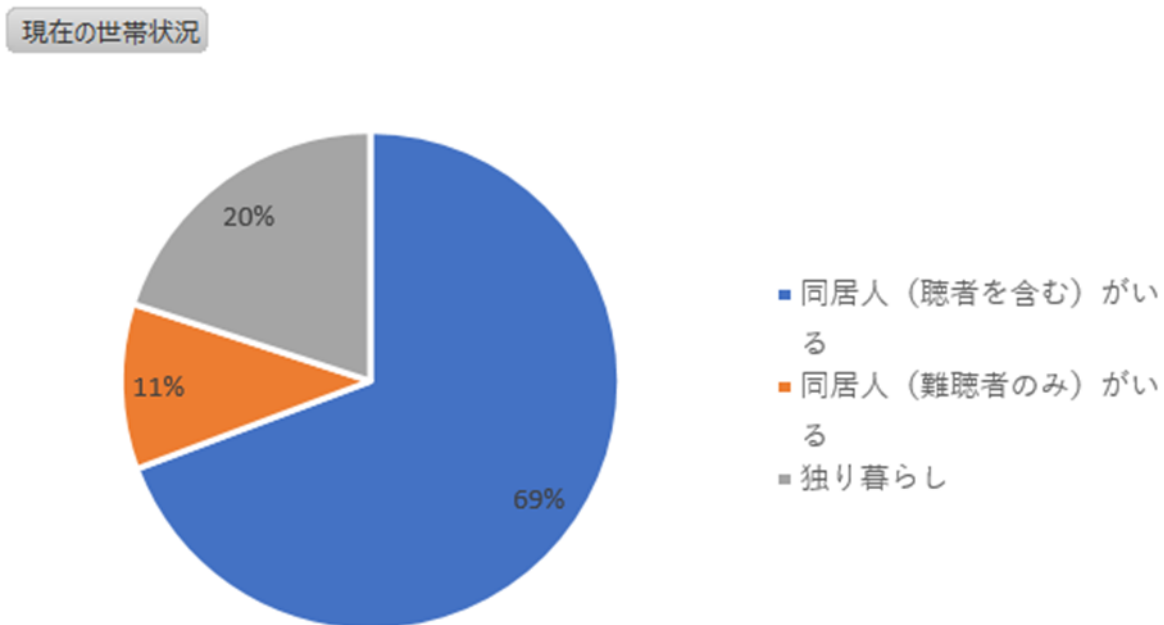
・性別では、女性57%、男性43%と女性の申し込みが多い。

図表7 登録モニターの性別



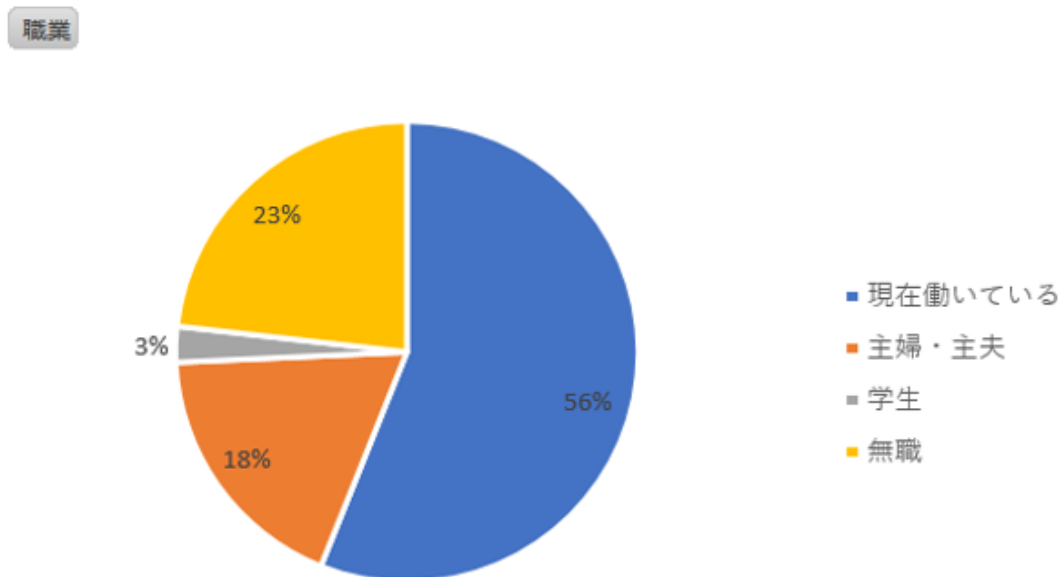
- ・世帯状況は聴者を含む同居人がいる世帯が 2/3 を占めている。

図表 8 登録モニターの世帯状況



- ・現在働いている人が 56% を占めている。職場で電話が必要となる時も多いと思われる。

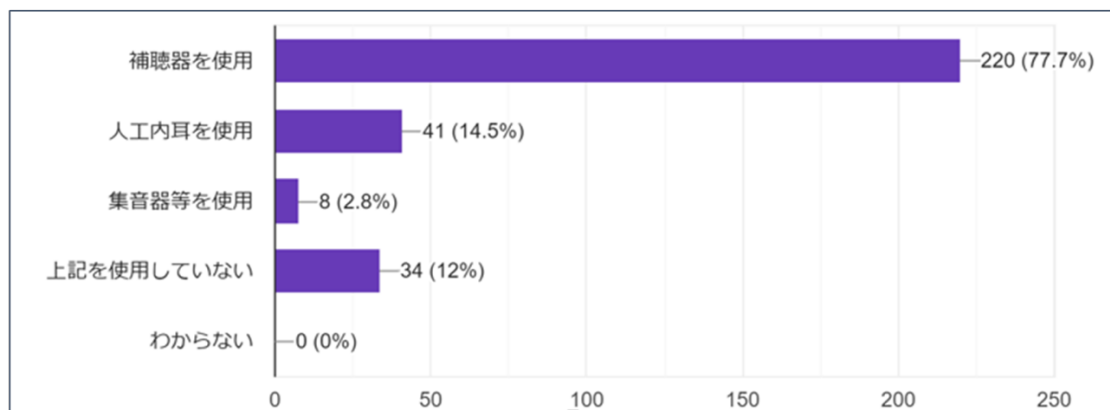
図表 9 登録モニターの職業



・補聴器は78%の人が使っている。人工内耳使用者も15%あり補聴器と併用している人が多いことが想定される

図表 10 登録モニターの補聴器使用の有無

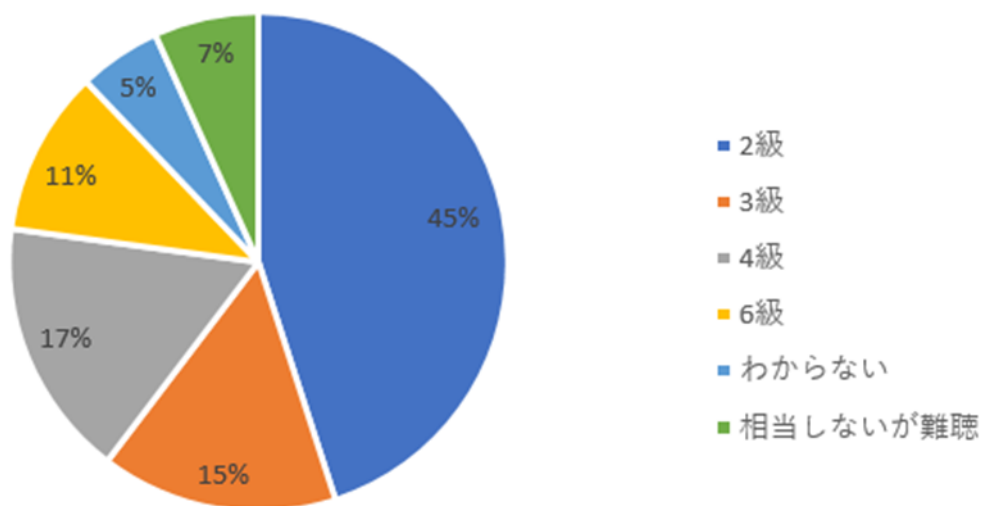
補聴器使用の有無(複数選択可)



・聴覚障害の程度は2級が最多で45%であり、障害者手帳保有者が88%を占める。申込者の主体が全難聴加盟協会の会員であるためと思われる。

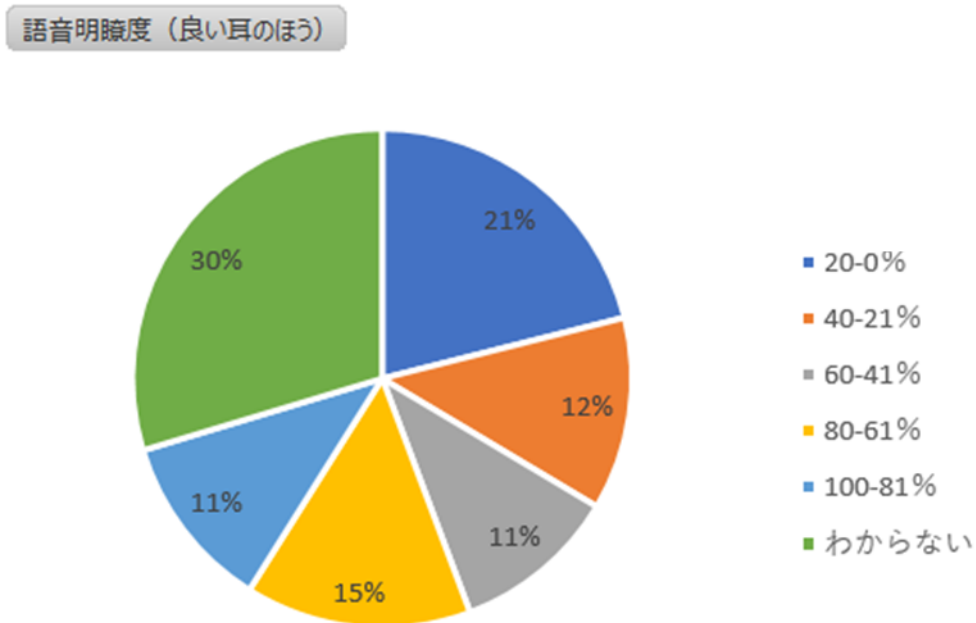
図表 11 登録モニターの聴覚障害の程度

聴覚障害の程度



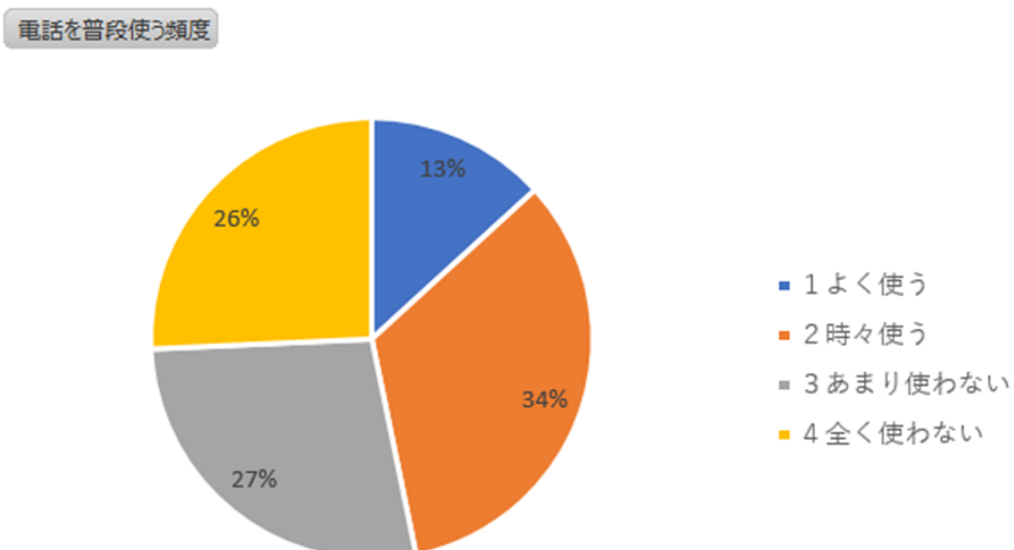
・ 語音明瞭度は軽度から重度まで広範囲に渡っている。語音明瞭度は音声の判別に大きな影響をあたえるが、その程度によって文字付き電話に対する期待内容が異なることが考えられる。

図表 12 登録モニターの語音明瞭度



・ これまで電話を使ってきたかの質問について、47%が「よく使う」、「時々使う」と答えており、これまで電話を使ってきている人たちも音声に加え文字情報が表示されることに期待していることがうかがえる。

図表 13 登録モニターの電話を普段使う頻度



実証実験時、文字付き電話利用のつどとられたアンケートデータとこれらの属性を紐づけ、電話利用における音声認識ソフトの実用性分析をおこなった。その結果は次章で述べる。

## (2) 実証実験の運用と通話件数

① 実証実験期間は2019年6月1日～2019年9月30日としたが、より多くデータを収集するために10月31日まで延長し5ヵ月間とした。期間中は土曜、日曜、祭日を含む毎日利用可とし、音声認識ソフト利用は0:00～24:00（24時間）オペレーター入力10:00～18:00（最終受付17:50）とした。オペレーター入力がない18:00～翌10:00の間は、2種の音声認識ソフトのどちらかが選ばれる。オペレーターは日本財団のモデルプロジェクトで文字リレーを委託されている沖縄のアイセック・ジャパン社の協力を得て、同社の通訳センターに配置した。

② 実証実験期間中の通話総件数は2,943件であった。内訳は以下の通りである。オペレーター入力の件数が少ないのは、夜間は音声認識ソフトのみになることによる。不成功も含めた発信、着信の割合はそれぞれ88%、12%であった。

A社音声認識ソフト	1,032件
B社音声認識ソフト	955件
オペレーター入力	317件
相手先出ずなど不成功	639件
発信	2,578件
着信	365件

## 第4章 利用者属性とアンケートデータ分析

本実証は、難聴者にとって文字付き電話が有用であるかを明らかにするために実施するものであり、ここでは、プログラム評価の5階層（ニーズ評価、セオリー評価、プロセス評価、インパクト／アウトカム評価、効率性評価）の中でもアウトカム評価にあたる部分である。なお、難聴者のニーズやその解決に向けたセオリーについては本報告書第2章、文字付き電話のサービス実施プロセスについては本報告書第3章で言及している。

### 1. 評価設計

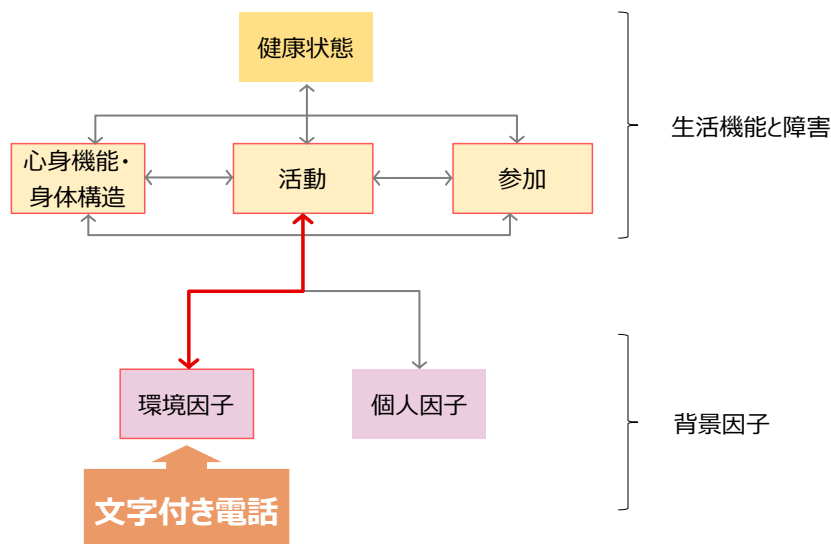
#### (1) 考え方の枠組み

評価とは一般に「事実特定＋価値判断」（日本評価学会認定評価士養成講座）とされており、評価を設計するにあたっては、事実特定のための枠組みと、価値判断のための着眼点が重要となる。

文字付き電話は、聴覚障害者がより簡便に電話を使いたいというニーズを解決する目的のサービスである。これをICF（国際生活機能分類）の構成要素として捉えると、背景因子のうち環境因子にあたり、生活機能と障害の面では主に活動を促すものと位置付けられる。そのため、事実特定の観点では、サービスの提供内容や利用状況を詳細に把握する必要がある。また、価値判断の観点では、「文字付き電話を通じて電話をする」という活動が、難聴者にとってどのような意義を有しているかについて分析することが求められる。

なお、背景因子のうち個人因子や、生活機能と障害のうち心身機能・身体構造もこれらに大きく影響することが見込まれるが、アウトカム評価を実施する上では、その影響を統制して厳密な評価を行うことが重要となる。

図表1 ICF構成要素でみた本実証の考え方の枠組み



## (2) 基本的な仮説

考え方の枠組みでみたように、文字付き電話は難聴者の活動や参加に影響を及ぼすと考えられる。この点について、ICFの「活動と参加」の大分類には様々な項目があるが、文字付き電話では、例えば以下のように「一般的な課題と要求」「コミュニケーション」「家庭生活」「対人関係」「主要な生活領域」「コミュニティライフ・社会生活・市民生活」といった項目で、難聴者の生活上の活動や参加に前向きな効果がもたらされると想定されている。

図表 2 本実証における基本的な仮説

ICFの大分類	具体的な例
一般的な課題と要求	<ul style="list-style-type: none"><li>・生活上の手续や契約を円滑に進められる</li><li>・急ぎの要件にも迅速に対応できる</li></ul>
コミュニケーション	<ul style="list-style-type: none"><li>・聴者が話している際のニュアンスを感じ取れる</li></ul>
家庭生活	<ul style="list-style-type: none"><li>・周囲の人に代理電話を依頼して申し訳なく思わずにすむ</li></ul>
対人関係	<ul style="list-style-type: none"><li>・友人・知人と楽しく意思疎通できる</li><li>・初めて会う人と率直に意思疎通できる</li></ul>
主要な生活領域	<ul style="list-style-type: none"><li>・工作上必要な意思疎通を円滑に進められる</li></ul>
コミュニティライフ・社会生活・市民生活	<ul style="list-style-type: none"><li>・人付き合いや一人での外出を避けずにすむ</li></ul>

## (3) 仮説を検証するための問い

文字付き電話を用いた本実証の特徴の1つとして、難聴者が電話をしている相手先の発話音声を音声認識ソフトで文字に変換するプロセスが取り入れられていることがあげられる。周知の通り、音声認識ソフトは様々な企業が技術開発を続けており、日進月歩で精度の改善が図られているものの、サービスに組み込めるほどの十分な実用水準に達しているかについては懐疑的な意見もある。そのため、文字付き電話が音声認識ソフトの組み込みを前提とした場合に、前項「基本的な仮説」で示したような前向きな効果が得られるかを検証する必要がある。

この「実用性」については、客観的な評価尺度で測定する方法と、主観的な評価尺度で測定する方法の2種類に大別できる。客観的な評価尺度としては、音声認識の精度テスト(例:音声認識ソフトが様々な環境下で音声を正確に文字化できるか)などが想定できるが、音声認識の精度は文字付き電話というサービスを構成する重要な要素であるものの、実際の通話では会話のシチュエーションや通話先の反応なども実用性に大きな影響があると考えられる。本実証は当事者である全難聴の実施事業ということもあり、サービス利用者である難聴者の主観的な評価尺度で実用性を評価することとした。すなわち、難聴者が文字付き電話のサービスの実用性についてどのように実感したかを評価尺度として用いる。

主観的な評価尺度を用いる場合には、回答者の価値基準が異なることから個人差が結果に反映されやすい側面がある。そこで、本実証では音声認識ソフトによる文字化だけでなく、既に通話として一定の品質水準が確認されているオペレーターによる文



字化も併用し、それらの比較により実用性を明確化しようと試みた。

以上でみてきたように、本実証の仮説を検証するための問い（リサーチ・クエスチョン）を定式化すると、下記のように整理できる。

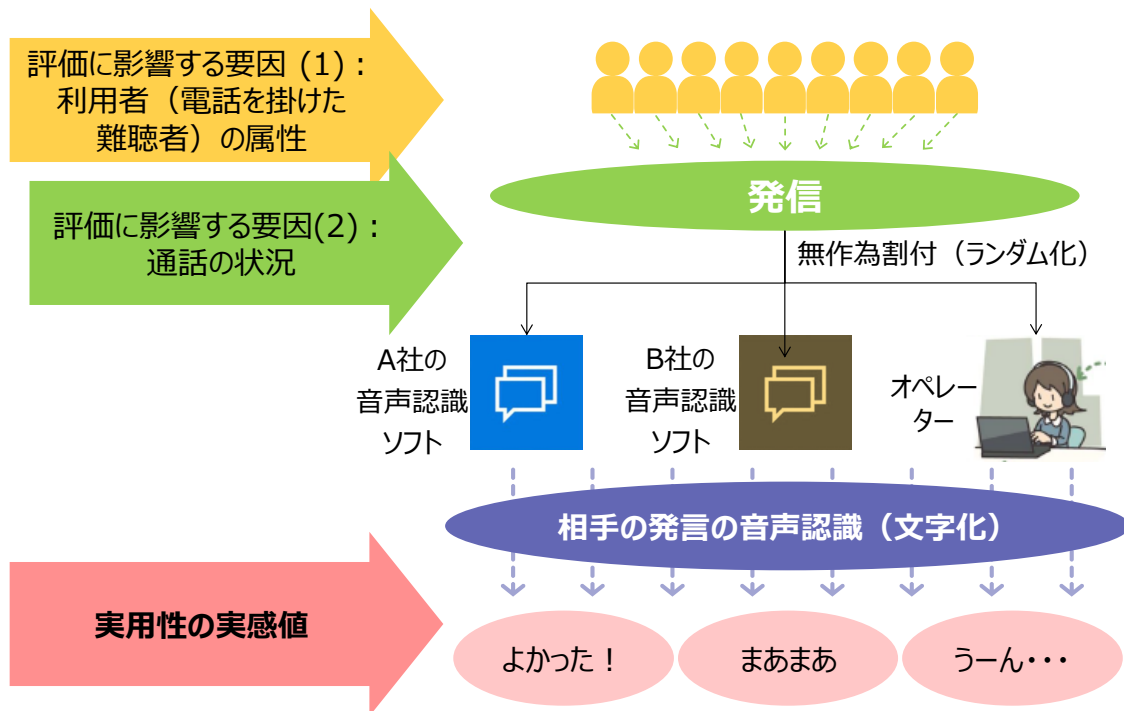
図表3 本実証における仮説を検証するための問い

課題の定式化	本実証への当てはめ
どのような対象について	難聴者は
どのような介入を受けると	音声認識ソフトを利用すると
どのような介入と比較して	オペレーターを利用した場合に比べて
何がどのように、どれくらい変わるか	利用満足度がどう変わるか

#### (4) 実証デザイン

本実証のアウトカム評価は、アウトカムの変化を厳密に測定できるランダム化比較試験（Randomized Controlled Trial：RCT）の手法で実施した。RCTは、複数の介入を無作為割付（ランダム化）することに特徴がある。本実証では難聴者による実用性の評価に影響する要因として、利用者（難聴者）の属性（ICFでいう個人因子）や通話の状況（シチュエーション、通話時間／等）が考えられるが、本来、これらによらず難聴者の利用満足度が高いのは音声認識ソフトとオペレーターのどちらか、ということが検証したい事項である。RCTでは、これらの影響を統制（コントロール）して分析することが可能となる。

図表4 実証デザインの概要



#### (5) 分析に用いたデータ

本実証の分析は、利用者（難聴者）の属性や通話の状況が、実用性の実感値にどのよ

うに影響したかを明らかにするものである。なるべく多くの項目を分析に用いることがより正確な分析結果の導出につながるものの、実証としての倫理性、データの入手可能性や正確性、及び委員会での検討を踏まえ、以下の項目のデータを収集することとした。

なお、実用性の実感値については、電話の品質評価で一般的に用いられる3つの尺度（内容が理解できたか、ニュアンスが理解できたか、テンポよくやり取りできたか）を採用し、これに通話の満足度を加え、利用者に各項目を6段階で評価してもらった。これら及び通話先（6分類）は、利用直後アンケートとして利用のつど、アプリにアンケート形式で設問が表示されるようにした。

**図表5 実証結果の分析に用いたデータ項目一覧**

分類	データ項目・内容	入手方法
利用者の属性	年齢 <sup>1</sup> （10代/20代/30代/40代/50代/60代/70代/80代）	実証参加登録情報
	性別（女性/男性/その他）	
	世帯（同居人（聴者を含む）がいる/同居人（難聴者のみ）がいる/独り暮らし）	
	職業（現在働いている/主婦・主夫/学生/無職）	
	補聴器等の使用状況（補聴器を使用/人工内耳を使用/補聴器と人工内耳を使用/補聴器と集音器等を使用/使用していない）	
	聴覚障害程度（2級/3級/4級/5級/6級/相当しないが難聴/わからない）	
	語音明瞭度（20-0%/40-21%/60-41%/80-61%/100-81%/わからない）	
	電話の習慣（よく使う/時々使う/あまり使わない/全く使わない）	
通話の状況	音声認識の種類（A社/B社/オペレーター）	通話記録
	文字付き電話利用回数（1回目/2回目/3回目…）	
	通話時間（1秒/2秒/3秒…）	
	バージョン（新バージョン/旧バージョン）	
	通話先（家族・親戚/友人・知人/仕事関係/公共機関/商業施設/その他）	利用直後アンケート
実用性の実感値	内容の理解（とてもそう思う/まあそう思う/ややそう思う/あまりそう思わない/ほとんどそう思わない/全くそう思わない）	
	ニュアンスの理解（とてもそう思う/まあそう思う/ややそう思う/あまりそう思わない/ほとんどそう思わない/全くそう思わない）	
	テンポのよさ（とてもそう思う/まあそう思う/ややそう思う/あまりそう思わない/ほとんどそう思わない/全くそう思わない）	
	通話に満足したか（とてもそう思う/まあそう思う/ややそう思う/あまりそう思わない/ほとんどそう思わない/全くそう思わない）	

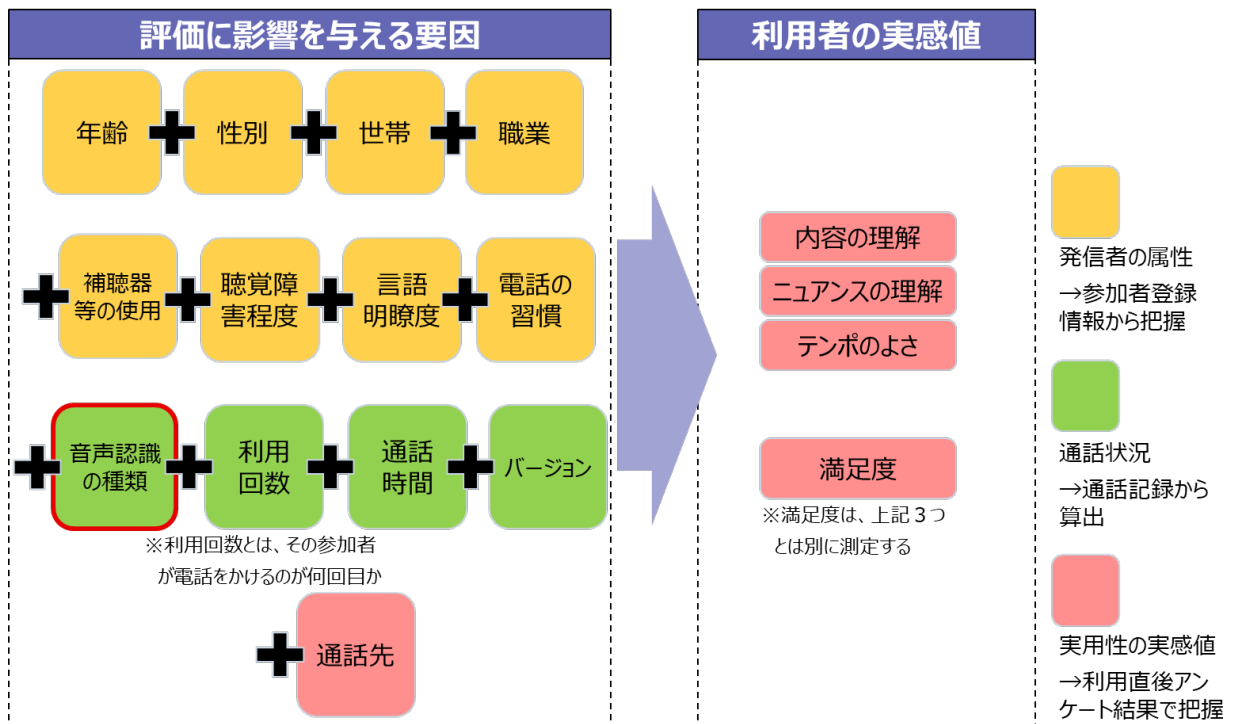
<sup>1</sup> 10代の登録者は1名だったが、実証期間内での利用はなかった（0回）。

図表 6 利用直後アンケート

No.	質問項目	選択肢
1	通話の内容はどのようなものでしたか	①家族・親戚との連絡、会話 ②友人・知人との連絡、会話 ③仕事での利用 ④公共機関（病院、役所、銀行など）への連絡 ⑤商業施設（飲食店など）への予約や問い合わせ ⑥その他
2	相手の発言の内容を理解できましたか (日時、金額、商品名、場所などの、会話にとって重要な情報を理解できたか)	①とてもそう思う ②まあそう思う ③ややそう思う ④あまりそう思わない ⑤ほとんどそう思わない ⑥全くそう思わない
3	相手の発言のニュアンスが理解できましたか (相手がどのような感情で発言しているかなどを理解できたか)	
4	テンポよくやり取りができましたか (相手の発言に対する返答がスムーズにできたか)	
5	今回の通話に満足しましたか	

※利用者が利用のつど、5問それぞれで選択肢から1つだけ選ぶ形式とした

図表 7 データ分析モデル



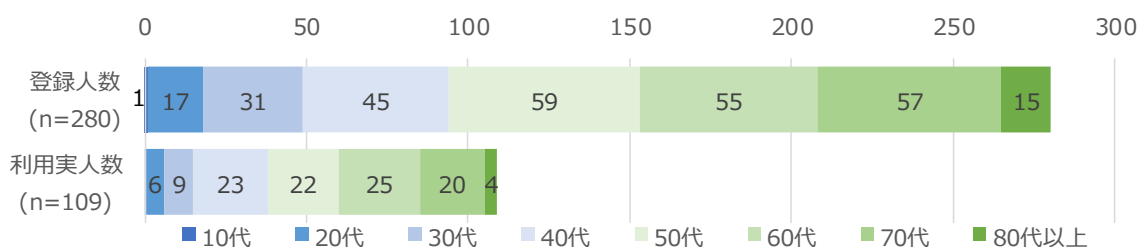
## 2. 利用状況

### (1) 利用の実人数とのべ回数

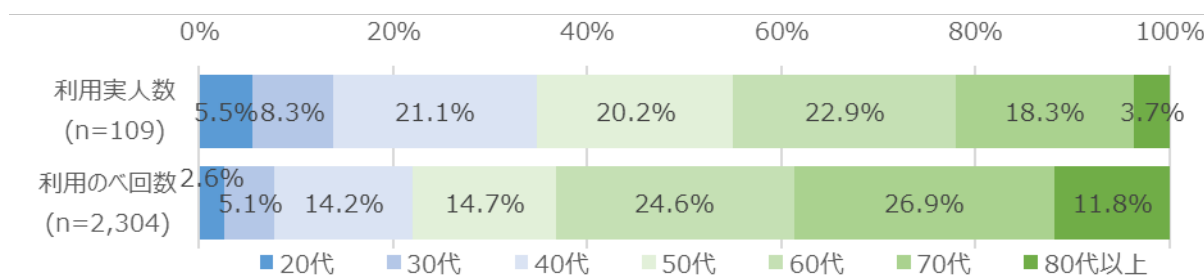
通話記録（個人が特定できない形で記録したログデータ）から文字付き電話の利用状況を集計したところ、登録者（重複等を除く実人数 280 人）の約 4 割に当たる 109 人が、のべ 2,304 回利用していた（通話失敗を除く）。

この 109 人について、利用のべ回数を年代別にみると、70 代が 26.9%、80 代が 11.8%と、70 歳以上の利用割合が高い傾向がみられた。

図表 8 年代別 登録人数と利用実人数



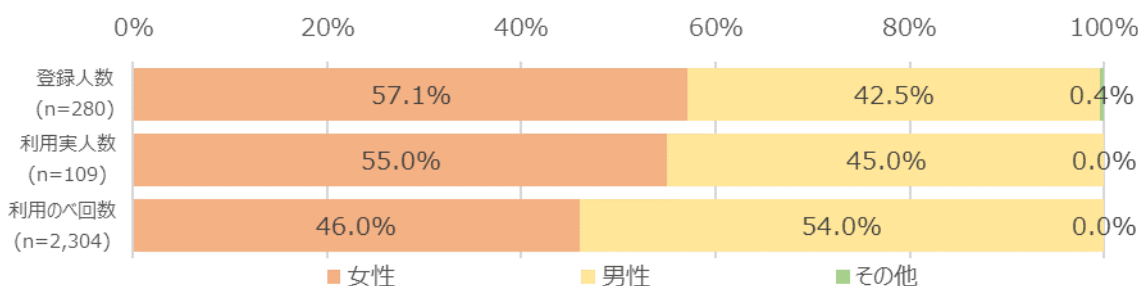
図表 9 年代別 利用実人数と利用のべ回数



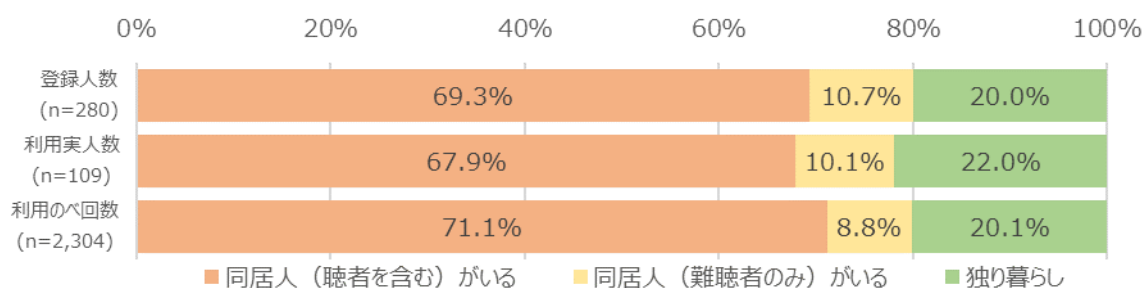
### (2) 利用者の属性

利用者の属性を実績ベースで見ると、登録者ベースの割合と比べて「男性」の割合がやや高かった（女性 46.0%に対して男性 54.0%）。世帯状況では大きな違いはみられなかった。就業状況では、「無職」の割合がやや高かった（32.3%）。電話の習慣では、電話を比較的「よく使う」「時々使う」人の利用割合がやや高かった（それぞれ 18.6%、31.3%）。

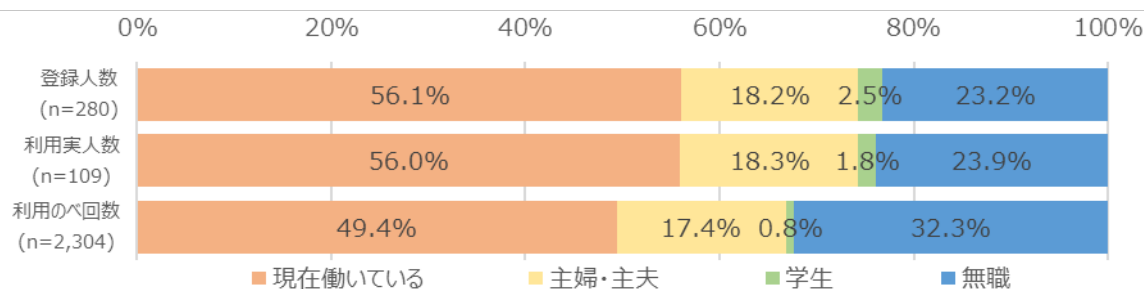
図表 10 利用者の性別



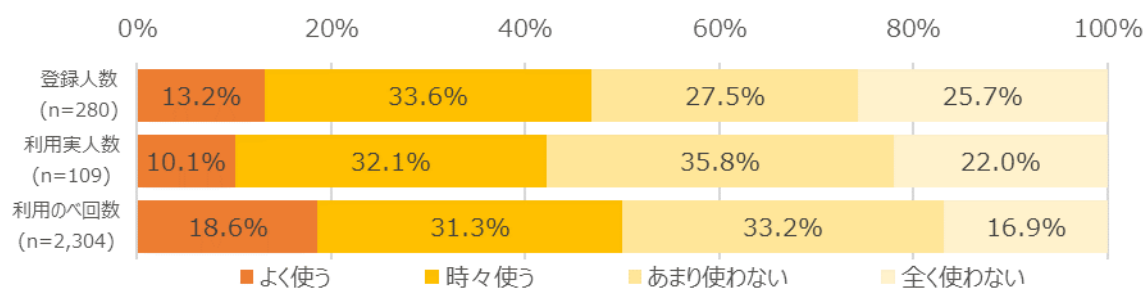
図表 11 利用者の世帯状況



図表 12 利用者の就業状況



図表 13 利用者の電話の習慣

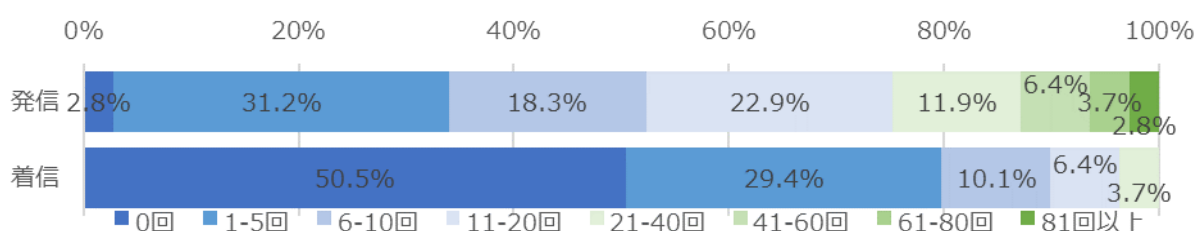


(3) 一人あたり利用回数

109人の利用者について、発信による利用回数は、1～20回以下が全体の7割以上を占めた（75.2%）。なお、発信が最も多かった人は、6～10月で合計134回利用していた。

受信による利用回数は「0回」の割合が約半数を占めており（50.5%）、受信が最も多かった人は、6～10月で合計33回だった。

図表 14 一人あたり利用回数

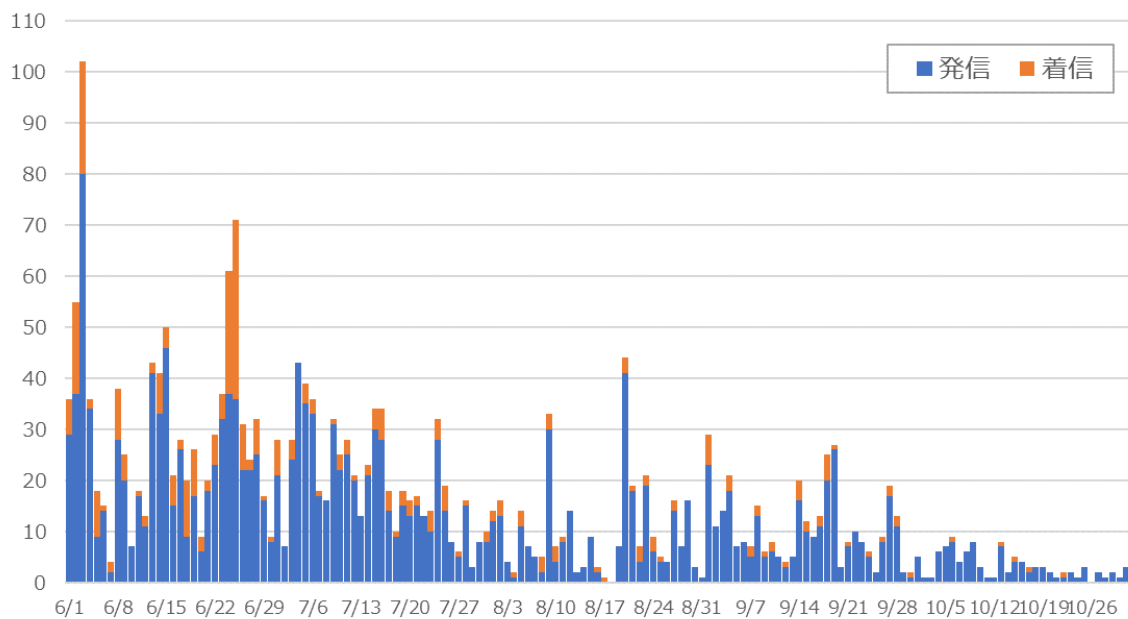


#### (4) 利用総数

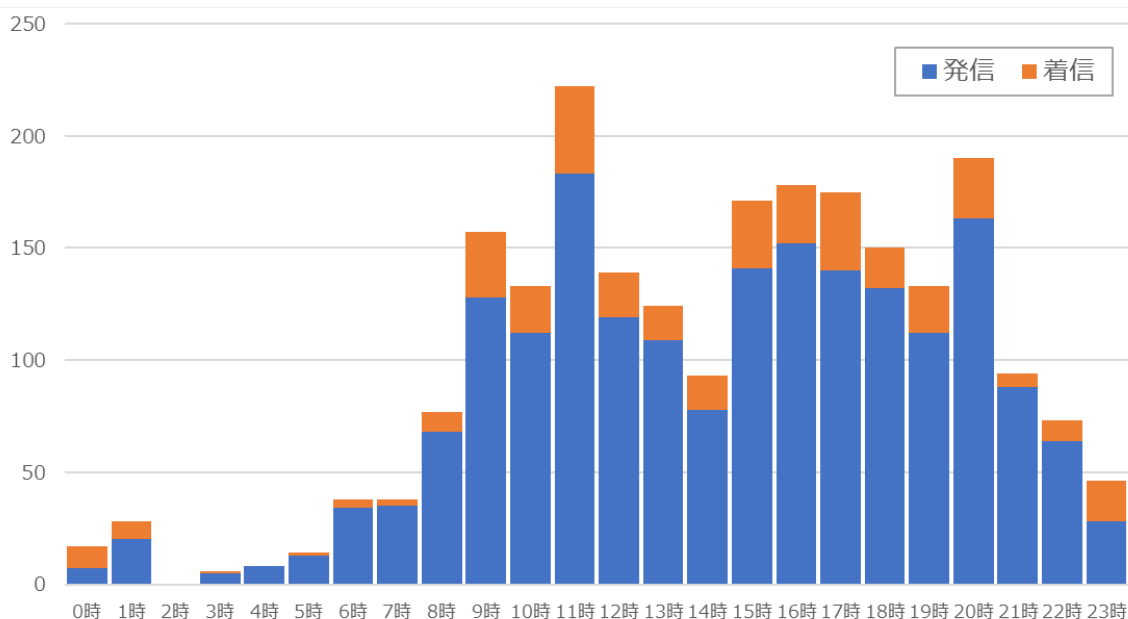
利用者の利用総数をみると、実証実験開始前半の6～7月の利用回数が最も多い傾向がみられた。利用が最も多かった日は6月3日で、合計102回（発信80回、着信22回）の利用があった。

これを時間帯別にみると、幅広い時間帯で利用があったが、利用のピークは11時と20時にあった（それぞれ222回、190回）。今回の実証実験では、オペレーターが10時から18時まで待機しており、それ以外の時間帯は音声認識のみを稼働させる方法をとっていたが、オペレーターが不在の時間帯（18～10時）でも一定程度の利用が確認された。

図表 15 日別 利用総数



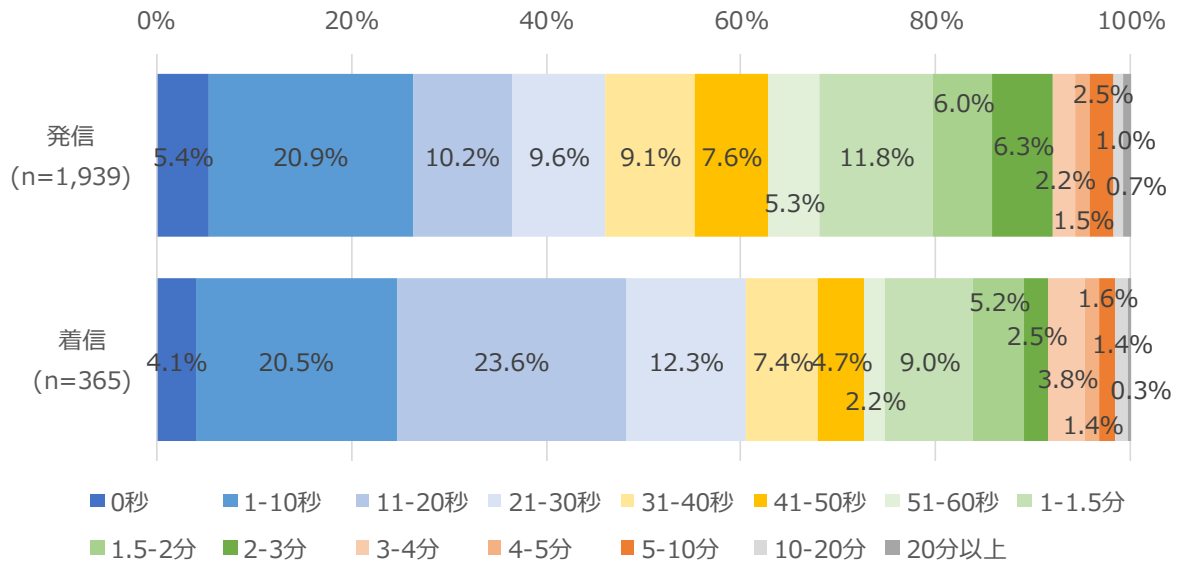
図表 16 時間帯別 利用総数



### (5) 通話時間

発信では 45.4%、着信では 60.3%が 30 秒以下の通話であり、比較的短い通話が約 5 割を占めていた。5 分を超える通話は、発信では 4.2%、着信では 3.3%だった。

図表 17 利用 1 回あたり通話時間



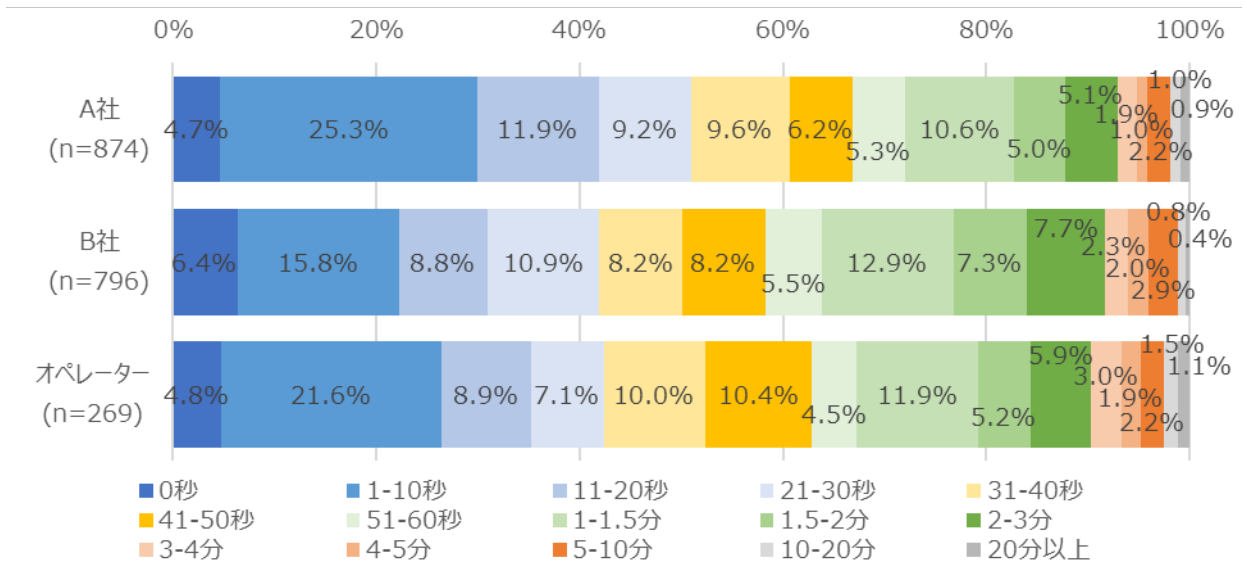
ここで、発信では 5.4%、着信では 4.1%が「0 秒」（つながった瞬間に通話を切った）だったこと、及び 30 秒以下の比較的短い通話がおおむね半数を占めていたことに着目し、どのようなパターンの通話時間が短かったのか・長かったのかを追加的に分析した。

まず、発信に限って、文字化する 3 種類の音声認識（A 社／B 社／オペレーター）別に集計した。その結果、A 社では 30 秒未満の通話が合計で 51.0%とやや高い割合だったのに対し、B 社は 42.0%、オペレーターは 42.4%だった。発信については、5 分を超える通話はいずれの文字化の方法でもほぼ同様だった。

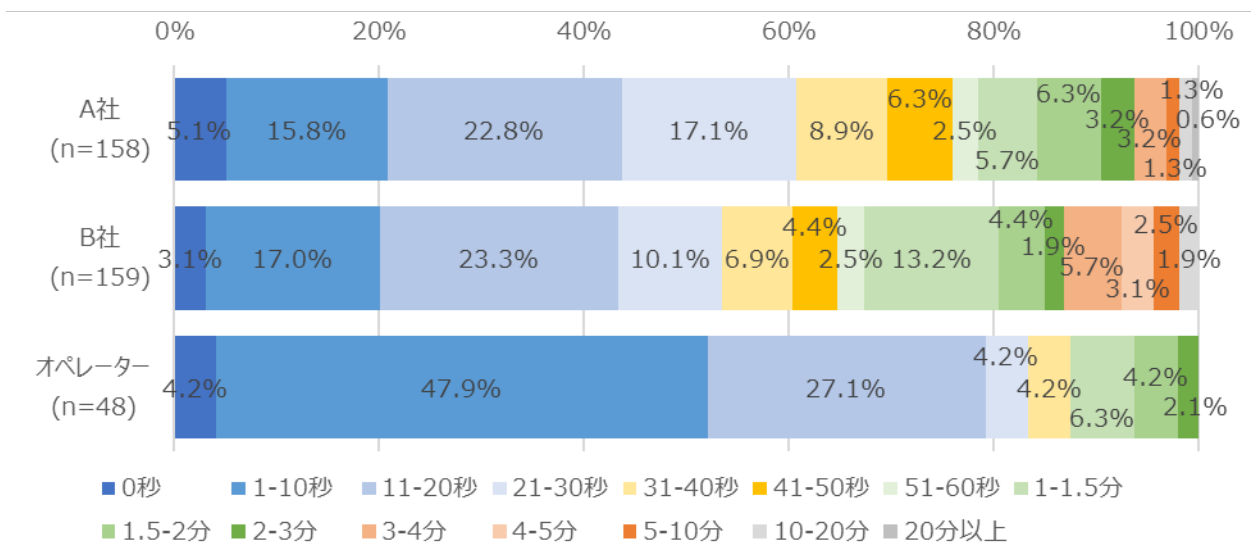
次に、着信に限って、文字化の方法別に集計したところ、A 社と B 社は 30 秒以下の割合がそれぞれ 60.8%、53.5%だったが、オペレーターでは 83.3%と特に高い割合だった。

以上を踏まえた考察として、発信では利用者のタイミングに応じて通話が開始されるため、どの文字化の方式でもほぼ同様の傾向を取ると考えられるものの、A 社だけは通話時間がやや短い傾向がみられており、この背景は以降の実用性に関する評価結果でさらに詳しくみていく。一方、着信は利用者が必ずしも通話の準備をしていない可能性があり、オペレーターにつながる時間帯（10～18 時）は特にその傾向が強かったことが示唆されるが、オペレーターにはつながらず音声認識ソフトのみにつながる夜間帯（18～10 時）は比較的長時間の通話もおこなわれていたとみられる。

図表 18 文字化の方法別 利用 1 回あたり通話時間【発信】



図表 19 文字化の方法別 利用 1 回あたり通話時間【着信】





### 3. 発信時の実用性に関する評価結果

ここでは、利用者に実用性として評価を求めた項目（内容の理解、ニュアンスの理解、テンポのよさ、満足度）について、発信時と着信時では利用者の評価結果に差が生じると想定されることから、利用者が発信した場合に限った分析を行った。

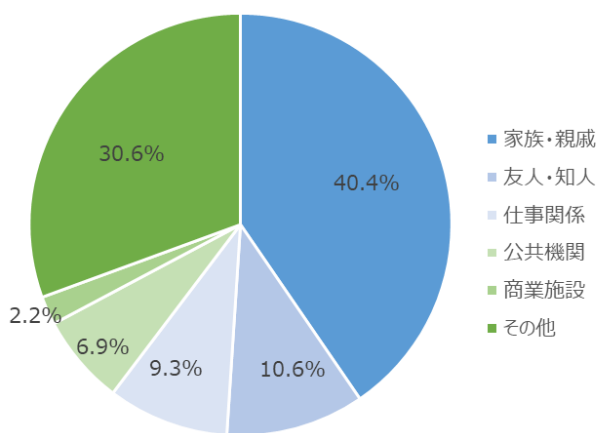
実証期間中の総利用回数 2,943 回のうち、エラーなく接続されたのは 2,304 回で、このうち（受信を除いた）発信回数は 1,939 回、さらにその中で利用直後アンケートに回答があった 1,056 回を分析対象とした。その際、前掲のように通話時間が 10 秒以下のものも発信全体の 26.3%を占めており、通話内容の質的な違いが想定されたため、通話時間が 1 分以上のデータに限った分析もあわせて実施した。

#### (1) 通話シーン

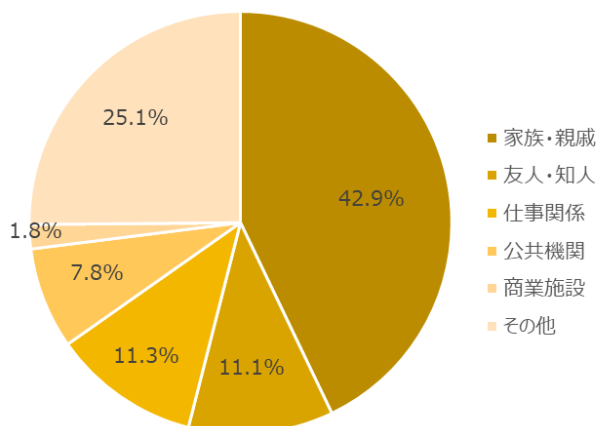
通話の目的は利用直後アンケートで選択肢を示して回答してもらう方法で把握した。回答では、家族・親戚との通話が 40.4%で最も割合が高く、次いで友人・知人との通話が 10.6%、仕事関係の通話が 9.3%と続いた。なお、選択肢に当てはまらない「その他」も 30.6%と高い割合となった。

通話時間を 1 分以上に限っても同様の傾向だが、「その他」の割合がやや減少した。

図表 20 利用者の通話シーン（発信のみ）



図表 21 利用者の通話シーン（発信のみ）【1 分以上】



## (2) 文字付き電話の実用性

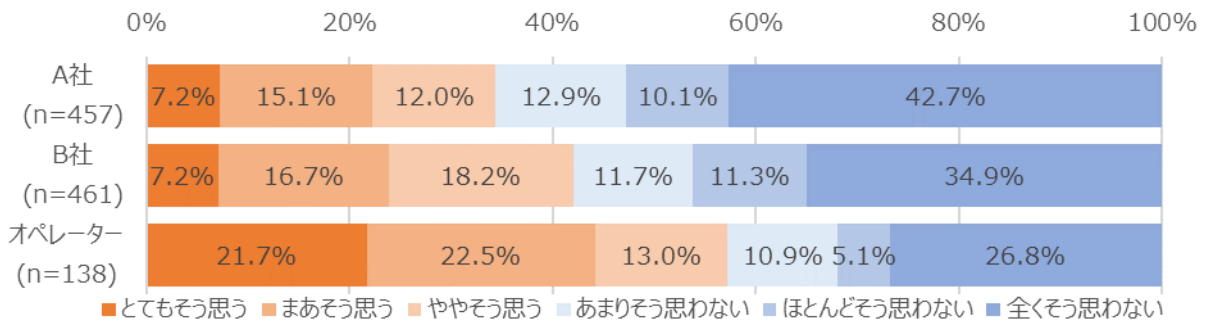
文字付き電話の実用性を、内容の理解、ニュアンスの理解、テンポのよさの3側面で利用者に利用直後アンケートで評価してもらった結果を、文字化する3種類の音声認識別に集計した。

### ① 通話の内容の理解度

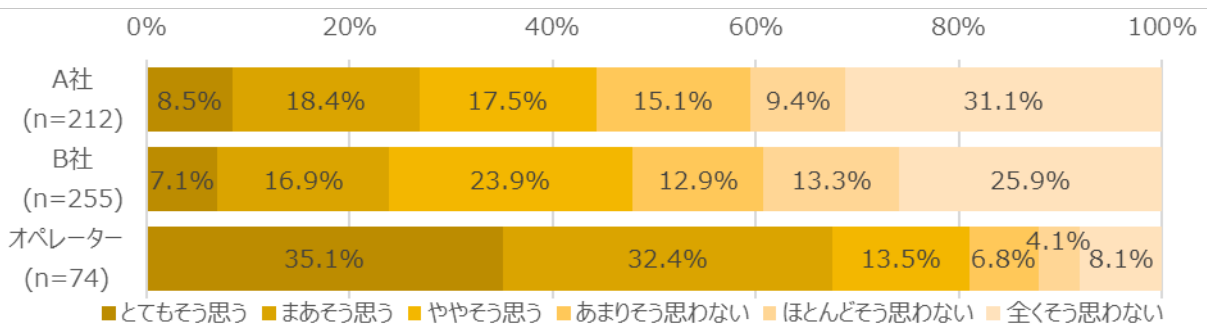
通話の内容の理解については、ポジティブな回答「そう思う」「とてもそう思う」「まあそう思う」「ややそう思う」の合計)との評価は、A社で約3割、B社で約4割、オペレーターで約6割だった。他方、音声認識ソフトは「全くそう思わない」も3~4割にのぼった。

これを、1分以上の通話に限って分析すると、A社・B社とも4~5割と、1分未満の回答も加えた集計結果よりもやや高かった。また、オペレーターでは8割(81.1%)を占め、理解度が高かった。

図表 22 文字化の方法別 通話の内容の理解度



図表 23 文字化の方法別 通話の内容の理解度【1分以上】

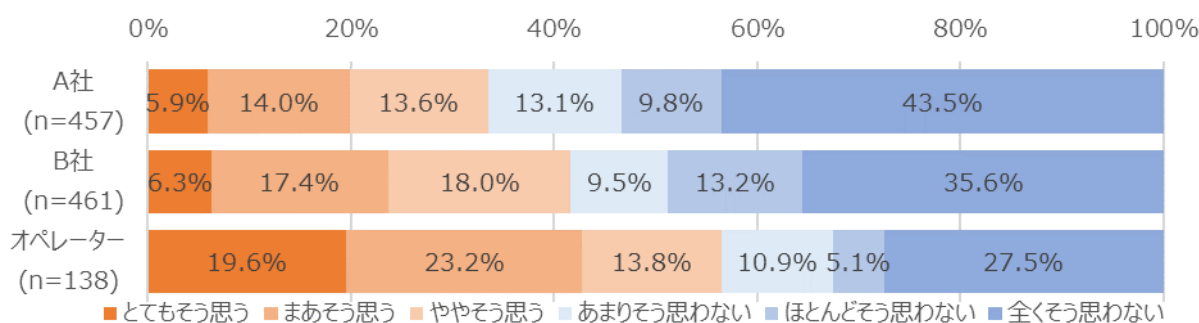


### ② 通話のニュアンスの理解度

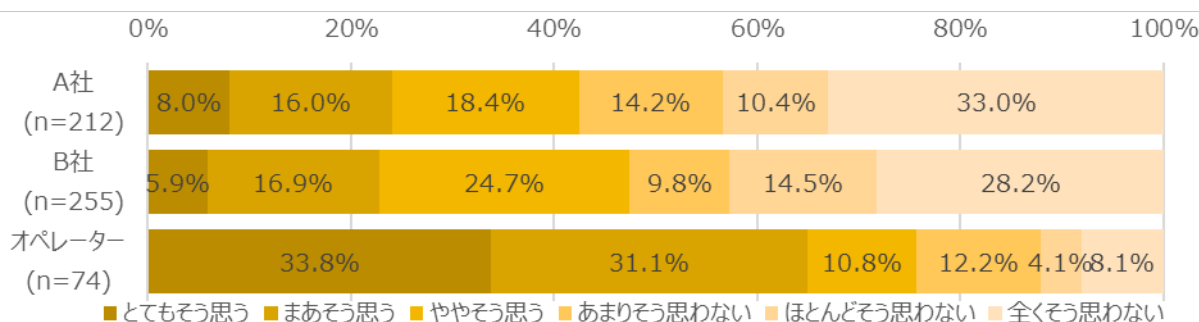
ニュアンスの理解度について、音声認識ソフト(A社・B社)はオペレーターと比較して「そう思う」の割合が1~2割低く、約3~4割であり、「内容の理解」とほぼ同様の傾向がみられた。

1分以上の通話に限った場合、音声認識ソフト(A社・B社)では大きな変化は見られないが、オペレーターはポジティブな「そう思う」との評価が75.7%になった。

図表 24 文字化の方法別 通話のニュアンスの理解度



図表 25 文字化の方法別 通話のニュアンスの理解度【1分以上】

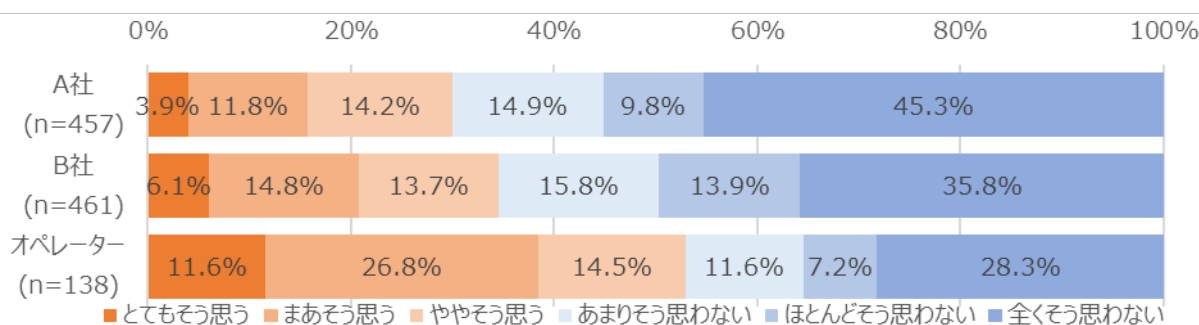


### ③テンポのよさ

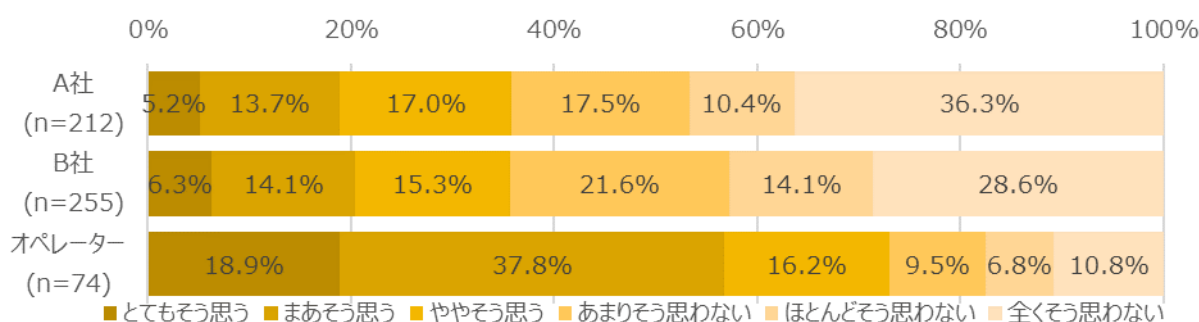
音声認識ソフトはオペレーターより2割程度「そう思う」割合が低く、「内容」や「ニュアンス」の評価と比較すると、全体的にややネガティブな結果だった。

通話時間1分以上に限ると、オペレーターで約7割が「そう思う」と回答していた。

図表 26 文字化の方法別 通話のテンポのよさ



図表 27 文字化の方法別 通話のテンポのよさ【1分以上】

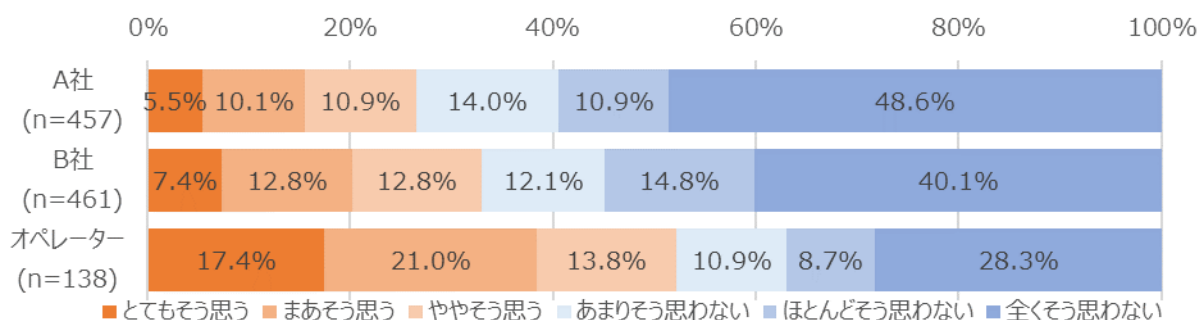


### (3) 通話の満足度

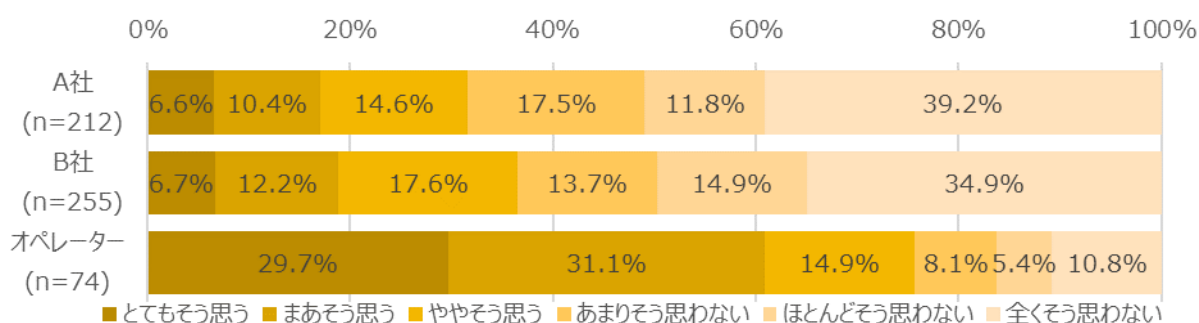
利用者に文字付き電話を利用したことについて満足できるものだったか、利用直後アンケートを通じて評価してもらったところ、音声認識ソフトはA社でもB社でも「思う」が約3割（それぞれ26.5%、33.0%）だったのに対し、オペレーターは「思う」が半数を超えていた（52.2%）。

これを通話時間1分以上に限定して分析すると、音声認識ソフト（A社・B社）はともにやや増加（それぞれ31.6%、36.5%）した。オペレーターは75.7%と大きく増加し、高く評価されていたと言える。

図表 28 文字化の方法別 通話の満足度



図表 29 文字化の方法別 通話の満足度【1分以上】



#### (4) 仮説を検証するための問いに関する分析

本実証の仮説を検証するための問い（リサーチ・クエスチョン）について、多変量解析の手法を用いて分析した。

##### ①目的変数の設定

これまでにみてきたように、実用性に関する 3 種類の評価尺度（内容の理解度、ニュアンスの理解度、テンポのよさ）と満足度は総じて類似した集計結果となっており、総合的な評価結果である満足度を多変量解析における目的変数（原因と結果のうち結果にあたるもの、被説明変数）に位置付けられるか、その妥当性について検討した。

まず、回答の一貫性を示すクロンバック  $\alpha$  の係数<sup>2</sup>は 0.977 と極めて高い数値となった。このことから、実用性に関する 3 種類の評価尺度に満足度を加えた 4 種類の変数は一貫性が非常に高いと考えられることから、この中から満足度を取り出して目的変数として用いても、実用性の評価をする上では問題がないと考えられた。

次に、満足度について基本統計量をみたところ、1（とてもそう思う）～6（全くそう思わない）の 6 段階のうち中央値が 5 で、尖度及び歪度はマイナスの値となっており、正規分布よりも右側に重心があり、比較的広がった分布となっていることが確認された。

図表 30 通話の満足度の基本統計量

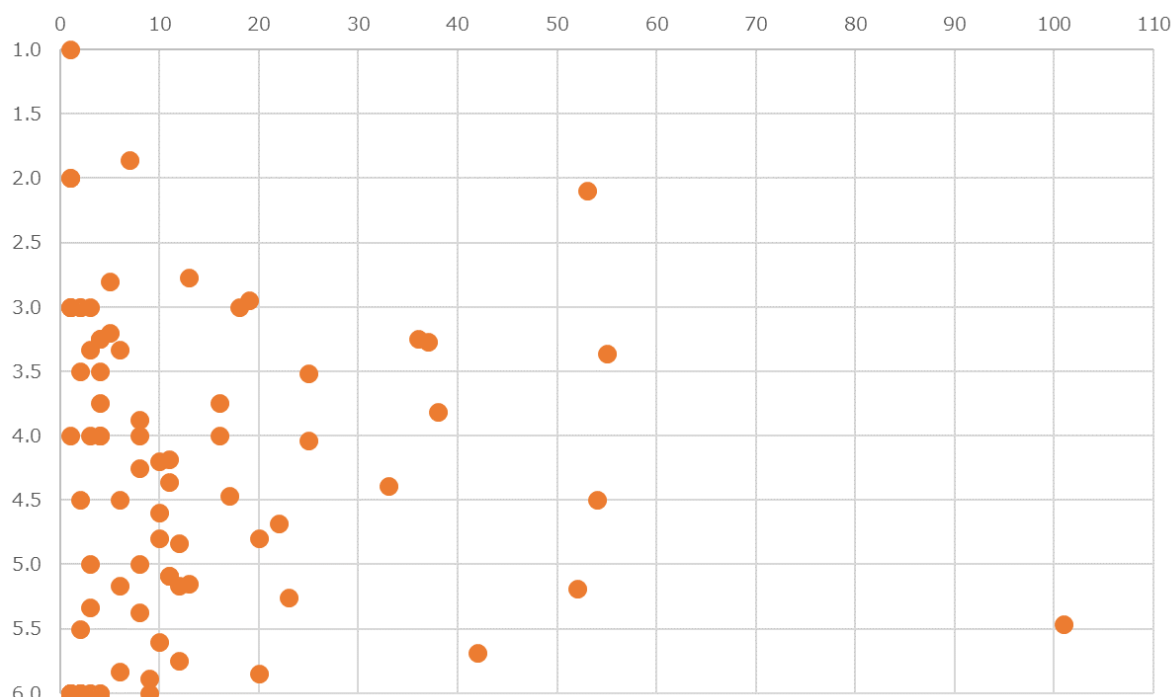
件数	平均値	中央値	標準偏差	尖度	歪度
1,056	4.36	5	1.74	-1.09	-0.58

さらに、1,056 件の回答は実証実験実施期間中の利用者によるアンケート回答のべ回数であるが、特定の回答者に偏った結果であることも想定される。発信で利用した 109 人の利用者のうち利用直後アンケートに 1 回以上回答があったのは 89 人（81.7%）であり、50 回以上の回答者も 5 人いた。

そこで、利用者別に全アンケート回答結果から通話の満足度の平均値を算出し、散布図で示したところ、最もアンケート回答数が多かった回答者（101 回）は平均 5.5 とネガティブな回答傾向がみられたが、総じて、平均満足度はバラついていることがみてとれる。また、サービス利用の習熟性（利用経験を重ねるに従って利用方法に習熟し、満足度も上昇する傾向）も想定されたが、散布図からは、利用総数が多いからといって満足度が上昇するわけではないことも示された。

<sup>2</sup> 本実証のように主観的な心理尺度が用いられた場合に、複数の項目を統計的に分析して、それぞれの回答に一貫性があるかを示すための係数。通常、 $\alpha$  係数が 0.8 以上であった場合は、それぞれの回答に一貫性があるとみなされる。

図表 31 利用者別 平均満足度



※縦軸：全利用における満足度の平均値、横軸：利用総数

以上より、利用直後アンケートの回答回数が利用者ごとにバラつきがあり、特定の利用者の回答傾向に影響を受けていることは否めないものの、回答結果は総じて広く分散しており、目的変数を通話の満足度とすることについては一定の妥当性があるものと考えられる。

## ②重回帰分析

通話の満足度を目的変数とした重回帰分析として、すべての変数（年齢、性別、世帯種別、職業、補聴器等の使用状況、障害の程度、言語明瞭度、電話の習慣、音声認識の種類、文字付き電話利用回数、通話時間、バージョン、通話シーン）をモデルに投入して分析を行った（Stata ver. 15.0 による最小 2 乗回帰分析）。なお、ここではアンケート全問に回答があり、かつ 1 分以上の通話に限って分析対象とした（n=339）。自由度調整済み決定係数は 0.3123 だった。

分析結果は次ページに示す図表 32 の通りである。ここでは＜女性、聴者を含む世帯、無職、A 社の音声認識ソフト、通話先「その他」＞のケースを基本とした比較結果である。p 値が統計的に有意水準にあるものに色を付けており、係数がプラスのものは満足度が下がる要因、係数がマイナスのものは満足度を高める要因として考えられる。

分析モデルに投入した変数のうち、オペレーターによる音声認識であること、利用回数が多いこと、通話時間が長いこと、通話先が公共機関であることは、満足度が高い要因として考えられた。また、男性であること、難聴者のみの世帯であること、通話先が友人であることは 5% 有意水準で係数の符号がマイナスであり、満足度が高かった。

主婦・主夫も 5%有意水準だが符号はプラスであり、満足度が低いという結果だった。

なお、A 社の音声認識ソフトを基本として B 社と比較した結果として、満足度について統計的に有意な差は認められなかった。

図表 32 回帰分析結果

変数名	係数	標準誤差	t 値	p 値	95%信頼区間
年代	0.0469	0.1154	0.41	0.685	-0.1801 ~ 0.2739
性別（「女性」と比較して）					
男性	-0.6541	0.2840	-2.30	0.022	-1.2129 ~ 0.0953
世帯（「聴者を含む世帯」と比較して）					
難聴者のみ	-0.7661	0.3481	-2.20	0.028	-1.4510 ~ 0.0812
独居	-0.4284	0.3468	-1.24	0.218	-1.1106 ~ 0.2539
職業（「無職」と比較して）					
働いている	0.6214	0.3447	1.80	0.072	-0.0569 ~ 1.2997
主婦・主夫	0.9255	0.3738	2.48	0.014	0.1900 ~ 1.6609
学生	0.8041	1.2327	0.65	0.515	-1.6213 ~ 3.2294
補聴器等使用あり	-0.6366	0.5226	-1.22	0.224	-1.6647 ~ 0.3915
聴覚障害程度区分	0.0991	0.0820	1.21	0.228	-0.0622 ~ 0.2605
言語明瞭度	0.0533	0.1033	0.52	0.606	-0.1499 ~ 0.2565
電話の習慣	-0.2251	0.1409	-1.60	0.111	-0.5025 ~ 0.0522
音声認識（「A 社」と比較して）					
B 社	-0.0946	0.1765	-0.54	0.592	-0.4419 ~ 0.2526
オペレーター	-1.8914	0.2863	-6.61	0.000	-2.4547 ~ 1.3280
利用回数	-0.0288	0.0082	-3.52	0.000	-0.0449 ~ 0.0127
通話時間（秒）	-0.0019	0.0004	-4.39	0.000	-0.0028 ~ -0.0011
バージョン	-0.0845	0.2443	-0.35	0.730	-0.5652 ~ 0.3962
通話先（「その他」と比較して）					
家族	0.0295	0.2343	0.13	0.900	-0.4314 ~ 0.4905
友人	-0.6858	0.3380	-2.03	0.043	-1.3507 ~ 0.0208
仕事	-0.1072	0.3004	-0.36	0.721	-0.6982 ~ 0.4838
公共機関	-1.6831	0.3690	-4.56	0.000	-2.4090 ~ 0.9572
店舗	-0.0467	0.9163	-0.05	0.959	-1.8495 ~ 1.7561

## 4. 考察

### (1) 高い利用ニーズが確認された

登録した 280 人のうち 109 人が合計 2,304 回の通話を行っており、5 ヶ月間で 100 回を超える利用者もいた。利用者の属性や利用シーンは多様であり、月別の利用実績や極端に短い通話時間が一定数あったことなどから、「まずは一度、試しに使ってみよう」という利用も多かったと推察されるものの、実証期間にわたって継続的に利用した登録者もいたことが明らかになった。

また、時間帯別の分析からは、オペレーターが不在の時間帯でも一定の利用実績がみられた。全体的な利用満足度は高くないながらも、難聴者の電話使用に対する潜在的なニーズが裏付けられた結果だと言えるのではないかと考えられる。

### (2) 音声認識ソフトよりオペレーターへの評価が高かった

主たる問いとした「音声認識ソフトの実用性の実感値」については、実用性の各項目（意味、ニュアンス、テンポ）すべてにおいて、オペレーターによる入力の前位性が確認された。特にテンポについては、改善の余地が残されていると示唆された。

全通話をみると、利用満足度としてはポジティブな回答をネガティブな回答が上回っており、社会サービスとしては大いに改善の余地があると思われる。しかしながら、通話を 1 分以上のものに区切って分析をしたところ、この傾向は逆転していた。中でもオペレーターによる文字化に対する満足度は総じて高かったことは着目に値する結果であった。

### (3) 利用満足度を高める取組みが求められる

既述のように、全体としては利用者から厳しい評価が示されているが、これは期待の裏返しという意味合いも含まれていると考えられる。利用シーンについてはより詳細な分析が求められるが、事実関係を明確にするやり取りが必要な行政機関とのやり取りで満足度が高かったり、一方で機微に触れるニュアンスを伝えることの多い友人とのやり取りも比較的満足度が高かったりするなど、様々な利用の場面で、様々なニーズに対応していたことがうかがえる。

今後、難聴者の QOL の改善に向けて、本サービスがより高精度で、平易に利用できるようになることが望まれる。




## 第5章 実証実験後の活動と今後の方向性

### 1. 全難聴福祉大会での中間報告と実証実験後のアンケート

#### (1) 第3分科会のセッション概要

2019年11月23～25日に滋賀県大津市ピアザ淡海にて、「全国中途失聴者・難聴者福祉大会 in 滋賀」が開催された。23日の第3分科会（情報文化部）において参加者93名のもと「電話で音声認識は使えるの？～文字付き電話の可能性と課題～」と題し、当事業の中間報告をおこなった。

第3分科会	(情報文化部)
『電話で音声認識は使えるの？ ～文字付き電話の可能性と課題～』	
1 電話リレーサービス状況 2 「文字付き電話」の実証実験概要 3 「文字付き電話実証実験」の評価と分析 4 パネルディスカッション・フロアとの質疑応答 5 分科会まとめと今後の方針	
パネリスト 星川 安之氏（共用品推進機構） 一瀬 宗也氏（シーコミュ） 家子 直幸氏（三菱UFJリサーチ&コンサルティング株式会社） 小川 光彦氏（全難聴 情報文化部） 座長 新谷 友良氏（全難聴 理事長）	

前半は当事業の背景や文字付き電話の説明と、実証実験での評価・分析の中間報告がおこなわれた。後半はパネリストによるディスカッションや会場の参加者との質問、意見交換が活発にされた（画像参照）。

分科会では聴覚に障害がある参加者にも内容が伝わるように、PCによる要約筆記、手話通訳、聞こえを補助するヒアリンググループの情報保障をおこなった。



## (2) 参加者との質疑応答

会場参加者からは以下のような意見がだされた。

- ・ 1200 万人の人が電話の利便性を利用できていない。電話が使えず、不自由な生活をしている
- ・ 電話は使い慣れないと、不安がある。誤変換がたくさん出ると不安が高まり気軽に電話ができない。もっと使いやすい形にする必要がある
- ・ この会場でも電話リレーサービスを使ったことがある人は少ない。電話リレーサービスの活動をしているが関心が薄いと思う。難聴者団体も電話リレーサービスの啓発をもっとしてほしい
- ・ 今回はスマートフォンだが、スマートフォンやタブレットはデータ通信で使うだけだ。自宅でパソコンを使って利用したい人は増えてくるだろう
- ・ 電話直後のアンケートは選択で選ぶだけで細かい実感を伝えられない

分科会終了後、参加者に対しアンケートを実施したところ、参加者 93 名のうち 45 名から回答があった。聴覚障害の当事者であるという回答が 33 名 73%。

分科会の内容の理解度について、

よく分かった…9 人、まあまあ分かった…22 人、ふつう…8 人、少し分かった…4 人、無回答…2 人と、「まあまあ分かった」以上の方が 31 名 69%という結果となった。

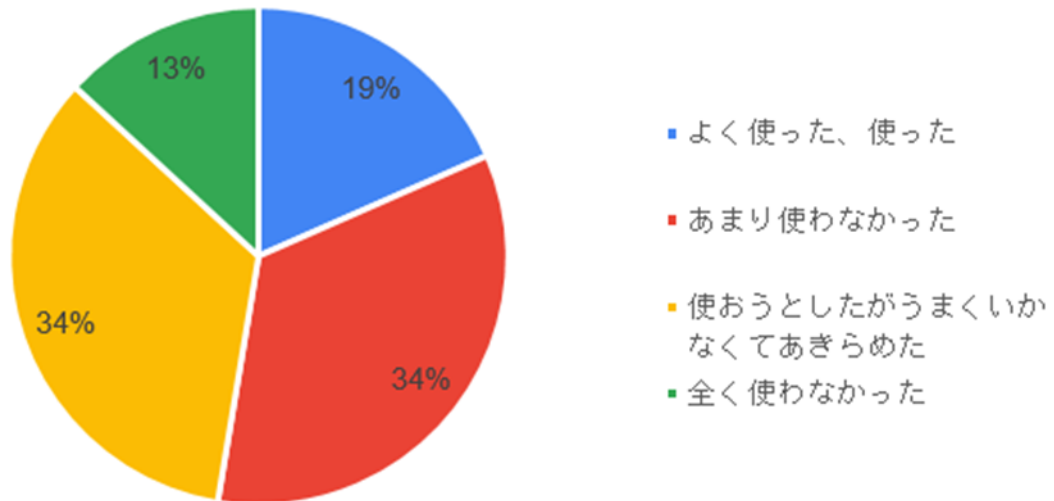
## (3) 実証実験終了後アンケート

実証実験で実施したアンケートは、利用者本人が電話終了のつど、通話内容の理解度等を評価する方法をとり、使わなかった人等を除外したデータを集約して第 4 章の分析・評価をおこなった。そのため使い勝手の評価やフリーコメントなどの項目はなく、上記分科会でも要望があったため、モニター申込者全員に実証実験終了後アンケート調査をおこない、76 件の回答を得た。質問は文字表示に関連するもの以外にも操作性やアプリケーション導入のしやすさ、今後の利用意向など多岐にわたる。

①Q1 と Q2 では、「文字付き電話を実際に使ったか、使わなかったか」と、その理由を質問している。今回の実証実験では半数近くが実際の電話会話に至らなかった。理由としては使い方がわからなかった、やってみたがうまく変換できなかった、などの意見が多数をしめた。原因として、第 3 章で述べたように外付けアダプターが必要になったことで操作がむずかしくなったことや、アダプター経由によるトラブルが多かったなどが考えられる。その他、今回アプリケーションを Google Play に登録できなかったため、インストールでつまづいた人もいると思われる。

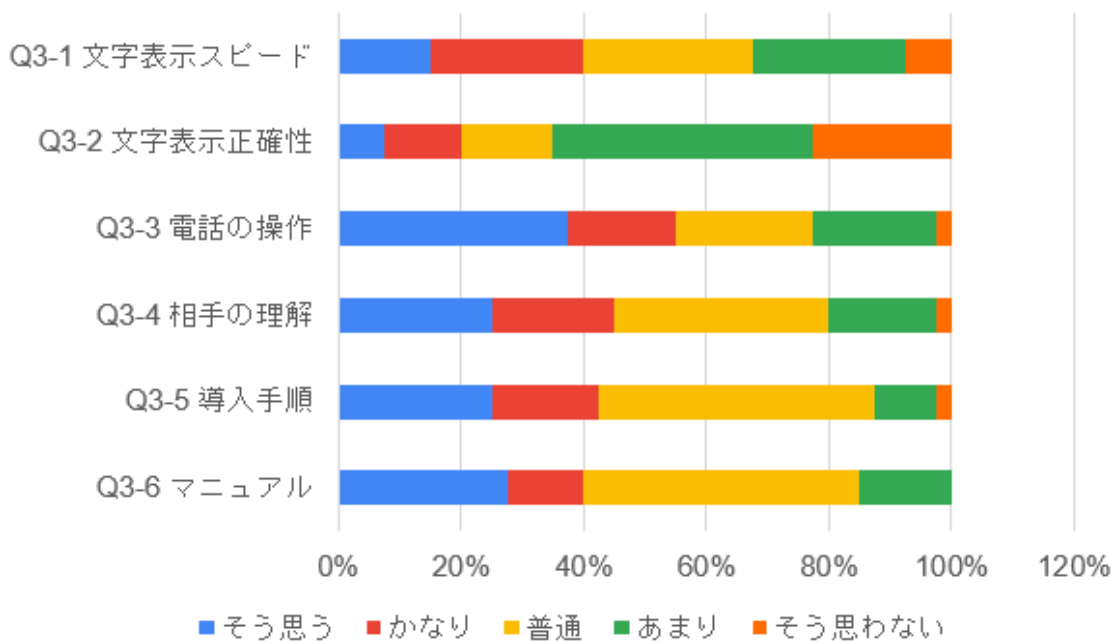
図表 1 文字付き電話の利用の有無

Q1 今回の実証実験で、文字付き電話を実際にお使いになりましたか？



② Q3 では、「よく使った、使った」、「あまり使わなかった」人を対象に文字付き電話全般の評価をたずねた。電話の操作性について、比較的肯定的評価が多かった。

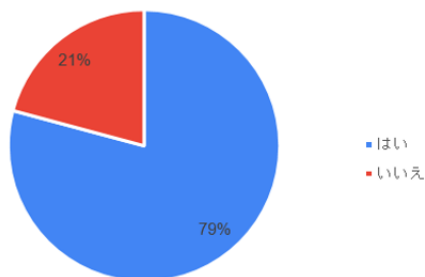
図表 2 文字付き電話の評価



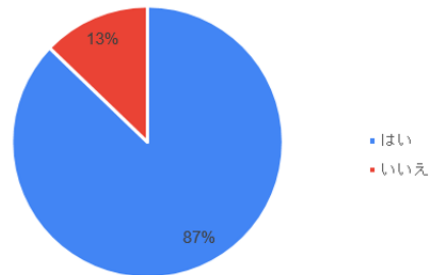
③Q4 では今後の利用意向などについて質問している。可能であれば今後も文字付き電話を利用したいがほぼ 80%、その時有料でもよいと答えたのはそのうち約 50%であった。また料金について約 80%は 990 円以下がよいとしている。さらに電話リレーサービスが国の制度として提供されるなら、文字付き電話も提供してほしいと答えた割合は 90%近くになる。

**図表 3 文字付き電話 今後の利用意向他**  
(左：全員、右：利用者のみ)

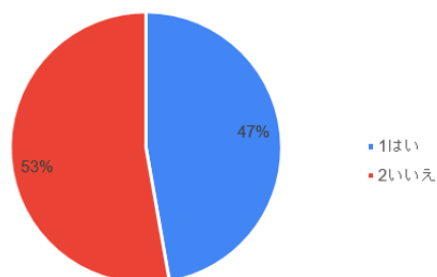
Q4-1 可能であれば今後も文字付き電話を利用したい



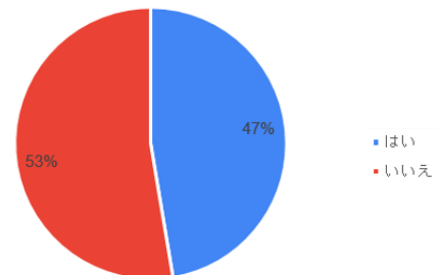
Q4-1 可能であれば今後も文字付き電話を利用したい



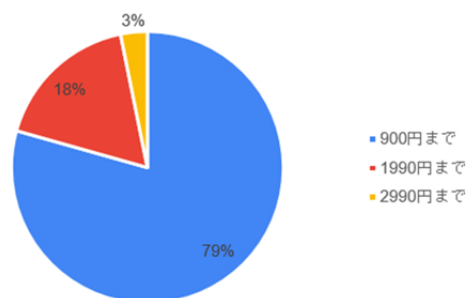
Q4-2 その場合、有料であっても文字付き電話を使いたい



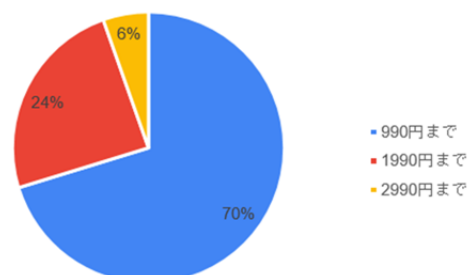
Q4-2 その場合、有料であっても文字付き電話を使いたい



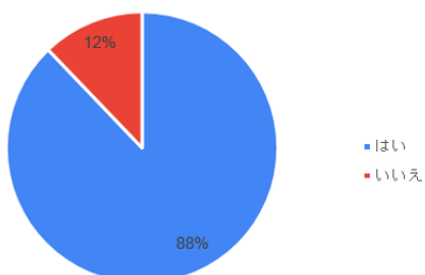
Q4-3 有料の場合、月額いくらまでなら払っても使いたいですか



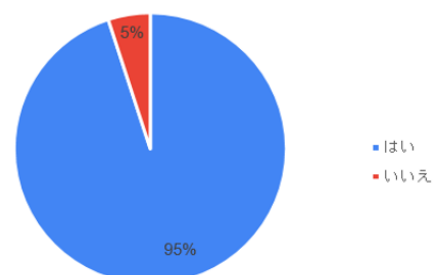
Q4-3 有料の場合、月額いくらまでなら払っても使いたいですか



Q4-4 電話リレーが国の制度として再来年4月から提供されますが、その時に文字付き電話も提供してほしい



Q4-4 電話リレーが国の制度として再来年4月から提供されますが、その時に文字付き電話も提供してほしい



## 2. 電話リレーサービス制度化の動きと文字付き電話

当事業を実施中の2018年11月7日に国会で安倍総理大臣が「電話リレーサービスは大切な公共インフラである」と答弁し、電話リレーサービスの公的制度化を認めるとともに、主管省庁を総務省とすることを明らかにした。

その結果2019年1月より制度化に向けて総務省と厚生労働省による、「安定的・継続的な提供、適正性かつ効率性、実現可能性などに配慮した電話リレーサービス」のあり方の検討が進み、2019年11月にそれまでの当事者を含む電話リレーサービスワーキンググループでの討議や、パブリックコメント、実務検討会を踏まえ報告書が発表されている。

これまでの発表では、文字付き電話は公的サービス開始時期には間にあわないとされているが、サービス開始後早い時期に実現に向けた検討がはじまることも考えられる。また音声認識については、「電話リレーサービスの実現を前提として、音声認識などの技術開発は『車の両輪』として並行して進める」とあり、積極的な取り組みが期待される。

## 3. 総括

全難聴は日本財団の助成を受けて、2年間にわたって「電話利用における音声認識ソフトの調査」事業を実施した。

この事業実施に当たっての問題意識は、「電話音声文字化する場合、オペレーターが電話音声を聞いて入力することが考えられるが、オペレーターのかわりに音声認識ソフトを活用し文字化する方法も考えられる。しかし、現時点での音声認識ソフトは通信環境や話者の発音などによって認識精度が左右されやすいと考えられるため、電話音声の文字化における音声認識ソフトの活用がどれだけ実用的なのかを分析・評価する」ことになった。

実証実験の結果は、各章に記述したとおりであるが、結論的に言うと「電話音声の文字化について高い利用ニーズが確認されたものの、音声認識ソフトの実用性については、各項目（意味、ニュアンス、テンポ）の評価は低く、オペレーターによる入力の優位性が確認された」というものである。

実証実験のこの結果は、日本財団の電話リレーサービスモデル事業の実績を前提に、事前に想定し得る結果であったが、音声認識を利用するにあたっての環境的制約からくる文字化アプリの未完成のもたらした部分も大きかった。第3章「開発アプリケーションと実証実験の概要」に説明されているように、実証実験用文字化アプリ開発にあたっては、iOSに止まらずAndroidにおいてもAPI(Application Program Interface)利用が不可能となった。その結果、やむなく「外付けアダプター」による電話音声利用をおこなったが、この方法は利用面の不便さ・煩雑さに加え、音質の劣化による音声認

識の精度にも大きな影響をもたらした。今回の実証実験が「調査環境を整えた音声認識率の数値比較（ラボ型）」ではなく、社会的な実証実験としてモニターが普段の生活のなかで文字付き電話を使い、利用者としての実用性を評価する方法（フィールド型）」で進められたことを考えると、文字化アプリの操作性が実験結果に与えた影響は非常に大きかったと考えられる。

電話リレーサービスにおける仲介者（手話・文字オペレーター）利用の有効性は確認されており、現在その事業モデルの課題は、事業運営上の課題（事業主体、オペレーター養成・配置・事業費用の負担方法など）に移りつつある。総務省に設けられた「電話リレーサービスに係るワーキングチーム」報告では、サービス提供の費用を中間値で34-43億円と想定、その費用の多くは仲介者に要する費用としている。また、同報告では利用者の範囲を限定しないとしているが、利用者は8万人/年（中間値）と予測されている。第2章で記述されている通り、我が国の聞こえに困難を抱えている人は、1,400万人以上いると想定されるが、この人たちの多くは電話利用が困難な状況に置かれている。この人たちのどの程度が電話リレーサービスを利用するか見極めは困難であるが、仮にこのうちの10%の人が電話リレーサービスを利用するとすれば、サービス利用者は約140万人となり、サービス提供費用の大幅な増加も想定される。

今回の実証実験で、音声認識利用の文字付き電話リレーに対する潜在的ニーズの確認は出来たが、音声認識利用の環境面からの制約で、フィールド型の実証実験は不十分に終わった。実証実験で浮かび上がった、

1. 文字化アプリの完成度を向上するための、携帯端末OSと外部システムとつなぐ技術仕様「API（アプリケーション・プログラム・インターフェース）」の公開の実現
2. 電話通信環境の音声認識精度に与える影響のさらなる評価
3. 騒音など送受信者の利用環境が音声認識精度に与える影響の評価
4. 文字付き電話の送受信両面での利便性の評価
5. 文字付き電話利用者が文字表示に求める文字表示精度の研究
6. 電話利用における福祉サービスと公共サービスの在り方の検討

といった課題の着実な検討が継続され、文字付き電話の実用化が近い将来実現されることを期待して、本報告書の総括としたい。

最後に、2年間にわたって今回の実証実験に多大なご支援をいただいた日本財団をはじめとする関係者各位に心からお礼を申し上げたい。

## (資料) 電話利用における音声認識ソフトの調査 会議開催状況

### 【事業委員会】

2018 年度

第 1 回	4 月 27 日	事業目的、委員の役割を説明。 電話の音声文字化の 2 つの方法（音声認識および手入力）について。音声認識の検証評価方法を検討。今後の進行について確認。
第 2 回	7 月 27 日	評価方法、評価対象を検討。 調査スケジュール（2018 年度は実験準備およびパイロットテスト。2019 年度に実証実験、データ分析、報告書作成）を確認。
第 3 回	2019 年 3 月 6 日	検証が必要な内容の詳細について意見交換。パイロットテスト実施について説明。2019 年度に実証実験をおこなうことを確認。

2019 年度

第 1 回	5 月 9 日	実証実験の日程（6 月 1 日～9 月 30 日）、全難聴福祉大会での中間報告、報告書完成までのスケジュールを確認。文字付き電話の紹介とデモ。モニター申し込み状況の確認及び募集方法の検討。
第 2 回	2020 年 1 月 23 日	報告書案について質疑応答。完成までのスケジュールを確認。

### 【事業進捗会議】

2018 年度 8 月 16 日・10 月 12 日・12 月 13 日・2019 年 2 月 15 日

2019 年度 4 月 22 日・6 月 26 日・8 月 1 日・9 月 4 日・11 月 14 日・  
2020 年 1 月 16 日

## 電話利用における音声認識ソフトの調査 事業委員会委員名簿

(順不同、肩書は就任時のもの)

星川 安之： 公益財団法人 共用品推進機構 専務理事（委員長）

大山 潤爾： 国立研究開発法人 産業技術総合研究所 人間情報研究部門  
感覚知覚情報デザイン研究グループ 主任研究員

井上 正之： 国立大学法人 筑波技術大学 産業技術学部産業情報学科准教授

新谷 友良： 一般社団法人 全日本難聴者・中途失聴者団体連合会 理事長

川井 節夫： 一般社団法人 全日本難聴者・中途失聴者団体連合会  
情報文化部副部長

(2018、2019年度日本財団助成事業)

「電話利用における音声認識ソフトの調査」

～文字付き電話の可能性と課題～ 報告書

発行日：2020年3月15日

編集・発行：一般社団法人 全日本難聴者・中途失聴者団体連合会  
〒162-0066 東京都新宿区市谷台町14-5 MSビル市ヶ谷台1階

TEL 03-3225-5600 FAX 03-3354-0046

<https://zennancho.or.jp>