

# 障害者自立支援機器等開発促進事業

言語障害者向けに人間味のある声で会話補助する支援機器の  
低価格化に向けた開発  
(3年計画の1年目)

平成22年度 報告書

開発代表機関 株式会社ウォンツ

平成23(2011)年4月

# 言語障害者向けに人間味のある声で会話補助する支援機器の 低価格化に向けた開発 概要

株式会社ウォンツ

## 全体の概要

声を失う人やその恐れのある方に対して供されている、本人の肉声感を保った会話補助支援装置 Polluxstar（以下 PXT）がある。本人の肉声感を保つため、声を失う前の本人の声から作られた音声ライブラリにより実現している。その会話補助支援装置が抱えている2つの課題を克服するための試作開発を行い、その効果をモニター評価等により確認した。

まず、既に発声障害がある方に対しては、性別、年齢や声の高さなどが異なる様々な代用音声ライブラリを用意し、このライブラリから好みの声質を選択することとした。次に録音作業の負担軽減対策として、ご自身のペースで自分の声を保存し、音声ライブラリ化を支援するユーザ作業用録音ツールや録音作業ガイドラインの整備と、自動音声収録ライブラリ化ソフトウェアの試作導入を進めた。

## 1. 代用音声ライブラリ

声の年齢や声の高さなどが異なる様々な声を代用音声ライブラリとして収録し、これら代用音声ライブラリから好みの声質を選択する会話補助支援装置を実現ための試作開発である。今年度は作業時間の制限などもあり、代用音声ライブラリのモニター評価をするための最低限の試作となった。

Webアンケート形式で予備的な評価を実施した。評価方法は、被験者に代用音声ライブラリを使った合成音を数秒程度試聴もらい、それぞれの声が自分の声質としてのふさわしさを5段階で主観評価をおこなった。

表1：代用音声ライブラリ

No	性別	年代	記事
N1-1	男性	50代後半	
N1-2	男性	60台半ば	
N1-3	女性	60台前半	
N1-4	男性	60台前半	
N1-5	女性	50台半ば	
N1-6	男性	50代前半	
N1-7	女性	50代前半	
N1-8	男性	40代後半	
N1-9	男性	40代後半	関西訛り



写真1：会話補助支援装置 Polluxstar

## 2. ユーザ録音

ご自身のペースで自分の声を保存し代用音声ライブラリ化を支援するユーザ作業用の標準的な録音ツールを用意するとともに、具体的な録音作業の進め方と注意点などを解説したガイドラインを作成した。

発話機能の低下はないが、発話時間が一日の内の短時間しか確保できない場合について適用おこなった。家族の協力の下、自宅録音でどの程度の音声作業ができるかを検証した。本ケースでは本人と家族の生活に支障ないレベルでの録音作業は週に30分程度であり、数週間にわたる録音期間が必要であったが録音作業は完了した。



写真2：ユーザ録音の様子

## 3. 自動音声収録ライブラリ化ソフトウェア

音声収録から音声ライブラリ化作業に掛かる作業コストの削減を目指して、自動的に音声録音すると同時に音声認識を利用した自動録音、自動加工ツールを導入した。

- ・ 画面のガイドに従って発声する。
- ・ いつでも中断、再開可能。
- ・ 発声終了後、自動加工を開始し音声ライブラリ化をおこなう。

ツールの使い勝手について開発プロジェクト内部で実験をおこなった。音声収録時にいくつかの注意点に気を配る必要はあるが、機能に問題はない。



写真3：自動音声収録  
ライブラリ化ソフトウェア

# 目 次

．報告	1
言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発 株式会社ウォンツ	
．開発成果の公表に関する一覧表	10
．開発成果の公表に関する刊行物・別刷	10
．参考資料	11
1．代用音声ライブラリ等の話者選定ガイドライン	11
2．話者評価方法ガイドライン	15
3．収録原稿に関するガイドライン	20
4．代用音声ライブラリ開発に関する実証試験研究計画書	23
5．代用音声ライブラリ開発に関する実証試験研究計画書	31
6．モニター評価参加者への説明文書	38
7．モニター評価参加者の同意書	44
8．会話補助支援装置に関するアンケート調査（モニター評価）参加者への説明文書	46
9．代用音声に関する Web アンケート	49
10．支援機器、音声収録ツール 使用の様子	52

# 障害者自立支援機器等開発促進事業 報告書

## 言語障害者向けに人間味のある声で会話補助する支援機器の 低価格化に向けた開発

開発代表機関：株式会社ウォンツ

開発分担機関：合同会社テクノマインズ

### 開発要旨

声を失う人やその恐れのある方に対して提供している、本人の肉声感を保った会話補助支援装置 Polluxstar（以下 PXT）がある。本人の肉声感を保つため、声を失う前の本人の声から作られた音声ライブラリにより実現している。その会話補助支援装置が抱えている2つの課題を克服するための試作開発を行い、その効果をモニター評価等により確認した。

まず、既に発声障害がある方に対しては、性別、年齢や声の高さなどが異なる様々な代用音声ライブラリを用意し、このライブラリから好みの声質を選択することとした。次に録音作業の負担軽減対策として、ご自身のペースで自分の声を保存し、音声ライブラリ化を支援するユーザ作業用録音ツールや録音作業ガイドラインの整備と、自動音声収録ライブラリ化ソフトウェアの試作導入を進めた。

株式会社ウォンツ

取締役ソリューション事業部長

川島 紳

ソリューション事業部グループリーダー

本田 武司

合同会社テクノマインズ

代表社員社長

三木 敬

音声収録作業の身体的負荷(通例1日～3日)とその後の音声加工作業費用(数十万円)により、その普及が進まないことである。

この2つの課題を克服すべく、

1つ目の課題に対しては、声の年齢や声の高さなどが異なる様々な声を代用音声ライブラリとして収録し、これら代用音声ライブラリから好みの声質を選択する会話補助支援装置を試作する。

2つ目の課題に対しては、ご自身のペースで自分の声を保存し代用音声ライブラリ化を支援するユーザ作業用の録音ツールの整備や、さらなるコスト削減を目指して、自動的に音声録音すると同時に音声認識を利用した自動録音、自動加工ツール（PCソフトウェア）の導入を目指すものである。健常者や軽微な発話障害者において、これら代用音声ライブラリ支援ツールの有効性を検証する。

### ．報告

#### A．開発目的

声を失う人やその恐れのある方に対して、声を失う前に本人の声を収録することで、本人の肉声感を保った会話補助支援装置 Polluxstar（以下 PXT）を提供しているが、現状2つの課題がある。

1つ目は既に発声障害がある方への適用ができないことであり、2つ目は長時間にわたる

「代用音声」の利用においては、多くの音声ライブラリから好みの声質を選びお使いいただくことで、

- 「人間味のある声」が安価に利用できる。

「自分の声」の利用においては、この取り組みの結果生まれる製品は、従来製品と比べ、製品利用の面では、ほぼ同等の性能で 사용할ことができるが、個々の完成品を作り出すまでの製造過程が改善され、

- 長時間にわたる音声収録作業による身体的負荷が軽減される。
- 音声加工作業の削減によりコストが削減される。

## B．開発する支援機器の想定ユーザー

自分の気持ちや意思などをうまく自分の声で伝えられない発話障害をお持ちの方、あるいは将来の発話障害に備えたい方。

## C．開発体制

### 1．開発者

開発代表者：川島 紳

株式会社ウォンツ ソリューション事業部

開発分担者：三木 敬

合同会社テクノマインズ

### 2．協力者

久保 健彦（アドバイザー）

麻生リハビリテーション専門学校

言語聴覚学科 学科長

川田 明広

都立神経病院 神経内科 医長

長尾 雅裕

都立神経病院 神経内科 医長

中山 優季

東京都神経科学総合研究所

難病ケア看護研究部門

## D．試作した機器またはシステム

### (1)代用音声ライブラリ(男性6名、女性3名)

声の年齢や声の高さなどが異なる様々な声を代用音声ライブラリとして収録し、これら代用音声ライブラリから好みの声質を選択する会話補助支援装置を実現ための試作開発である。今年度は作業時間の制限などもあり、代用音声ライブラリのモニター評価をするための最低限の試作となった。

表1 代用音声ライブラリ

No	性別	年代	記事
N1-1	男性	50代後半	
N1-2	男性	60台半ば	
N1-3	女性	60台前半	
N1-4	男性	60台前半	
N1-5	女性	50台半ば	
N1-6	男性	50代前半	
N1-7	女性	50代前半	
N1-8	男性	40代後半	
N1-9	男性	40代後半	関西訛り

- 発話者の年齢、性別は発声障害の方の布(仮説)を考慮
- 話者の実発話時間は平均1.5時間
- 収録作業ノウハウの蓄積  
全9名中、3名は専門スタッフの立会い  
で音声収録をおこない(訪問録音)、6名  
は録音作業ガイドを渡して、本人(プロ、  
セミプロ)と現場スタッフだけで音声収  
録をおこなった(ユーザ録音)。  
録音作業ガイドの改善につながった。
- 声の年齢、声の高さなどの声質バリエー

ションをどれくらい用意すべきか、今後の更なる検討が必要と考えられる。

予備評価実験。

## (2) ユーザ録音

ご自身のペースで自分の声を保存し代用音声ライブラリ化を支援するユーザ作業用の標準的な録音ツールを用意するとともに、具体的な録音作業の進め方と注意点などを解説したガイドラインを作成した。

録音機器セット(標準セット)

録音作業ガイド(冊子)

- ・ 現状では、録音作業( IC レコーダ等)の経験者が身近にいること。
- ・ 将来的には、作業説明ビデオがあると良い。  
安全かつリラックスして録音をするためのノウハウ
- ・ 体調の確認
- ・ 録音ペースの調整
- ・ しっかりとした事前準備( 対面での作業説明、読み上げ原稿の事前確認等)

## (3) 自動録音、自動加工ツール ( PC ソフトウェア)

音声収録、音声ライブラリ化作業にかかるコストのさらなる削減を目指して、自動的に音声録音すると同時に音声認識を利用した自動録音、自動加工ツールを導入した。

- ・ 画面のガイドに従って発声する。
- ・ いつでも中断、再開可能。
- ・ 発声終了後、自動加工を開始し、音声ライブラリ化。
- ・ ツールの使い勝手：開発プロジェクト内部実験では、種々の改善要望点はあるが、実際上問題なし。

今後、一般健常者での実験を予定。

・ 自動加工( 音声ライブラリ化)の精度と合成音質

- 音声認識(加工処理)精度が比較的良好な場合、専門スタッフ加工との差はあまりない。
- 音声認識の精度確保には、録音品質を確保することも重要。
- 録音ボリュームの調整、リップノイズ(舌打ち、鼻息混入)対策など音声収録時の発話者による注意点がある。

## E . 開発方法

### 1 . 代用音声ライブラリの収集と評価

発話者は健常発話者を基本とし、喉頭摘出予定者やその近親者も可とする。

専門録音スタッフ立会いによる集中録音(以下、訪問録音)を含めて予定人数37名とする(3年間の総数)。

声質ライブラリ収集にあたっては、今回暫定的に声質の属性として発話者の年代と性別属性でセグメント化し、セグメント毎に数名の声を収録する。

声の必要数ならびに収集基準(属性)の検証等は、今回の検討結果を踏まえた今後の検討で明確にしていく。

代用音声ライブラリに必要な声質の広がりや分布を検討し、開発作業を効率化するための技術検討を行う。

また、代用音声ライブラリに必要な音声を効率的に収録するための発声リストについての技術検討も合わせて実施する。

代用音声の発話者は健常者プロナレータ等から選定した。発話者への同意は、失声者向け

の支援機器に限定し、編集・加工した音声データとして利用するという条件で、著作権、その他の知的財産権、利用権ならびに使用権を弊社に帰属することの同意書によっている。(有償ライセンス)

代用音声ライブラリを利用した音声合成音の評価は、発話補助装置等を使用している障害者を含めてプロトタイプモニター評価を実施する。モニター評価は、従来型会話補助支援装置で使われている一般健常者利用向けの声メッセージ(ホームページ読み上げ、カーナビ、構内連絡放送等)や、失声者の声質分布を考慮した声質ライブラリを利用した声メッセージを、会話補助支援装置利用候補者に対して、声質毎に2~3文(数秒)を試聴して頂き、自分の音声としてふさわしい声の選択回答式無記名質問紙調査を行う。

インフォームド・コンセントは、文書ならびに口頭で説明をおこなう。質問紙調査は無記名式であるため、個人情報収集しない。

## 2. 音声の収録手法、自動化ツールの評価

専属録音スタッフの立会い元での声の記録・作業は、録音スタッフコストがかかることに加えて、どうしても集中収録作業となり、本人の自然なペースでの収録や都度の追加収録などは実際上困難である。このことがPXTを利用する際の制約条件となっているため、専属録音スタッフの立会という条件をはずして、補助者、支援者を含むユーザ自身で録音可能な「ユーザ録音」ツールを開発する。また熟練専門家の手作業による音声コーパスの構築では、1人の声を処理するのに数十日の作業を要し、PTXの提供コストを押し上げる原因となっていた。そこで、ユーザ自身で録音すると同時に、PCソフトで自動的に音声コーパス化作業を行う「音声コーパスセルフ構築ツール」の導入を目指し、実運用での性能評

価と課題や改良点を洗い出す。

これらの実施において、開発対象者に対する人権擁護上の配慮、開発方法による開発対象者に対する不利益、危険性の排除や説明をおこない、同意を得た上で遂行する。

- インフォームド・コンセントは、文書ならびに口頭で説明、同意書により実施する。説明書は、参考資料に添付。
- 個人情報保護は、実験中は連結可能匿名化、実験終了後またはモニター機器の返却後は連結不可能匿名化にして5年間保存し、その後は廃棄するものとする。
- 同意書は鍵付きキャビネットに保存、データはハードディスクに暗号化して保存する。
- PC情報管理、個人情報管理、連結表は実験担当責任者を設け管理者とする。

### (1)ユーザ録音

専属録音スタッフの立会という条件をはずして、補助者、支援者を含むユーザ自身で録音可能な「ユーザ録音」ツールを開発する。具体的には、標準的な録音機器一式と発話原稿および録音作業ガイド冊子を貸し出し、発話者が自ら(補助者の支援を含む)録音するものである。今年度の評価人数は10名で、健常者7名、軽度発話障害者3名でおこなう。

最初に健常者による先行試行を実施し、課題等の洗い出しをおこなった後、軽度発話障害者で実施する。

軽度発話障害者はALS(進行性脊索硬化症)障害者をコア評価対象とし、その後他の神経系障害や一般会話補助機器使用(予定含む)者へ展開する。

軽度発話障害者の録音作業では、モニター評価を行う。評価項目は、録音作業負荷や音声ライブラリを利用した会話補助支援装置の音質満足度や会話支援効果等を障害者と介護者、支援者、御担当の方を含めて実施する。

## (2)自動録音、自動加工ツール（PCソフトウェア）

音声ライブラリー・セルフ構築ソフトで発話者が自ら（補助者支援を含む）操作し、録音と音声ライブラリ化まで行う（以下、自動録音）。予定人数31名。ただし今年度は開発者自身による機能評価のみを実施。

## F．モニター評価

### (1) 代用音声ライブラリのプロトタイプ評価（倫理審査済）

【評価手法、評価項目、対象者数、時間、モニター評価時の環境条件など】

声の年齢や声の高さなどが異なる様々な声の収録した代用音声ライブラリを利用した会話補助支援装置の声メッセージを、声質毎に2～3文（数秒）を試聴して頂き、その声質を評価する。

評価項目は、自分の音声としてふさわしい声であるかどうかの選択回答と会話補助支援装置に求められる声質についての自由意見を無記名質問調査にて実施する。

予備調査として、Webによるアンケート調査を実施する。対象者は男女計15名程度。調査時間は10～20分程度。同時に被験者を集めた対面によるアンケートの準備を進める。対面での調査時間は20～40分、対象者は男女計20名。

モニター評価に関する倫理面への配慮は倫理審査基準に則って実施。

### 【プロトタイプ評価結果】

今年度は対面アンケート形式の評価作業が整わなかったため、Webアンケート形式で予備的な評価を実施した。評価方法は、被験者に代用音声ライブラリを使った合成音を数秒程度試聴もらい、それぞれの声が自分の声質としてのふさわしさを5段階で主観評価した。

5段階評価は、

5（ふさわしい）

4（まあふさわしい）

3（どちらともいえない）

2（あまりふさわしくない）

1（ふさわしくない）である。

6種類の男性代用音声ライブラリを評価した結果、被験者男性11名の評価平均値において、6種類の代用音声ライブラリ全体の平均評価値は3.1で「どちらともいえない」レベルであったが、代用音声ライブラリ中から最良の話者を選んだ場合の評価値は4.2で「まあふさわしい」レベルであり、代用音声ライブラリから話者選択を行った場合の効果が期待できる結果となった。ただし、正規分布を仮定した有意水準5%の検定では、両者の差が有意であると判断するには至らず、更に被験者を増やす必要がある。

なお、女性3種類の代用音声ライブラリの評価は、現時点での女性の被験者が4名と少ないため、今後被験者を増やして評価する必要がある。

### (2) 進行性神経性疾患の発話障害者の適用評価（倫理審査済）

【評価手法、評価項目、対象者数、時間、モニター評価時の環境条件など】

発声障害が顕著でないステージ初期の進行性神経性疾患の発話障害者に対して、音声収録モニター作業と「自分の声」利用型の会話

補助支援装置の使用評価を実施する。調査項目は音声収録作業負荷、会話支援装置によるコミュニケーションの利用満足度、失声不安に関する対面アンケート調査と対面ヒアリングである。標準調査時間は30分、対象者は2名。モニター評価に関する倫理面への配慮については、倫理審査基準に則って実施。

#### 【評価結果】

##### モニターケース1

発話機能が軽度低下障害においても、1日2時間の収録で、話者性を満足できる音声合成が可能

今回のモニター本人の音質レベル評価は、家族や知人などにインフォーマルな会話・談話をする場合は使えると感じる。それ以外のフォーマルな会話や電話などの使用はもう少しいろいろな文章の音声合成音を評価しながら、判断したいというものであった。

また声の記録・保存という観点では、録音時点での声が保存・確保されているので、今後の更なる発話機能の低下に対して大きな安心感があるとの結果を得た。

##### モニターケース2

発話機能の低下はないが、発話時間が一日の内の短時間しか確保できない場合について、家族の協力での自宅録音でどの程度の音声作業ができるかを検証した。本ケースでは本人と家族の生活に支障ないレベルでの録音作業は週に30分程度であり、数週間にわたる録音期間が必要であったが、録音作業自体は完了した。今後録音品質の確認作業をしていく。

## G. 開発で得られた成果

### (1) 代用音声ライブラリ

自分の代わりに喋ってくれるふさわしい声

質の供給源としての代用音声ライブラリを試作し、利用できるための道筋を確認できた。

### (2) ユーザ録音

音声収録支援システムの整備では、ユーザ録音による収録作業の制約緩和実験を実施し、収録経験のある健常者においては、作業上の問題がないことを確認した。軽度の発話困難障害者でも、録音機器や録音作業を工夫することにより、自分の声を記録・録音できることを確認した。またこの音声ライブラリ利用による自分の声の合成音は、本人性や肉声感が感じられるレベルにあり、コミュニケーションの品質向上に寄与する可能性を感じていることをモニター評価で確認した。

### (3) 自動録音、自動加工ツール（PCソフトウェア）

開発者自身による機能評価を行い、合成音の品質において本人性や肉声感が感じられるレベルであること、また専門スタッフによる人手の加工処理による合成音と比較してもその品質差が少ないことが確認された。

開発における当促進事業が有用であった点

#### 【参考】

### (1) 開発内容・計画について構成員からのアドバイスの効果

進行性の発話障害者に関する知見と、本装置の適用・利用方法についての理解が深まったこと。

先天性や一過性の発話障害者に関する知見と、本装置を含む会話補助支援装置の利用方法についての理解が深まったこと。モニター評価の進め方や倫理審査の進め方について、具体的なお指導が頂けたこと。

## (2) 独自開発との差

代用音声ライブラリ開発、音声収録支援システム開発、モニター評価の実施は困難であり、利用者拡大、利用制限緩和、提供価格低減等の課題解決の取り組みを進めることができなかつたと考えられる。

## (3) 促進事業予算の効果ならびにモニター評価および社会的効果

医療機関との連携による進行性発話障害者モニター評価

専門家を含むナレータとの商用利用契約（障害者向け支援機器）の（低価格）締結

## H. 予定してできなかったこと

### (1) 進行性発話障害者の訪問録音と音声ライブラリ製造

予定 2 名の完了に対して、達成度は 75% となった。

- ・ 1 名分については予定通り、訪問録音ならびに音声ライブラリ製造を完了。
- ・ 1 名はは先方希望・医療関係者確認によりユーザ録音作業中。先方の希望ペース(30分/週)によるユーザ録音作業で録音作業が3月末完了。録音品質確認ならびに音声ライブラリ製造は今後実施予定。

### (2) 進行性発話障害者モニター評価

計画内容に対して、達成度は 80% となった。

- ・ 倫理審査は完了。
- ・ モニター開始条件の整った 1 名に対してモニター実施中。
- ・ 初期モニター評価は完了。
- ・ モニター都合により開始時期が遅れたため、現在もモニター評価継続中。  
モニター最終評価は 4 月下旬を予定。

### (3) 代用音声ライブラリの音声利用可能性モ

## ニター評価

計画内容に対して、達成度は 60% となった。

- ・ 倫理審査は完了。
- ・ Web アンケートによる予備評価分析中
- ・ 対面モニター作業準備を継続中。対面モニター募集をお願いする NPO 団体殿との意識すり合わせと調査同意確認が遅れているため、対面モニター調査時期は未定。

## I. 考察

(1) 代用音声ライブラリ等の話者選定(詳細は参考資料参照)

- ・ 音声の話者性を表現する軸は高々 2 次元が現実的。次元の例としては、声の高さ(ピッチ)や話速など。
- ・ 話題や音声合成音の利用シーンを指定した場合の話者選定事例を紹介。
- ・ 汎用的な利用シーンにおいての紹介事例の手法の有効性は未検証である。

(2) 話者性評価方法 (詳細は参考資料参照)

- ・ 本人性らしらの従来指標をサーベイ。
- ・ 本人性らしらの評価事例の紹介。
- ・ 合成音がどれだけ本人の声に似ているか、どれだけ本人を想起させるか、どれだけ本人に受け入れられる音声であるかを評価している。
- ・ 評価方法は原音と合成音とを 1 対比較し、評価指標に沿って、5 段階で評価する。

以下評価値の抜粋。

4: 一部、再現できていない箇所を含むが、全体としては、ほぼ再現できている。(再現できていない割合<再現できている割合)

3: 一部、再現できている箇所を含み、全体としても、再現できていると感じる。(再現できていない割合 = 再現できている割合)

#### ・ 評価結果

1.5 時間程度の音声ライブラリ 6 名分の自動録音、自動加工ツールによる音声合成音の話者性（本人の声質の再現度）の被験者 10 名の平均主観評価値は 3.5 となった。また、自動加工による音声ライブラリと人手による加工の音声ライブラリとの本人性評価指標は差がない。

#### (3) 収録原稿（詳細は参考資料参照）

- ・ 収録原稿の音素バランスおよびアクセントバランスがとれていることが必要である。
- ・ 原稿の読み上げやすさの観点からは、文の長さ、読点の数・位置、構文、言葉の難しさ、話し言葉への近さ、分野への親しみ度合い、意味的な自然さ、などを指標として原稿を構築する必要がある。
- ・ また、よく使う語やフレーズを切り貼りして利用することも考えられるため、上記バランスのとれた原稿とは別に、自由に語やフレーズを追加できるようなデフォルト原稿と仕組みを用意しておくことが重要となる。

#### (4) 進行性発話障害者モニター評価

- ・ 本人、親族の評価ともに本人性は再現できている。
- ・ 会話文の品質に課題あり。
- ・ 会話文の品質を確保するキャッシュの作り方に工夫が必要。
- ・ キャッシュの重要性は事後に判ることが多い。追加キャッシュ機能が必要。
- ・ アクセントがおかしい。標準語ではないと感じる。

アクセントの修正機能が望まれる。そのためにはアクセント修正ツールに加えて、アクセント指定音源となる音声データの追加録音等が必要となる。

#### (5) 代用音声ライブラリの音質評価

6 種類の男性代用音声ライブラリを評価した場合、被験者男性 11 名の評価平均値において、6 種類の代用音声ライブラリ全体の平均評価値は 3.1 で「どちらともいえない」レベルであり、分散は 0.12、危険率 5% の信頼区間は 2.82 ~ 3.37 であった。代用音声ライブラリ中から最良の話者を選んだ場合の評価値は 4.2 で「まあふさわしい」レベルであり、分散は 0.38、危険率 5% の信頼区間は 3.32 ~ 5.04 であった。代用音声ライブラリから話者選択を行った場合の評価値が大きいが、正規分布を仮定した有意水準 5% の検定では、両者の差が有意であるとは判断できず、もう少し被験者を増やす必要がある。

なお、女性 3 種類の代用音声ライブラリの評価は、現時点での女性の被験者が 4 名と少ないため、今後被験者を増やして評価する必要がある。

・ 今回はあくまでプロトタイプ予備評価の位置づけであり、今後実験条件や被験者数をそろえた評価が必要である。

## J. 結論

発話困難者や失声者が人間味のある代用音声を獲得、利用できるための道筋が確認できた。

・ 発話困難者や失声者向けの代用音声ライブラリを試作し、ライブラリ利用の合成音の予備的な品質評価を実施した。

・ 軽度の発話困難障害の状態でも録音機器や

録音作業を工夫することにより、自分の声が残せる場合があることをモニター作業により確認した。声の合成音の本人性や肉声感は感じられるレベルにあり、被験者はコミュニケーションの品質向上に寄与する可能性を感じている。また声の保存ができているという安心感も得られている。

・合成音の利用面では、発音アクセントの指定や本人の生録音声のはめ込み利用などの音質面の向上の仕組みに加えて、利用者のPCスキルや障害特性に応じたインタフェースなどのソフト面での工夫が望まれる。

## K．健康危険情報

本研究では、標準的なマイクを使用した音声収録、PCのキーボード操作によるものであるため、開発者側、当事者側共に健康に及ぼす危険はない。

## L．成果に関する公表

1．ホームページ、刊行物等の紙面などでの発表

開発社、開発協力社のホームページで公開中。

・ ウォンツ ホームページ

<http://www.wantsinc.jp/Solution8.html>

・ テクノマインズ ホームページ

<http://www.technominds-jp.com/>

2．展示会などでの発表

厚生労働省主催の一般公開(2011年3月9日)で公開

## M．知的財産権の出願・登録状況

1．特許取得 該当なし

2．実用新案登録 該当なし

3．その他 該当なし

## **. 開発成果の公表に関する一覧表**

開発社、開発協力社のホームページで公開中。

- ・ ウォンツ ホームページ  
<http://www.wantsinc.jp/Solution8.html>
- ・ テクノマインズ ホームページ  
<http://www.technominds-jp.com/>

## **. 開発成果の公表に関する刊行物・別刷**

該当なし

## ．参考資料

### 1．代用音声ライブラリ等の話者選定ガイドライン

#### 第1章 はじめに

#### 1. 人間味のある声を持つ(持てる)ということ

- ・ 声は、個々人の存在を示し、その人であることを識別する手段である。
- ・ ふさわしい声により自分自身の存在を確認するとともに、行動への自信につながる(Polluxstar ユーザの声)
- ・ 現状、発話機能を失った方にとっての代用音声手段は、食道発声、電気喉頭、シヤント発声、VoCA など
- ・ 人間味のある声を利用できる状況が整備されている社会
  - さまざまな方にとって、自身がふさわしいと感じられる声を持てる状態
  - 発話機能に不安を覚えたとき、人間味のある声が用意されている状態

#### 2. 人間味のある声の整備にむけた話者性(本人性)定義

- ・ 話者性とは
  - 音声に含まれる個人性である
  - 誰の声かわかること
  - その人らしさ
- ・ 話者性をもたらす要素とは
  - 身体的特徴、生い立ち
  - 言語的特長：母国語・方言
  - 文化・所属組織
- ・ 合成音声における話者性の評価観点(大枠)
  - 合成音がどれだけ本人の声に似ているか
  - どれだけ本人を想起させるか
  - どれだけ、その声に納得できるか

#### 3. 話者設計の既存(現有)技術

##### (1)JEITA 規格 IT-4005

<http://www.jeita.or.jp/cgi-bin/standard/list.cgi?cateid=8&subcateid=46>

「音声認識エンジン性能評価方法のガイドライン」

##### 5.3.2 不特定の話者を対象とした音声認識

性別・年齢・出身地(地域性)・健康状態  
発声の状態(発話内容のゆれ、発話速度、  
発話音量、発話の自然性、読み上げ/主体的は  
発声、協力的音声、疲労など)

##### (2)宇都宮大学工学部 粕谷 英樹

<http://ci.nii.ac.jp/naid/110003283799>

「声質評価システムの構築 Construction of the evaluation system of voice quality」  
声の質を表す語として、「高い」、「若い」、  
「落ち着きのある」など、さまざまな表現語が  
日常的に用いられている

過去の著者らの研究において、音声研究者ではない一般の人々が用いる表現語を対象に、通常発話の声質に関連した日常表現語を抽出した。これらの声質表現語と音響関連量との関係を交差検証法による決定木分析で調べ、人間の聴取印象を推定するプログラムと聴取印象による話者検索プログラムからなる声質評価システムの構築を試みた。その結果、聴取印象の推定では、かなり誤差が小さく抑えられ、話者検索においては、上位 10%内に 94.1%を示すなど、良好な性能が得られた。

##### (3)電気通信大学 吉田研

<http://www.lib.uec.ac.jp/limedia/dlam/M463317/2.pdf>

「声質検索システムの構築」

多数話者音声データベースの中から本人の声に最も似ている話者を選び、その話者の音声データを用いて音声合成をするという方法である。1,035 名の話者を擁する音声データベースを用いて、その中から目的の声質を検索する方法を検討している。その結果、声が似ているかどうかを判断するには、何らかの聴覚的尺度(声の高さ、かすれ具合等)による判断基準を設けることが有効であること、SOM という、特徴の似たものは近くに、似ていないものは遠くに配置する 2 次元マップを利用することで、声質検索の速さが上がることが分かった。

#### 4. 既存技術のまとめと課題

- ・ 既存技術のまとめ
  - すでに大量の声質ラインナップがある時に、その中からよりよい声を選ぶときの指針になる重要技術である。
  - ただし、あくまで、そこに存在する大量の声質の整理法を示すものである
- ・ 課題
  - 既存技術は、例えば以下のような現場での構築ニーズを満たすものではない
    - ・ 構築すべき声質・話し方について、関係者間で、どのようにその情報を設計段階から共有・具体化していけばよいか
    - ・ 収録話者に対して、収録現場でどのような収録指示をおこなえばよいのか
- ・ 本報告の位置付け
  - 狙った声質・話し方を関係者間で整合し、実際に設計・構築・評価した既存事例について紹介する。
  - 上記事例について、品質作り込みの手順とその限界を示すことで、上記事例での手法のメリット・デメリットを明らかにする。

## 第2章 事例

### 1. 事例の概要

- 事例における構築概要
  - ・ 音声対話システムの構築
- 音声合成機能に関する実証内容
  - ・ システムにふさわしい声を設計し、提供すること
    - 寄り添い感のある声 など
  - ・ 対話インタフェースとしての音声出力部として、対話効率および内容の深さを促進させる(あるいは阻害する)要因を特定し、対処すること
    - 対話ターンの自然さ など
- 話者選定に関する検討内容
  - ・ 合意プロセス
  - ・ 作りこみプロセス
  - ・ 評価プロセス

### 2. 話者選定・コーパス構築のフロー概要

以下の流れで、話者選定し、コーパス構築・評価までをおこなった。それぞれのステップでの作業について紹介する。

#### 1. オーディション

##### 目的

- ・ 音声の具体イメージ化・関係者間での整合
- ・ 収集・利用可能な発話者の特定

##### 実施概要

- ・ 使用目的を明らかにし、商用利用同意を得られる話者の募集
- ・ 音質および印象をチェックできる短文原稿による簡易収録の実施
- ・ 大規模収録に備え、発話再現可能な声質・口調のチェック
- ・ 収録費用の概算見積もり

#### 2. 話者選定(バリエーション設計)

## 目的

- 声質・話し方の特徴の漏れ・ダブリ防止
- 声質構築優先度の決定、作業計画の立案
- 構築後の評価観点の明確化

## 設計観点

既存技術・保有ノウハウを参考にし、観点を洗いだす

- 観点の例
  - ・ 収録話者の属性
    - 年代・性別・出身地域（母語）
  - ・ 音響パラメータ
    - 声の高さ（ピッチ）・話速・スペクト
    - 上記の平均・ダイナミックレンジの広さ、あるいは統計的グループ化
  - ・ 発話シーン
    - うれしい／悲しい、 明るい／暗い、 フォーマル／カジュアル など

## 軸設計

- オーディション話者を見渡し、軸をさだめ、軸にネーミングをおこなう。

（実施してみたところ、管理可能な次元は、高々2次元であった）

複数話者（バリエーション）で実現可能な空間全体を定義する。
- オーディション話者をグループ化し、それぞれのグループに名前をつける。

漏れ・重複を減らすようグループを見直し、構築にあたっての優先度を選定する。

## 話者選定

- 過去の構築経験をもとに、各話者の仕上がり品質・工数を予測する。
- 話者候補順を、効果性・効率性の両面から、定める。

## 3. 話者作りこみ

作りこみにあたっての考え方

- コーパスベース音声合成では、収集（収録）音声から統計学習で構築された韻律（話者）モデルを構築する。
- 音声の収録時間が十分でない場合、収集に用いた原稿の文構造と、音声合成の際に入力される文の構造との差異を少なくすることで、音声合成の構築から利用にわたる全体のコストパフォーマンスを高めることが可能となる。

## 作り込み戦略

2つの観点でのアプローチが必要

- 利用シーンにあった音声を集める仕組み  
頻出語・重要語を集める仕組みを入れ  
収録原稿を、利用シーンに合わせる
- 収録音声と、音声合成時の入力文との間の、文構造を合わせる仕組み  
入力文を、収録原稿の文構造に合わせる  
仕組みを入れる  
利用シーンを限定することで、音声品質  
保証の範囲を明確にする

## 4. 話者選択（バリエーションからの選択）

### 評価観点

- 話者の評価
  - ・ 違和感  
主に定量的なギャップ、年代、性別など
  - ・ 納得性  
定性的、本人・自分らしさを感じられるか
  - ・ 文構造  
定性的、本人らしい文表現や区切り・息継ぎ位置など
- 話者カバレッジ評価
  - ・ 話者バリエーションの観点で、漏れ・ダブリが少ない声質・話し方に仕上がっているか
  - ・ 今後、ラインナップとして充実させるにあたり、現状の話者軸設計を見直す必要があるか

## ➤ 音質評価

### 第3章 考察と結論

#### 全体について

- ・ 一般的にコーパスベース音声合成では、音質保証と構築コストとがトレードオフの関係にある。
- ・ 特に、声質のバリエーションの多様さに価値がある場合に、取り揃える話者数と、それぞれの話者音質の保証レベルとをどこに定めるかが、非常に難しい問題となる
- ・ ここで、音響处理的対処（エフェクト処理あるいはスペクトル平滑化など）にて、音質保証をおこなうことが1つの解となるが、この場合、そもそもの「人間味のある声」が、機械的音質に劣化するため、本プロジェクトの趣旨にあわない解決策である場合がある。

#### 本書にて紹介した事例手法の長所

- ・ ある1つの、固有のドメインおよび利用シナリオに特化した場合に、有効である。典型的な利用シーンを洗い出し、そこで対話内容のカバレッジを上げることで、

利用アプリケーションのタスク達成率およびタスク達成効率を効率的に上げていくことが期待できる。

- ・ 各ユーザごとの利用シナリオ特化の仕組み（自動化）などが今後期待できる。
  1. 実際の対話ログ分析の結果をもとに追加収録原稿として有効な文をピックアップする
  2. 上記有効文を追加収録する
  3. 追加収録音声もコーパスに追加する、などの機能を持たせることで、自己成長型のシステムに向かうことが期待できる。

#### 本書にて紹介した事例手法の短所

- ・ 汎用的な利用における音質向上については、構築作業の効率を高められない。あらゆる利用場面における利用アプリケーションのタスク達成率およびタスク達成効率を評価指標とした場合、その指標向上にむけた製品作りこみの手段にはならない。

## 2. 話者評価方法ガイドライン

### 第1章 はじめに

#### 1. 音声における話者性（本人性）

##### 1-1 話者性とは

音声に含まれる個人性であり、声質から誰の声かわかることである。また、その人らしい話し方を指すものである。

##### 1-2 話者性をもたらす要素

- ・ 身体的特徴（性別・年齢・病歴・体調など）
- ・ 言語的特長（母国語・方言・文化・風習・符丁など）
- ・ 意図の表現（語尾表現、ため息、感情・意図に対応する抑揚など）

##### 1-3 音声合成における話者性

- ・ 合成音がどれだけ本人の声に似ているか
- ・ どれだけ本人を想起させるか
- ・ どれだけ本人に受け入れられる音声か

#### 2. 話者性の評価

##### ・ 従来技術

###### 主観評価

MOS評価など

###### 客観評価

ピッチ・話速・スペクトル

話者認識における認識パラメータ

##### 事例

キャラクター性の評価等がある。

#### 3. 話者性についての課題

- ・ 自分自身の声に対する評価方法が確立されていない。
- ・ 被験者が本人を知っているかどうかにより、大きく結果が異なる。
- ・ 自然性・明瞭度と話者性を分離して評価することが現実的に難しい。

### 第2章 事例

自動録音、自動加工ツールの性能と話者性評価の方法について、自動収録ライブラリ化ソフトの性能を測るために行われた評価試験の内容を結果と共に報告する。

#### 1. はじめに

##### 1-1 音声における話者性（本人性）

###### (1) 話者性とは

- 音声に含まれる個人性
- 誰の声かわかる声質
- その人らしい話し方

###### (2) 話者性をもたらす要素

- 身体的特徴  
（性別・年齢・病歴・体調など）
- 言語的特長  
（母国語・方言・文化・風習・符丁など）
- 意図の表現  
（語尾表現、ため息、感情・意図に対応する抑揚など）

###### (3) 音声合成における話者性

- 合成音がどれだけ本人の声に似ているか
- どれだけ本人を想起させるか
- どれだけ本人に受け入れられる音声か

##### 1-2 話者性の評価

###### (1) 従来技術

###### 主観評価

MOS評価など

###### 客観評価

ピッチ・話速・スペクトル

話者認識における認識パラメータ

###### 事例

キャラクター性の評価等

###### (2) 話者性についての課題

- 自分自身の声に対する評価方法が確立されていない。
- 被験者が本人を知っているかどうかにより、大きく結果が異なる。
- 自然性・明瞭度と話者性を分離して評価することが現実的に難しい。

### 1 - 3 評価項目

評価したい項目として以下の5項目が挙げられる。

#### 話者間

特定の性別、年齢に関係なく、様々な話者で構築した場合の性能差。

#### 境界線データ精度 1

境界線データを自動または手動で付与した場合の性能差。ただし、全体に対する境界線データ量の割合は区別せず。

#### 音声データ量

データベース構築に必要な音声データを小規模にした場合の性能差。

#### 境界線データ量

手動で付与した境界線データを小規模にした場合の性能差。

#### 利用シーンに応じた原稿

利用シーンのカテゴリを細分化した場合の性能差。

これら5項目の中から、、、の3項目を評価対象として絞り込む。

1:境界線データとは、音声データベース構築の前段階におこなう、音素分けにおける、各音素の境界線を定めるデータの事である。

### 1 - 4 評価性能

ここで評価したい性能差は、合成音声の『再現性』指標により測定する。

『再現性』の中には、『話者・口調の雰囲気』の再現』と『完全性(なめらかさ、不連続)』の2種類の視点が含まれる。

さらに、『話者・口調の雰囲気』は、『声質』と『イントネーション』に分割される。

これにより現時点での各指標の項目は、以下の3つとなる。

『声質』

『イントネーション』

『完全性』

それぞれは、独立なものとみなし、評価をおこなうものとする。

### 2 . 評価項目・性能の仮説

#### 話者間

特定の性別、年齢に依存せず、『声質』の再現度は一点に集中するだろう。

早めの話速と感じる話者、または、女性話者の『イントネーション』、『完全性』の再現度はやや低いだろう。

また、他の『イントネーション』、『完全性』の再現度は同等になるだろう。

#### 境界線データ精度

自動、手動付与の『声質』の再現度は同等になるだろう。

自動付与の『イントネーション』、『完全性』の再現度は、手動付与の再現度と比較すると、やや低いだろう。

#### 音声データ量

通常の音声データ量に比べ、小規模なデータ量の『声質』の再現度は同等になるだろう。

小規模な音声データ量の『イントネーション』、『完全性』の再現度は、通常に比べると低いだろう。

さらに、音声データ量を増やすごとに、通常の再現度に近づくだろう。

### 3 . 評価方法

#### 3 - 1 評価実験 仕様

以下4種類のカテゴリから、記憶に残りやすい長さ十数秒を目安にした原稿を作成する。

### < カテゴリ >

- 「ポジティブな話」
- 「マニュアル説明」
- 「講演」
- 「無意味な文」

サンプル音声（刺激音声）は、  
 < 原音 + 一定時間の無音 + 合成音 >  
 で構成される 1 対セットとする。  
 原音と合成音内の、話速、読点位置は同じに  
 合わせる。

サンプル音声の提示順はランダムに行う。

#### 話者間

境界線データ精度、音声データ量を共通とし  
 た男女複数のサンプルとする。

#### 境界線データ精度

手動付与の境界線データ量は、自動付与全体  
 の 4 割程度とする。

#### 音声データ量

音声データ量を、100,200,400 文とする。（文  
 の選定に工夫なし）

### 3 - 2 評価手法

原音と合成音とを 1 対比較し、評価指標に沿  
 って、5 段階に評価値を与える。

#### 『声質』

再現できているか？

#### 『イントネーション（口調や話速含む）』

再現できているか？

#### 『完全性（なめらかさ、不連続）』

再現できているか？

### 5 段階の評価レベル値

5	原音を聞いているように再現できている。
4	一部、再現できていない箇所を含むが、全体としては、ほぼ再現できている（再現できていない割合<再現できている割合）。
3	一部、再現できていない箇所を含み、全体としても、再現できていないと感じる。（再現できていない割合<=再現できている割合）逆に、一部、再現できている箇所を含み、全体としても、再現できていると感じる。（再現できていない割合>=再現できている割合）。
2	一部、再現できている箇所が存在するが、全体としては別人、又は、再現できていない。（再現できていない割合>再現できている割合）。
1	別人のようだ。または全く再現できていない。

### 3 - 3 被験者毎の評価値の正規化

評価値の基準バラツキをなくすために正規化  
 を行う。

#### < 手順 >

5 段階の評価値を、被験者と話者毎にグルー  
 プ化して、平均値と偏差を求める。

その平均値と偏差を用いて、正規化した値を  
 求める。

### 4 . 結果

#### (1) 「話者間」

特定の性別、年齢、話速に依存せず、『声質』、  
 『イントネーション』、『完全性』の再現度は  
 同等に近い結果となっている。

さらに、声質 完全性 イントネーションと  
 いう傾向も、どの話者においても生じている。  
 一部の話者において、声質 イントネーショ  
 ン 完全性という傾向が生ずるが、境界線デ

ータの精度が一番の要因と考えられる。

## (2) 「境界線データ精度」

自動境界線データの精度が比較的によく、かつ、データベースに含める手動付与の割合が半分以上という条件下で、『声質』、『イントネーション』、『完全性』の再現度は同等になり、自動境界線データの精度を向上させる重要性もわかる。

## (3) 「音声データ量」

音声データ量が小規模になるほど、『声質』、『イントネーション』、『完全性』の再現度が低くなる。また、傾向も同じとなる。ただし逆に音声データ量を増やすごとに、通常の再現度に近づくことも分かった。

## 5 . 技術課題

### 収録

ヘッドセットマイクの使用を想定している。プレスノイズやリップノイズが含まれた音声への処理対応は自動ではおこなえないため、発話者自身が注意を払う必要がある。同様に背景ノイズが含まれた環境に対応ができていない。

また、何日かの収録に分けた場合の性能評価は得られていないが、発話者への注意点を与えることが可能である。

### 原稿

原稿の量と質ともに、現在の原稿より、音素バランスが整った、少量でかつ、読みやすくしていく必要がある。

### 構築処理

読み誤りが生じた場合のカナ修正に対応していない。そのため、音声と表記内に含まれる間の位置を正確に対応できなくなる。自動境界線データの精度を更に向上させることが望まれる。

### 手動境界線データの反映

自動化が目的の1つであるため、手動境界線

データへの対応は考えないものとしたいが、手動境界線データの反映に対応する機能を具備することが検討に値する。

### 評価方法

主観的な評価の進め方を見直す必要がある。合成音の評価によくあることで、「評価基準が曖昧である」、「判断に迷う音声をどのように扱うのか」といった課題がある。

## 第3章 考察と結論

より良い話者性評価のために、評価結果から得られた今後へ向けた改善ポイントを以下に示す。

### ポイント1

話者再現性を、声質、イントネーション、完全性の3つに分けた。

互いに独立としたが、声質・イントネーションは話者再現性であり、完全性は（誰の声かは別として）人間の肉声としての再現性である。完全性が欠落した際、影響がイントネーションに及ぶ可能性が大きいことがある。

#### < 提言1 >

完全性と、話者再現性を分離して考えるとよいと思われる。

完全性が一定スコア以上のデータに対して、話者再現性を評価（集計）する。

その話者の完全性（一定スコアを超える率）は、別途評価する。

Ex)話者再現性 = X.XMOS 完全性 = 80%

### ポイント2

話者の好みをなくすため、スコアの正規化を実施した。

原音との1対評価（C-MOS）なので、話者の好みによる影響は少ないのではないかと判断できる。

好みは、各スコアではなく、イントネーションと完全性の振り分け加減に落ちることもあ

と思われる。被験者にとって、話者の声が既知（知人・著名人等）か否かがスコアに影響してしまうと思われる。またリファレンスだけでなく、被験者の頭の中にある話者の声がスコアに影響するかも知れない。

< 提言 2 >

原音との違いを評価するという趣旨をしっかりと伝える必要がある。

実験の冒頭で、話者の「好み」ではないことを強くアナウンスする。

あえて評価項目に「好ましさ」等を設け、話者の好みはこの項目で吸収するようにする。等の考慮が必要である。

ポイント 3

被験者の評価基準に重きを置いて主観評価を行った。

技術課題にもあるように、評価基準が曖昧になる。主観評価である以上避けがたいが、ある程度の統一性は必要である。

< 提言 3 >

主観評価実施前に、声質が不十分な例として、

イントネーションが不十分なものなどを具体的な合成音サンプルとして、いくつか聞かせる。

実施側での評価セットを、ある程度設計する必要がある。

ポイント 4

・本人性の評価にあたっては、これはある側面では評価結果へのバイアスと

なりうるものであるが、評価者の収録への立ち会い関与度合いが、

できあがり品質の満足度の評価値に大きな影響を及ぼすことがわかった。

この効果について検討を深め、この効果を生かすような製品構築のプロセス

（社会面を含めた、ユーザとの製品構築時のフィードバックを得る仕組み）を

考えることが、製品の幅広い展開に向けた1つの解となりうる。

### 3. 収録原稿に関するガイドライン

#### 1. はじめに

音声コーパスを作成するための収録原稿の要件について検討を行った。

コーパスベース音声合成に適用するためには、原稿の音素バランスおよびアクセントバランスがとれていることが必要である。

また、よく使う語やフレーズを切り貼りして利用することも考えられるため、上記バランスのとれた原稿とは別に、自由に語やフレーズを追加できるようなデフォルト原稿と仕組みを用意しておくことが重要となる。

原稿の読み上げやすさの観点からは、文の長さ、読点の数・位置、構文、言葉の難しさ、話し言葉への近さ、分野への親しみ度合い、意味的な自然さ、などを指標として原稿を構築する必要がある。

読み上げやすくするために収録現場で原稿を加工する場合があるが、あらかじめ加工基準を定めておき、収録データとともに適用した基準を保持する仕組みを用意することにより、合成時に適切に情報を利用することができる。

以下、詳細に記述する。

#### 2. コーパスベース音声合成(CTTS)の合成音生成の観点から

Polluxstar や XIMERA などの CTTS の収録データとしては、一連の語から構成される文（あるいは節や句）単位での原稿を用意する必要がある。

この原稿は、読み上げた際にさまざまな音素をできる限り網羅するように音素バランスが考慮されている必要がある。これにより合成音の自然さを担保する。

同時に、さまざまなアクセントのパターンを網羅するようにアクセントバランスが考慮さ

れていることが望ましい。これによりイントネーションの自然さを担保する。

これらの条件を満たすような読み上げ原稿の量は、最低でも 200～300 文（1500 語程度）は必要である。しかし、音素・アクセントのバランスが完全にとれた原稿を用意するのは困難なので、できれば 1000 文程度を読み上げてもらふことにより、結果的にある程度の音素・アクセントがカバーされるという方法をとるのが現実的である。

現在の Polluxstar に標準添付されている原稿は、音素バランスについてはある程度考慮されているが、アクセントバランスについては十分に考慮されていない。合成音声のイントネーションに違和感を覚えるという指摘を受けるのはそれが原因ではないかと考えられる。アクセントバランスに配慮して原稿を改修・追加していくことが望まれる。

なお、原稿を改善していくにあたり、原稿全体の音素バランスおよびアクセントバランスを計測するためのツールを用意しておくことで作業が効率的に進むと考えられる。このようなツールは、収録現場で自由に原稿を追加・変更するような事態が生じた場合にも有用である。

#### 3. 録音した音声をそのまま利用する観点から

CTTS であっても、よく使う語やフレーズに関しては録音した音声をそのまま貼り付ける手段をとる場合があり、ユーザ満足度の向上には有効である。

この場合、既存の原稿に加えて、バランスを考慮せずに自由に語を追加することができる。日常語なら 300 語程度録音しておけば効果が期待できるのではないかと、ユーザの一人である牧先生からのコメントがある。

また、長時間の読み上げに対応できない利用者のためには、CTTSではなく、単音節の録音音声と語・フレーズの録音音声を切り貼りする方式の音声合成を採用することも考えられる。録音時間を最少に抑えつつ、(十分ではないが)本人の声質を多少なりとも残すことができ、ソフトウェアとしてのコストもユーザの負担も軽減することができるため、普及型の音声合成エンジンとして一考に値する。このようなエンジンの場合、時間がとれるときに単語やフレーズを追加することができれば合成音の質が向上するため、録音および音声データの合成音への反映をユーザ自身で行える仕組みを用意することにより満足度を高められる。

#### 4. 原稿の読み上げやすさの観点から

##### (1) 指標に関して

読み上げの難易度に関連がある指標としては以下のものが考えられる。

##### ➤ 文の長さ

音声の収録はある程度の単位で実施するため、通常は一文を読み上げてもらうことになる。このとき、文が長いと(全部を頭に入れることができないせいか)読み間違ふことが多くなる。理想的には漢字かな混じり文で20字程度に抑えることが望ましい。ATRが用意したXIMERAの標準原稿は長文のものが多く、読み上げが大変だとのコメントが得られたため、Polluxstarでは文を短くするよう工夫されている。

##### ➤ 読点の数・位置

実際に発話文を合成するにはある程度の長さの文を対象とすることもあるので、長い文をいくつか登録したい場合もある。その際は、(意味的には読点をはさむ必要がない位置で

も、読み上げの容易にするために)適切な位置に読点をはさんでやる必要がある。

##### ➤ 構文

文の構成も読み上げには影響がある。原稿はなるべく単文(述語を一つだけ持つ文)であることが望ましい。「～して、…する」のような重文や、「～した が…する」のような複文は読みにくい。特に複文は構造が複雑になりがちで、読み上げ時に間をとりにくいので避けるべきである。

関連して、修飾語や修飾部が長いもの・多いものも読み上げを困難にすることが多い(複文は長い修飾部の一種である)。

##### ➤ 言葉の難しさ

使う用語は平易なものであることが望ましい。漢字やひらがなのどちらかに偏った用語はなるべく避けたい。個人差もあるが、ふりがながないと読めない語は原稿を目で追う際に見間違える要因になる。同じ意味で、読みが複数ある用語はできれば排除したい。

(読み上げ者は不特定多数なので事前に設計しておくことは困難ではあるが)なるべく読み上げ者が普段から使い慣れている語を用いることが望ましい。

##### ➤ 話し言葉への近さ

声を出して読むことが前提でない文章を原稿に流用すると、発声には不適な語の組み合わせが用いられていたりするので、できれば声を出して読んで不自然でない文章を用意したい。

このとき、友達同士の対話のようにあまりに話し言葉に寄った文章を原稿とすると、現状の日本語(形態素)解析の技術的問題によりうまくいかない場合があるので注意が必要である。

##### ➤ 読みなれている分野か

原稿が、読み上げ者が読みなれている分野の内容であるとよい。用いられている用語が使い慣れている言葉であるということも含める。読み上げ者が内容に精通しているものであればなおよい。これにも個人差があるので、現実的には様々な分野の原稿を用意しておき、読み上げ者によって選択できるようにしておくのがよい。

➤ 意味的におかしくないか

読みなれている・いないにかかわらず、適当に作文した文は意味的に不自然な場合があるので注意が必要である。何度も原稿を読む可能性があるため、意味的な不自然さは少ないほうが読み上げ者のストレスは少ない。

(2) 原稿の現場での加工に関して

長文を読み上げるのに苦労がある場合、収録現場で文章を分割するという手段を取ることがある。

合成のためのコーパスを作る際には分割したものを結合する必要があるが、収録時にどこで分割したのかの情報を保持していることが精度の向上につながるが、実際にはその情報を残しておくのは現場では難しい。

そのため、原稿を分割するための基本的な指針をあたえておき、現場で対処する場合にはその指針に従うようにしておけば、結合をする際にもデフォルトでそれを考慮することができる。

4. 代用音声ライブラリ開発に関する実証試験研究計画書

(ヒトを対象とする支援機器の実証試験)

**実証試験研究計画書**

作成日 2010年 10月 4日  
 修正日 2010年 11月 5日

作成責任者 氏名 三木 敬  
 所属・職名 合同会社テクノマインズ 代表社員

**1. 実証試験研究課題**

研究課題	言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発
研究期間	平成 22年 9月 1日から平成23年 3月 31日まで

(A) 研究組織

	氏名	所属・役職・職種	分担項目	連絡先
研究代表者	川島 紳		実験全般管理	
実験担当責任者	本田武司		実験実施管理	
分担研究者	三木 敬		実験技術管理と評価作業	
助言を担当する医師	川田明広 長尾雅裕			

(B) 共同研究実施機関・組織・施設・研究実施場所

機関・組織名	実施組織・場所	実施内容	倫理審査状況
該当なし			

(C) 研究協力機関

機関・組織名	実施組織・場所	実施内容	倫理審査状況
都立神経病院 脳神経内科		(必要に応じ) ユーザの声の録音作業	申請予定
(D) 研究資金	厚生労働省 障害者自立支援機器等研究開発プロジェクト		

## 2. 研究の概要

### (A) 支援機器の目的・目標

発声障害者用の会話補助支援装置として、さまざまな音声を発する機器が提供されているが、人格が見えないメッセージ的な声質ではなく、その人らしさ（肉声感）をもって伝える装置が望まれていた。

そこでわれわれは、専属録音スタッフの元で本人の声を記録・録音し、熟練専門家の手作業により声データベース（音声コーパスと呼ぶ）を構築し、そのデータベースを使った個人声質の音声合成（テキスト文字列を音声に変換する）ソフト Pol luxstar（以下 PXT と呼ぶ）を提供している。PXT の最大の特徴は、失う前のその人自身の声質で音声メッセージを伝えることである。

### (B) 開発する支援機器の概要

専属録音スタッフの立会い元での声の記録・作業は、録音スタッフコストがかかることに加えて、どうしても集中収録作業となり、本人の自然なペースでの収録や都度の追加収録などは実際上困難であった。このことが PXT を利用する際の制約条件となっていた。

そこで、専属録音スタッフの立会という条件をはずして、補助者、支援者を含むユーザ自身で録音が可能なる「ユーザ録音」ツールを開発する。

また熟練専門家の手作業による音声コーパスの構築では、1 人の声を処理するのに数十日の作業を要し、PTX の提供コストを押し上げる原因となっていた。

そこで、ユーザ自身で録音すると同時に、PC ソフトで自動的に音声コーパス化作業を行う「音声コーパスセルフ構築ツール」を導入し、実運用での性能評価と課題や改良点を洗い出す。

### (C) 実証試験の目的

発声収録上、健常者と異なる注意や工夫が必要と思われる ALS 等の神経変性疾患等の発話障害者に対して、

- (1) 「ユーザ録音」の作業負荷と必要な性能（会話補助機能に支障がない音質）が得られるかを確認する。また作業上の注意点をリストアップする。
- (2) 「音声コーパスセルフ構築ツール」が必要な性能が得られるかを確認する。

### (D) 研究の概要（この研究によって実証すべき機器の性能、研究デザイン、研究方法の概要）

「ユーザ録音」や「音声コーパスセルフ構築ツール」による音声コーパス構築が、専属スタッフ録音や専門家による通常構築と比較して、その会話補助支援機能や作業負荷の差異を調べる。

会話補助機能では、コミュニケーションの品質、コミュニケーション範囲や目的等について、アンケート等で PXT 使用前との差異や変化を調べる。作業負荷は作業時間と主観的な作業負荷をアンケート等で調査する。

### (E) インフォームド・コンセントの取得方法、個人情報保護の方法の概要

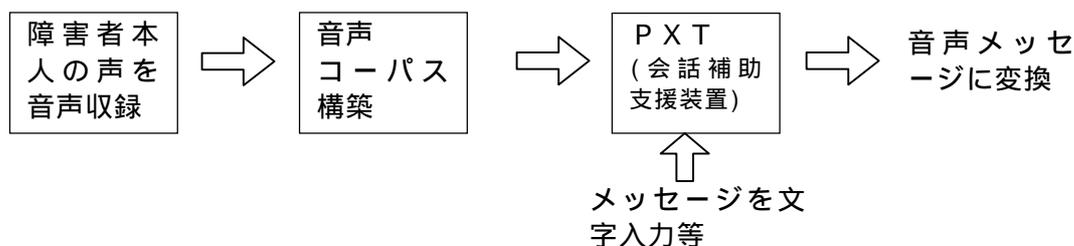
- インフォームド・コンセントは、文書ならびに口頭で説明、同意書による
- 個人情報保護は、実験中は連結可能匿名化、実験終了後またはモニター機器の返却後は連結不可能匿名化にして 5 年間保存し、その後は廃棄する。
- 同意書は鍵付きキャビネットに保存、データはハードディスクに暗号化して保存。
- PC 情報管理、個人情報管理、連結表の管理者は実験担当責任者（本田武司）とする。

### 3. 機器の詳細

#### (A) 支援機能について

発声障害者用の会話補助支援装置として、さまざまな音声を発する機器が提供されているが、人格が見えないメッセージ音声ではなく、声に含まれるその人らしさ(肉声感)をもって伝える装置が望まれていた。

そこでわれわれは、専属録音スタッフの元で本人の声を記録・録音し、熟練専門家の手作業により声のデータベース(音声コーパスと呼ぶ)を構築し、そのデータベースを使った個人声質の音声合成(テキスト文字列を音声に変換する)ソフト Polluxstar(以下 PXT と呼ぶ)を提供している。PXT の最大の特徴は、失う前のその人自身の声質で音声メッセージを伝えることである。



P X T (会話補助支援装置) の作成の流れ

#### (B) 課題

専属録音スタッフの立会い元での声の記録・作業は、録音スタッフコストがかかることに加えて、どうしても集中収録作業となり、本人の自然なペースで収録や都度の追加収録などは実際上困難であった。このことが PXT を利用する際の制約条件となっていた。

また、熟練専門家の手作業による音声コーパスの構築は1人の声を処理するのに数十日の作業を要し、PTX の提供コストを押し上げる原因となっていた。

#### (C) 実験目的

PXT の録音作業による制約を軽減し、また PTX の提供コストを下げる改良を進めて、利用者を拡大させることにある。

#### (D) 先行研究など

すでに ALS などの神経変性疾患により徐々に発話機能が失われていく障害者を対象に PXT の提供を行い、個別事例の観察やヒアリングにより、PXT の会話補助機能は高く評価を頂いている[1]。

[1]神経変性疾患患者への自分の声ソフトウェア Polluxstar 適用事例、HIS2010、本田 武司 他

#### (E) 開発する支援機器(プロセス改良含む)の概要

- 専属録音スタッフの立会という条件をはずして、補助者、支援者を含むにユーザ自身で録音が可能なる「ユーザ録音」ツールを開発する。
- ユーザ自身で録音すると同時に、PC ベースで自動的に音声コーパス化作業を行う「音声コーパスセルフ構築ツール」を導入し、実運用での性能評価と課題や改良点を洗い出す。
- 「ユーザ録音」ツール「音声コーパスセルフ構築ツール」の健常者による先行試行と課題等の洗い出しを実施する。
  - ◆ 発話負荷 「ユーザ録音」ツール作業と「音声コーパスセルフ構築ツール」作業を健常者で実施し、発話負荷に関するアンケートやヒアリングで問題点の洗い出しを実施する。
  - ◆ 音質評価等 両ツールで作成された健常者音声コーパスの PXT 合成音に対して、無記名、健常者による視聴実験等により音質評価等を実施し、従来の「通常録音」コーパスと比較し、問題点の洗い出しを実施する。

## 4. 研究方法

### (A) 研究デザイン

- ・対象者が従前の使用している従来型の VOCA ( 会話補助支援装置 : 例 HeartyLadder+smartTalk ) と本人録音型 VOCA ( PXT ) を比べて、その品質と適用範囲に関しての比較
- ・音声コーパス作成手順の作業負荷測定 ( 従来録音、ユーザ録音、音声コーパスセルフ構築ツールの 3 種類 )
- ・3 種類の作成手順の違いによる、合成音声の音質の違いを比較評価 ( 本人による評価は伴わない )

### (B) 仮説ならびにエンドポイント

- ・**仮説 1 「どの 3 種類の複数の音声コーパス作成手順においても、日常コミュニケーションに必要な音質が得られる」**  
エンドポイントは、各音声コーパス作成手順毎の合成音声の画一的な音質比較評価法の結果評価とする。
- ・**仮説 2 「どの 3 種類の複数の音声コーパス作成手順においても、コミュニケーション品質が向上する」**  
エンドポイントは、障害者本人の声を熟知している近親者等のコミュニケーション品質指標のアンケート結果評価とする。
- ・**仮説 3 「どの 3 種類の複数の音声コーパス作成手順においても、コミュニケーション範囲・目的が広がり、QOL が向上する」**  
エンドポイントは、近親者等のコミュニケーション範囲指標のアンケート結果評価とする。
- ・**仮説 4 「どの 3 種類の複数の音声コーパス作成手順においても、本人の声を記録し保存することで、失声不安が軽減される ( モニュメント価値 ) 」**  
エンドポイントは、本人、近親者等のモニュメント価値のアンケート結果評価とする。
- ・**仮説 5 「どの 3 種類の複数の音声コーパス作成手順においても、本人の声の録音作業の負担は許容範囲内である」**  
エンドポイントは、本人、近親者等の録音作業のアンケート結果評価とする。

### (C) 仮説の立証のために記録する事実

録音作業に関わる記録事項と作業負荷

- (1) 従来録音 ( 訪問録音型 )  
被験者は発話するのみである ( 指定された文・単語を読み上げる )。被験者の PC 等操作は不要である。  
収録が必要な発話量は、健常者の場合 5 時間 ~ 6 時間の収録時間を要する。
  - (2) ユーザ録音 ( ご本人録音型 )  
もし近親者などが録音補助作業 ( マイク設定、IC レコーダなどの録音開始、終了操作 ) をしていただけるのであれば、訪問録音型と同じに操作は不要であり発話するのみである。ご自分で録音補助作業をするのであれば、IC レコーダなどの録音開始・終了操作も必要である。  
発話量は従来録音と同じく健常者の場合 5 時間 ~ 6 時間の収録時間を要する。
  - (3) 音声コーパスセルフ構築ツール ( 自動録音型 )  
もし近親者などが録音補助作業 ( PC 自動録音ソフトの操作 ) をしていただけるのであれば、訪問録音型と同じに操作は不要である。ご自分で録音補助作業をするのであれば、自動録音ソフトの操作 ( PC 操作 ) が必要である。  
発話量は基本的に従来録音と同じく健常者の場合 5 時間 ~ 6 時間の収録時間を要する。
- ・本人の声の録音作業の負担などの検査項目は、録音完了時に個別に質問紙調査とヒアリングで記録する。( 全 1 回実施、実施時間 3 0 分 )

利用操作に関わる記録事項と作業負荷

- (1) 従来型の音声合成音を出力する VOCA と PXT による合成音を出力する VOCA ( 標準 VOCA

- として HL:Harty Ladder を使用) の操作説明と簡単な訓練を実施する。( 60 分程度)
- (2) 従来型 VOCA と PXT 使用の VOCA を定められた期間(1ヶ月程度)試用していただく。  
試用期間中は、近親者や協力者をお願いし、利用記録をとって頂く。
- (3) 試用期間終了後、こちらで指定した双方の比較操作実験(30分程度)をしていただき、その後コミュニケーション満足度、コミュニケーション範囲・目的、あるいは失声不安に関するアンケート調査とヒアリングを実施する。(全1回実施、実施時間60分程度)コミュニケーション品質、コミュニケーション範囲・目的、失声不安の検査項目は、本コミュニケーション支援機器を1ヶ月使用後に、個別に質問紙調査とヒアリングで記録する。(全1回実施、実施時間30分)

音声、映像等を記録する場合の頻度と所要時間

モニター期間中、1、2回実施予定、所要時間は各10分程度

(D) 記録した事実からエンドポイントを導出する手続き

アンケート項目毎の平均値の有意差検定による。

ただし、録音作業量の評価項目等は除く。

(E) 国外の施設における実証試験の実施予定の有無

(有りとした場合の相手国における研究倫理に関する対策)

なし

## 5. 対象者

(A) 対象者の選定基準(選択基準、除外基準、禁忌)

選択基準:

- ・疾患・障害は進行性神経性疾患(主に筋萎縮性側索硬化症:ALS)とする。(必須条件)
- ・発話発声が可能な方(必須条件)

除外基準:

- ・アンケート回答が不可能である、未成年

禁忌:

- ・発話発声が不可能

(B) 予定人数(年齢層、性別、疾患・障害別等)

従来録音2名、ユーザ録音1名、自動録音ツール1名(可能ならば)

いずれも、年齢は成人、性別不問、疾患・障害は進行性神経性疾患による発話障害患者とする

注)

- ・従来録音とは、訪問録音型のことで、専任録音スタッフが立会い、録音機器の設定や録音操作などの全て行います。録音された音声から手作業で音声コーパスを構築します。
- ・ユーザ録音とは、専任録音スタッフの立会という条件をはずして、補助者、支援者を含むユーザ自身で録音をしていただくことです。録音された音声は、手作業で音声コーパスを構築します。
- ・自動録音ツールとは、PC操作により、PCにつながれたマイクなどから、自動的にPC上に音声を記録します。記録が終了後に、PC上のソフトウェアの自動作業で音声コーパスを構築します。

(C) 対象者への特別の配慮

なし

(D) 対象者の募集・選定手続き ( 機縁募集 公募 )

機縁募集先、機縁先との関係 ( 機縁先への依頼状等を添付すること )

都立神経病院の実験協力者を通じた機縁募集とする。

対象者候補との接触方法。主治医、担当セラピスト、担当ソーシャルワーカー等と研究者の関係、役割分担。

協力される先生方に先の基準で該当者をスクリーニングしていただく。  
スクリーニング後の候補者に対して、実験責任者が実験の説明と同意手続きを行う。

施設の入所者、病院等の入院患者を対象とする場合、威圧、強制などを伴わないための特別の配慮

特になし

(E) 対象者の被る危害と便益 ( リスクとベネフィットの可能性 )

この研究に必然的に伴う侵襲

なし

予見される身体的・心理的・社会的不利益、危害とそれへの対象者保護対策  
声を出す作業、あるいは会話補助支援装置の操作に限られるので、心理的、社会的危害はない。さらに身体的危害も極めて少ないが、録音作業に伴う発声行為からくる疲労等からくる体調不良は想定しておく必要がある。

録音作業はいつでも中断、再開が可能であり、疲労兆候が現れたら直ちに作業を中断し、本人、親族、介護者、補助者と相談連絡しながら、必要な処置をとる ( 休憩をとる、体勢を工夫する、帰宅させる、医療機関へ連絡など )

危害・有害事象のために対象者を除外あるいは中断するための判断基準

対象者除外、中断は実験担当責任者が判断する。実験全体の中止判断等は研究代表者が行う。実験開始前に比べて、発話行為が困難・疲労の度合いが明らかに増大、会話補助支援装置の操作が困難、アンケートの回答が不可能といった状況になったと本人、親族、介護支援者から意見が上がり、実験担当者も同様な判断した場合は実験を一時中断し、実験担当責任者が対象者を除外するなどの判断を行う。

この研究のために健康被害が発生した時の措置

なし

この研究によって対象者が直接受ける便益

コミュニケーション手段の品質が向上し、コミュニケーションコストが下がり、本人の生活活動の範囲も広がる。

この研究の結果社会が受ける便益

- ・「ユーザ録音」や「音声コーパスセルフ構築ツール」を実運用評価し、評価終了後1年以内の販売を目指す。その結果、本支援装置の利用者負担を1/3以下に削減させる。年300人利用を想定している。
- ・将来的には低コストで音声ライブラリや音声バンクを構築し、この音声バンク活用により、全ての福祉機器に人間味のある音声応答メッセージが使えるようになる。

対象者に提供する謝金、謝礼

録音作業終了時 ( アンケートを含む ) : 1回あたり1000円程度

モニター利用終了時（アンケートを含む）：1回あたり1000円程度

(F) インフォームド・コンセントの手続き

説明の方法

個別に文書を添えて口頭にて説明する

説明実施者（氏名、所属）

川島 紳、株式会社ウォンツ ソリューション事業部（研究代表者）  
本田 武司、株式会社ウォンツ ソリューション事業部（実験担当責任者）  
三木 敬、合同会社テクノマインズ（研究分担者）

インフォームド・コンセントの具体的手順

- ・実験協力者に先の選定基準で該当者のスクリーニングを依頼する。
- ・スクリーニング後の候補者に対して、実験責任者が実験紹介のアポイントをとる
- ・アポイントがとれた対象者に対して、順番に個別面談し、実験の説明と同意手続きを行う。

限能力者のための特別の配慮

該当なし

(G) 対象者の個人情報保護・収集したデータのための安全管理

匿名化の措置

匿名化しない。 連結可能匿名化する。 連結不能匿名化する。

連結可能匿名化のときの連結表の管理者：本田武司（実験担当責任者）

ある時点で連結不能匿名化する場合：

連結不能匿名化の時期：実験終了5年経過後またはモニター終了後

連結不能匿名化担当者名：本田武司（実験担当責任者）

匿名化しない場合および連結可能匿名化する場合、その理由

モニター評価中の機器操作や手順の問い合わせ対応、アンケート調査訪問の連絡などが必要なため

匿名化する場合の匿名化担当者（氏名・所属）

本田武司、株式会社ウォンツ ソリューション事業部（実験担当責任者）

研究期間中あるいはモニター機器使用中の個人情報、連結表、データ・試料等の保管  
保管責任者：本田武司（実験担当責任者）

保管場所：株式会社ウォンツ内 ファイル管理サーバ

保管方法：パスワード付PCファイルで保管、紙媒体は鍵付キャビネットで保管

研究終了後の個人情報、連結表、データ・試料等の保管法、

保管期間：実験終了5年経過後またはモニター機器終了後、連結不能匿名化し継続保管もしくは廃棄

保管責任者：本田武司（実験担当責任者）

保管場所：株式会社ウォンツ内 ファイル管理サーバ

保管方法：パスワード付PCファイルで保管、紙媒体は鍵付キャビネットで保管

データ等の処分・破棄の方法：紙媒体はシュレッダー廃棄、PCファイルは削除

同意書の保管

保管責任者：本田武司（実験担当責任者）

保管場所：株式会社ウォンツ内

保管方法：鍵付キャビネットで保管

破棄の時期：実験終了5年経過後またはモニター機器終了後

破棄の方法：シュレッダー廃棄

## 6. 起こりうる利益相反とその管理

### (A) 経済的な利益相反

なし

### (B) その他の利益相反（研究者が対象者となる利益相反、学生や従業員を対象者としたときの利益相反、患者と担当医療職との利益相反等の利益相反があれば、それを指摘し、その管理策について記載すること）

なし

## 7. 特記事項

なし

## 8. 研究者の素養

氏名	現職	最終学歴・専攻	この分野の研究歴、臨床経験等
川島 紳	株式会社ウォンツ ソリューション事業 部 取締役		
本田 武司	株式会社ウォンツ ソリューション事業 部 グループリーダー		
三木 敬	合同会社テクノマ インズ 代表社員		

## 9. 文献リスト

[1] 神経変性疾患患者への自分の声ソフトウェア Polluxstar 適用事例、HIS2010、本田 武司他

5. 代用音声ライブラリ開発に関する実証試験研究計画書

(ヒトを対象とする支援機器の実証試験)

実証試験研究計画書

作成日 2010年 11月 18日

修正日 2010年 12月 27日

1. 実証試験研究課題

研究課題	言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発(健常者音声コーパス開発分)
研究期間	平成 平成 23年 1月 21日から平成23年 3月 31日まで

(A) 研究組織

	氏名	所属・役職・職種	分担項目	連絡先
研究代表者	川島 紳	株式会社ウォンツソリューション事業部 取締役	実験全般管理	名古屋市中区栄三丁目 10-22 東朋ビル 6F Tel:052-251-0571 [Redacted]
実験担当責任者	本田武司	株式会社ウォンツソリューション事業部	実験実施管理	同上 [Redacted]
分担研究者	三木 敬	合同会社テクノマインズ 代表社員	実験技術管理と評価作業	川口市南前川 2-21-10-304 TEL 090-6471-4144 [Redacted]
助言を担当する医師	川田明広 長尾雅裕			

(B) 共同研究実施機関・組織・施設・研究実施場所

機関・組織名	実施組織・場所	実施内容	倫理審査状況
該当なし			

(C) 研究協力機関

機関・組織名	実施組織・場所	実施内容	倫理審査状況
*****		メッセージ音声試聴とアンケート記入(無記名)	

(D) 研究資金 厚生労働省 障害者自立支援機器等研究開発プロジェクト

## 2. 研究の概要

### (A) 支援機器の目的・目標

発声障害者用の会話補助支援装置として、予め録音した本人の声データベース（音声コーパス）を使ってその人らしさ（肉声感）をもつ音声を合成するソフト Polluxstar（以下 PXT と呼ぶ）を提供している。

音声コーパスの作り方は、さまざまな発音を万遍なく網羅するような文章を用意し、本人がその文章を数時間にわたって喋った声を録音する。

しかし、既に声を失った方には、録音したデータがないため、その適用は不可能であった。

### (B) 開発する支援機器の概要

そこで様々な声質の健常者の音声コーパスを録音し、それを集めた音声コーパスの集合体（以下、声質ライブラリと呼ぶ）を用意し、その中からふさわしい声質を選択して PXT 等に利用する手法を検討している。

### (C) 実証試験の目的

想定利用者へのアンケート調査を通じて、声質ライブラリの利用可能性検討（検証）を行う。

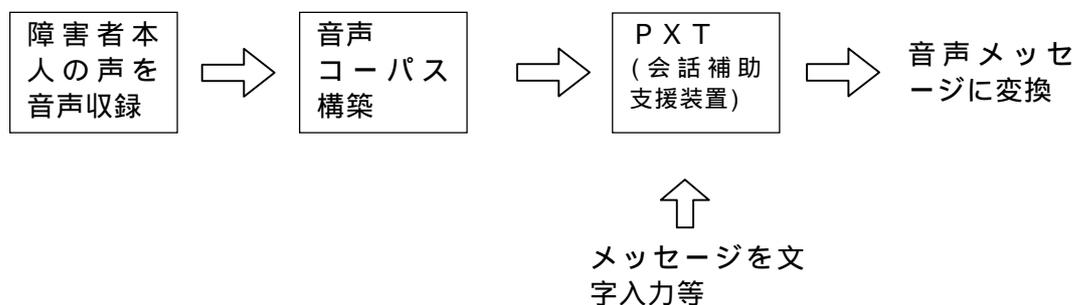
### (D) インフォームド・コンセントの取得方法、個人情報保護の方法の概要

- インフォームド・コンセントは、文書ならびに口頭で説明する。  
個人情報は収集しない。

### 3. 機器の詳細

#### (A) 支援機能について

発声障害者用の会話補助支援装置として、予め録音した本人の声データ（音声コーパス）を使ってその人らしさ（肉声感）をもつ音声を合成するソフト Polluxstar（以下 PXT と呼ぶ）を提供している。



PXT(会話補助支援装置)の作成の流れ

#### (B) 課題

しかし、既に声を失った方には、録音したデータがないため、その人らしさを再現することは不可能であった。

#### (C) 実験目的

そこで年齢別ならびに性別毎に数名分のコーパスからなる声質ライブラリを用意し、その中から最もふさわしい声質として選択した1名の声だけを使って音声を合成する手法を検証する。

#### (D) 先行研究など

特になし。

#### (E) 開発する支援機器(プロセス改良含む)の概要

失声者(会話補助支援装置利用者)の声質分布をカバーするような声質ライブラリを収集し、そのライブラリ等を使った音声メッセージの試聴アンケート調査実験を通じて、声質ライブラリ構築の進め方や有効性等を確認するパイロット評価を行う。

声質ライブラリ収集にあたっては、今回暫定的に声質の属性として発話者の年代と性別属性でセグメント化し、各セグメント毎に数名の声を収録する。

声の必要数ならびに収集基準(属性)の検証等は、今回の検討結果を踏まえた今後の検討で明確にしていく。

具体的な発話者は健常者プロナレータ等から選定した。発話者への同意は、失声者向けの支援機器に限定し、編集・加工した音声データとして利用するという条件で、著作権、その他の知的財産権、利用権ならびに使用権を弊社に帰属することの同意書によっている。(有償ライセンス)

#### 4. 研究方法

##### (A) 研究デザイン

従来型 VOCA( 会話補助支援装置 )で使われている一般健常者利用向けの声メッセージ( ホームページ読み上げ、カーナビ、構内連絡放送等 ) や、失声者の声質分布を考慮した声質ライブラリを利用した声メッセージを、会話補助支援装置利用候補者に対して、各声質毎に 2 ~ 3 文 ( 数秒 ) を試聴して頂き、自分の音声としてふさわしい声の選択回答式無記名質問紙調査を行う。

##### (B) 仮説ならびにエンドポイント

- ・ 仮説

「他人の音声コーパスから合成した音声を受け入れられる」( 傍証：複数利用者からの意見あり )

発声障害者の声質分布を踏まえた声質ライブラリから選択した声質でニーズを十分にまかなえる。

- ・ エンドポイント：

声質ライブラリ等から利用した声メッセージを自分の音声としてふさわしい声と判断する人がいる

##### (C) 仮説の立証のために記録する事実

男女別に複数名集ってもらい、音声を聞いてもらい、質問紙に回答する。

( 1 ) 声質( 男性：9種類、女性：5種類を予定 ) 毎の複数音メッセージ( 2 ~ 3 文( 各文数秒 ) ) を順に試聴し、

( 2 ) それぞれの声質が自分の会話メッセージとしてふさわしいどうか、あるいは違和感があるかないかをアンケート用紙に記入する

なお、複数音メッセージは、声質当り 2 回程度提示する。一回の実施時間は 40 分程度

##### (D) 記録した事実からエンドポイントを導出する手続き

アンケートの単純集計ならびに、属性( 個人、性別、年齢 ) とのクロス集計、並びに個人毎にふさわしい声質が存在する割合

##### (E) 国外の施設における実証試験の実施予定の有無

( 有りとした場合の相手国における研究倫理に関する対策 )

なし

## 5. 対象者

### (A) 対象者の選定基準（選択基準、除外基準、禁忌）

選択基準：

- (1) 声帯摘出者、あるいは舌摘出などで、発声機構、調音機構に異常がある方
- (2) 聴覚に異常がなく、TVの声などが通常に聞き取れる方
- (3) アンケート用紙に回答（筆記）できる方
- (4) 会話補助支援装置を使ってみたい、あるいは興味がある人

除外基準：

- ・アンケート回答が不可能である、未成年

禁忌：

- ・聴覚異常

### (B) 予定人数（年齢層、性別、疾患・障害別等）

- ・アンケート調査2回を予定。男性15名、女性5名程度、計20名を予定。

### (C) 対象者への特別の配慮

なし

### (D) 対象者の募集・選定手続き（機縁募集      公募）

機縁募集先、機縁先との関係（機縁先への依頼状等を添付すること）

日喉連（銀鈴会：声帯摘出者NPO等）を通じた機縁募集とする。

対象者候補との接触方法。主治医、担当セラピスト、担当ソーシャルワーカー等と研究者の関係、役割分担。

協力される関係者に該当者を募集。

アンケート調査実施前に、実験責任者が実験の説明と同意手続きを行う。

施設の入所者、病院等の入院患者を対象とする場合、威圧、強制などを伴わないための特別の配慮

特になし

### (E) 対象者の被る危害と便益（リスクとベネフィットの可能性）

この研究に必然的に伴う侵襲

なし

予見される身体的・心理的・社会的不利益、危害とそれへの対象者保護対策  
なし。

危害・有害事象のために対象者を除外あるいは中断するための判断基準  
危害・有害事象なし。この研究のために健康被害が発生した時の措置としては、  
万一の体調不調の場合は、医療機関へ連絡するなど、必要な手段をとる。

この研究によって対象者が直接受ける便益  
なし。

この研究の結果社会が受ける便益  
既に声を失った方にも声質ライブラリの中から、自分の声質としてふさわしい、ある  
いは違和感のない声質を選択して代替音声とする発話支援装置を提供できるよう  
なる。

(F) 対象者に提供する謝金、謝礼

モニター利用終了時（アンケートを含む）：1回あたり500円程度

(G) インフォームド・コンセントの手続き

説明の方法

個別に文書を添えて口頭にて説明する

説明実施者（氏名、所属）

川島 紳、株式会社ウォンツ ソリューション事業部（研究代表者）

本田 武司、株式会社ウォンツ ソリューション事業部（実験担当責任者）

三木 敬、合同会社テクノマインズ（研究分担者）

インフォームド・コンセントの具体的手順

- ・ 実験協力者に先の選定基準で該当者募集を依頼する。  
（依頼状添付：健常者コーパス被験者選定依頼 r2.doc）
- ・ あらかじめ実験協力者と打ち合わせて指定した日時に被験者候補者に集まってもらい、説明文書を配布するとともに実験の手順、個人情報保護などの説明を行う。併せて、無記名の質問表調査なので、提出していただくことで同意いただいたものとさせていただきますことを説明する。

制限能力者のための特別の配慮

該当なし

(H) 対象者の個人情報保護・収集したデータのための安全管理

個人情報は収集しない。

研究期間中あるいはモニター機器使用中の個人情報、連結表、データ・試料等の保管  
保管責任者：本田武司（実験担当責任者）

保管場所：株式会社ウォンツ内 ファイル管理サーバ  
保管方法：パスワード付PCファイルで保管、紙媒体は鍵付キャビネットで保管

研究終了後のデータ・試料等の保管法、  
保管期間：実験終了5年経過後は廃棄  
保管責任者：本田武司（実験担当責任者）  
保管場所：株式会社ウォンツ内 ファイル管理サーバ  
保管方法：パスワード付PCファイルで保管、紙媒体は鍵付キャビネットで保管  
データ等の処分・破棄の方法：紙媒体はシュレッダー廃棄、PCファイルは削除

## 6. 起こりうる利益相反とその管理

(A) 経済的な利益相反

なし

(B) その他の利益相反（研究者が対象者となる利益相反、学生や従業員を対象者としたときの利益相反、患者と担当医療職との利益相反等の利益相反があれば、それを指摘し、その管理策について記載すること）

なし

## 7. 特記事項

発話者とは、製作請負業者を介しての業務契約を結んでおりますので被験者には該当しません。

業務契約書「制作業務契約書.pdf」を添付

## 8. 研究者の素養

氏名	現職	最終学歴・専攻	この分野の研究歴、臨床経験等
川島 紳	株式会社ウォンツ ソリューション事業 部 取締役		
本田 武司	株式会社ウォンツ ソリューション事業 部 グループリーダー		
三木 敬	合同会社テクノマ インズ 代表社員		

## 対象者として支援機器実証試験に参加するための説明文書

### この実証試験研究について

1. 験課題： 言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発

#### 2. 実証試験実施者

実証試験研究代表者： 川島 紳 株式会社ウォンツ ソリューション事業部  
実証担当責任者： 本田 武司 株式会社ウォンツ ソリューション事業部  
分担研究者： 三木 敬 合同会社テクノマインズ  
総括責任者： 川島 紳 株式会社ウォンツ ソリューション事業部

#### 3. 研究の場所と期間

この実証試験は、XXXXXXXXXXにおいて全期間が平成22年9月1日（または「実証試験の実施が承認された日」）から平成23年3月31日までにまたがる予定です。ただし、対象者の方に参加していただく期間は1～3ヶ月間（または「XXXX年XX月X日からXXXX年XX月X日まで」）です。

#### 4. 実証試験の背景と目的

さまざまな工夫を施した会話補助支援装置や器具が開発され、いろいろな場面で使われるようになりました。しかし、その応答音や応答音声は、明瞭ではありますが、機械的で親しみに欠ける印象がありました。

そこで、あらかじめ記録・録音しておいた自分の声とそれら会話補助支援装置を組み合わせ、その人の口調やくせなどの個性を伝えることで、親しみやすい会話のやり取り（コミュニケーション）が可能になります。

その結果、色々な人とのコミュニケーションが個性豊かなものとなり、人との触れあいや交流がより楽しくなります。さらには活用頂いた方からは、積極的な人との交わりのなかで、一瞬ではありますが永遠に心に残る貴重な瞬間に出会えたという素晴らしい報告も頂いております。

この素晴らしさをより多くの皆さんに使っていただけるよう、多くの方のモニター評価実験を通じて、良い製品開発につなげたいと思います。

## 5. 実証試験の方法

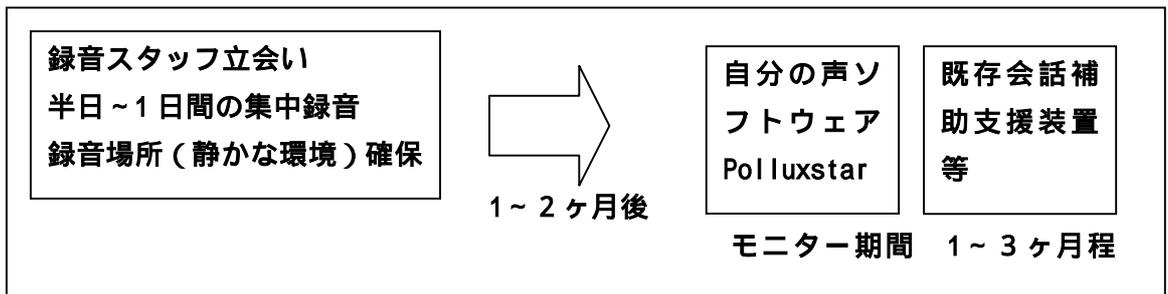
・自分の声を録音・記録します。録音作業終了後、作業時間や負担などをアンケート紙に記入、また聞き取り調査もいたします。操作性の確認のために、必要に応じて操作場面等の動画を記録することがあります。  
その後、新たな声の音声合成ソフトが完成して利用可能になったとき(1~2ヵ月後)から、以下の作業をお願いいたします。

- (1) 最初に、既存の会話補助装置(VOCA)とPXT利用型VOCAの操作説明と簡単な訓練を実施いたします。(60分程度)
- (2) その後、従来型VOCAとPXT使用のVOCAを定められた期間(1ヶ月程度)試用していただきます。試用期間中は、近親者や協力者の方に、適宜利用の記録を付けて下さるようお願いいたします。
- (3) 試用期間終了後、こちらで指定した双方の比較操作実験(30分程度)をしていただき、その後にコミュニケーション満足度、コミュニケーション範囲・目的、あるいは失声不安に関するアンケート調査と聞き取り調査を実施いたします。(全1回実施、実施時間60分程度)

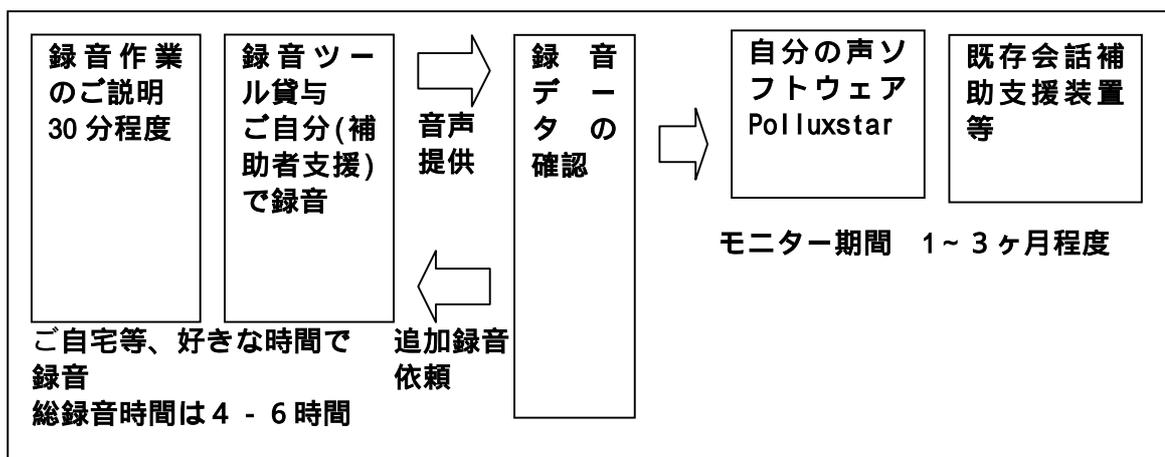
具体的なモニター作業の日程やスケジュールは個別に相談させていただきます。

モニター作業には2つタイプがあり、こちらから指定したどちらか一方のモニター作業をお願いいたします。

- (1) 訪問録音型 (平成22年9月以降開始予定)



- (2) ご本人録音型・自動録音型(平成22年12月以降開始予定)



## 6. 研究に関する資料の開示について

あなたのご希望があれば、他の対象者の個人情報保護や研究の独創性の確保に支障がない範囲で、この研究の研究計画および研究方法についての資料を開示いたします。また、この研究に関するご質問がありましたらいつでも担当者にお尋ね下さい。

## 7. 研究への参加の任意性

この研究への参加は任意です。あなたの自由な意思が尊重されます。研究に参加しないことによって、不利益な対応を受けることはありません。

一旦参加に同意した場合でも、いつでも不利益を受けることなく同意を撤回することができます。そのためには、この説明書の最終ページに添付してある同意撤回書に署名捺印して、この説明の最後に明示してあるこの研究に関する問い合わせ先まで撤回をお申し出下さい。

その場合、それまでに提供していただいたデータは廃棄され、それ以降はそれらの情報が研究のために用いられることもありません。ただし、同意を撤回したときすでに研究成果が論文などで公表されていた場合等、すでに公表済みの成果は取り消せないこともあります。

## 8. この研究への参加をお願いする理由、代諾手続きの場合の参加が不可欠である理由

本製品のような本人の音声を収録する会話補助支援装置の適用が可能な方(安定した発声ができること、その際に身体負荷が大きいこと)、さらにその利用に興味をもたれる多くの方(ご自分でPC操作が出来、アンケート調査に協力いただける方など)にモニター評価実験に参加いただいて、良い製品開発につなげたいと思います。

また会話補助装置の利用者だけでなく、モニター作業に協力頂いた近親者、補助者、支援者の方とも、改めてその同意を頂きながら、随時意見交換、アンケート調査をお願いするなどして、積極的に共同改良実験などに取り組み、その成果は御協力頂いた皆様方と共同で学会等へ発表につなげていければと考えております。

## 9. この研究への参加を中断する場合

実験中に、発話行為が困難になったり、収録作業にかかる疲労が大きいと思われる場合、あるいは会話補助支援装置の操作活用が困難になった場合、さらにアンケートの回答が不可能といった状況になった場合には、実験を中断、中止いたします。

## 10. この実証試験への参加に伴う危害の可能性、有害事象発生の際の補償について

モニター評価はいつでも中断、再開が可能ですので、ご本人のペースで進められます。ご本人ならびにその補助者・支援者の肉体的・精神的負担が大きくなるよう、事前・実験中実験後を問わず、モニター作業の進め方に工夫配慮するとともに、本人、その補助者、支援者、専門家、担当医師等が負担が大きいと判断した場合は、ただちにモニター評価実験を中断・中止して、休みを取ったり、親族や介護者への連絡や相談、必要に応じて医療機関へ連絡するなど、必要な手段をとります。

実験の継続が困難であると本人あるいはその補助者、支援者、専門家、担当医師等が判断した場合には、直ちに実験を中止いたします。

ただし、万一音声の収録時の発声などの行為が体調悪化の要因となった場合においては、

それに関する補償は出来かねますので、その点はご了解いただきますようお願いいたします。

#### 11. 研究により期待される便益

多くの方のモニター評価実験を通じて、より録音の作業負荷が少ないものに改良するなどして、本製品の適用となる方（発話障害を持つ方）全般のコミュニケーション支援機能開発に役立てていきます。また録音作業コストカットなど実現することで、本製品の提供価格を下げ、本支援装置の利用者負担の軽減を目指します。

またモニター期間中にも、本装置のコミュニケーション支援機能を活用することで、親しみやすい会話のやり取り（コミュニケーション）が期待されます。

#### 12. 個人情報の取り扱い

あなたのデータや個人情報は、この研究を遂行し、その後検証するために必要な範囲においてのみ利用いたします。この研究のために研究グループの外部にデータを提供する必要があった場合は改めて承諾をお願いします。また関連した実験のための対象者として依頼する場合があります。その際には改めて同意をお願いします。

個人情報の取り扱いに関しては、実験終了後5年間は、氏名や住所等の個人を特定できる情報と実験データを別々に管理し、実験データからは個人を特定できないものとしたします。5年経過後は、個人を特定できるデータは廃棄し、完全に匿名化された実験データのみを保管いたします。

#### 13. 研究終了後の対応・研究成果の公表

学会発表・報告等に使われる場合は、個人データが明らかとならない統計的データの形で使用し、それ以外の使い方場合は、本人・関係者へ事前に同意確認をいたします。

また、あなたの個人情報は、氏名・住所を実験データと別々に管理するなど通常の使い方では個人情報が確認できない方法で厳重に管理した上で保存し、研究終了5年経過後に個人情報が外部に漏れないよう完全に匿名化した上で音声データのみ保管します。

#### 14. 研究のための費用

負担はありません

#### 15. 研究に伴う対象者謝金等

録音作業終了時（アンケートを含む）：1回あたり1000円程度

モニター利用終了時（アンケートを含む）：1回あたり1000円程度

#### 16. 知的財産権の帰属

実証試験研究代表者 株式会社ウォンツに帰属します。

#### 17. 音声データの取り扱い

今回収録した音声は発話者個人の肖像権として保護され、発話者のコミュニケーションを支

援する機器・ソフトウェア・サービスと組み合わせでのみ使用されます。

## 問い合わせ先・苦情等の連絡先

この研究に関する問い合わせ先

**研究代表者 株式会社ウォンツ 川島 紳**  
**名古屋市中区栄三丁目10番22号**  
**TEL:052-251-0571**

この研究に関する苦情等の連絡先

-----

以上の内容をよくお読みになってご理解いただき、この研究に参加することに同意される場合は、別紙の「研究への参加についての同意書」に署名し、日付を記入して担当者にお渡し下さい。

## 同意撤回書

研究代表者: 株式会社ウォンツ ソリューション事業部 取締役  
川島 紳 殿

私は、「言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発」の研究に対象者として参加することに同意し、同意書に署名しましたが、その同意を撤回することを担当研究者

..... 氏

に伝え、同意書は返却され、受領いたしました。ここに同意撤回書を提出します。

平成 年 月 日

(対象者本人による同意書を提出された場合は以下に署名、捺印をお願いします。)

対象者氏名(自署).....  
生年月日  
住所・連絡先

(代諾者による同意書を提出された場合は以下に署名、捺印をお願いします。)

代諾者(家族等)氏名(自署).....  
(注)家族等とは、後見人、保佐人、親権者、父母、配偶者、成人の子又は兄弟姉妹等をいう。  
対象者(患者)との続柄  
生年月日  
住所・連絡先

本研究に関する同意撤回書を受領したことを証します。

担当研究者 川島 紳 ..... 印  
所 属 株式会社ウォンツ ソリューション事業部  
職 取締役

## 同意書

実証試験代表者

株式会社ウォンツ ソリューション事業部 取締役 川島 紳 殿

試験課題： 言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発

私は、研究計画名「言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発」に関する以下の事項について説明を受けました。理解した項目については自分で の中に ✓印を入れて示しました。

研究を実施する研究者（説明文書 項目 2）

研究の場所と期間（説明文書 項目 3）

研究の背景と目的（説明文書 項目 4）

研究の方法（説明文書 項目 5）

研究に関する資料の開示について（説明文書 項目 6）

研究への参加が任意であること（研究への参加は任意であり、参加しないことで不利益な対応を受けないこと。また、いつでも同意を撤回でき、撤回しても何ら不利益を受けないこと。）（説明文書 項目 7）

私がこの研究への参加を依頼された理由（説明文書 項目 8）

この調査への参加を中断する場合（説明文書 項目 9）

この試験への参加に伴う危害の可能性について（説明文書 項目 10）

研究により期待される便益について（説明文書 項目 11）

個人情報の取り扱い（被験者のプライバシーの保護に最大限配慮すること）（説明文書 項目 12）

研究終了後の対応・研究成果の公表について（説明文書 項目 13）

研究のための費用（説明文書 項目 14）

研究の参加に伴う被験者謝金等（説明文書 項目 15）

知的財産権の帰属（説明文書 項目 16）

音声データの取り扱い（説明文書 項目 17）

問い合わせ先・苦情等の連絡先

なお、この実証試験において撮影・記録された私の映像（静止画、動画）・音声の公開につきましては以下の の中に ✓印を入れて示しました。（説明文書 項目 5）

公開に同意しない

研究者を対象とする学術目的に限り、下記条件の下に公開に同意する。

顔部分など個人の同定可能な画像も含んで良い

顔部分や眼部などを消去・ぼかすなど個人の同定不可能な状態に限る

その他（特別な希望があれば、以下にご記入ください）

これらの事項について確認したうえで、被験者として研究に参加することに同意します。

平成.....年.....月.....日

被験者署名.....

本研究に関する説明を行い、自由意思による同意が得られたことを確認します。

説明担当者（所属・職名・氏名）.....

## 同意書（代諾者用）

研究代表者:

株式会社ウォンツ ソリューション事業部 取締役 川島 紳 殿

研究課題名: 言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発

私は、研究計画名「言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発」に関する以下の事項について説明を受けました。理解した項目については自分で✓印を入れて示しました。

研究を実施する研究者（説明文書 項目2）

研究の場所と期間（説明文書 項目3）

研究の背景と目的（説明文書 項目4）

研究の方法（説明文書 項目5）

研究に関する資料の開示について（説明文書 項目6）

研究への参加が任意であること（研究への参加は任意であり、参加しないことで不利益な対応を受けないこと。また、いつでも同意を撤回でき、撤回しても何ら不利益を受けないこと。）（説明文書 項目7）

この研究への参加を依頼された理由、この研究の重要性と、研究対象者が参加することが不可欠である理由（説明文書 項目8）

この調査への参加を中断する場合（説明文書 項目9）

この試験への参加に伴う危害の可能性について（説明文書 項目10）

研究により期待される便益について（説明文書 項目11）

個人情報の取り扱い（被験者のプライバシーの保護に最大限配慮すること）（説明文書 項目12）

研究終了後の対応・研究成果の公表について（説明文書 項目13）

研究のための費用（説明文書 項目14）

研究の参加に伴う被験者謝金等（説明文書 項目15）

知的財産権の帰属（説明文書 項目16）

音声データの取り扱い（説明文書 項目17）

問い合わせ先・苦情等の連絡先

なお、この実証試験において撮影・記録された私の映像（静止画、動画）・音声の公開につきましては以下の の中に✓印を入れて示しました。（説明文書 項目5）

公開に同意しない

研究者を対象とする学術的に限り、下記条件の下に公開に同意する。

顔部分など個人の同定可能な画像も含んで良い

顔部分や眼部などを消去・ぼかすなど個人の同定不可能な状態に限る

その他（特別な希望があれば、以下にご記入ください）

これらの事項について確認したうえで、（.....）がこの研究に参加することに同意します。

平成.....年.....月.....日

家族等署名 .....

（注：家族等とは、後見人、保佐人、親権者、父母、配偶者、成人の子又は兄弟姉妹）

住所・連絡先（電話）〒 .....

被験者名・被験者との続柄・被験者生年月日

.....年.....月.....日

本研究に関する説明を行い、自由意思による同意が得られたことを確認します。

説明担当者署名（所属・職名・氏名） .....

## 会話補助支援装置に関するアンケート調査の説明文書

倫理審査コメント対応分 2010年12月10日

倫理審査コメント対応分 2010年12月22日

### この実証試験研究について

1. 試験課題：言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発

#### 2. 実証試験実施者

実証試験研究代表者：川島 紳 株式会社ウォンツ ソリューション事業部

実験担当責任者：本田 武司 株式会社ウォンツ ソリューション事業部

分担研究者：三木 敬 合同会社テクノマインズ

総括責任者：川島 紳 株式会社ウォンツ ソリューション事業部

#### 3. 実証試験の場所と期間(時間)

この実証試験は、XXXXXXXXXXにおいて全期間が平成23年1月21日(または「実証試験の実施が承認された日」)から平成23年3月31日までにかかると見込まれています。ただし、対象者各人のアンケート調査にかかる時間は40分程度です。

#### 4. 実証試験の背景と目的

さまざまな工夫を施した会話補助支援装置や器具が開発され、いろいろな場面で使われるようになりまし。しかしながら、その声質は必ずしも本人の声の代用として会話補助支援装置向けに開発されたものばかりではなく、一般利用向けの音声メッセージ(ホームページ読み上げ、カーナビ、構内連絡放送等)の声質に近いものまでも含まれていました。

そこで今回、会話補助支援装置としてどんな声質がふさわしいかを、さまざまな方にお聞きして、望まれる声とはどんな声質か、また品揃えはどれほど必要なかを調査したいと思っております。皆様のご協力をお願いいたします。

例えば衣服なら、各人の体系や好みに応じて、サイズ、形、色、模様、素材など、いろいろ品揃えが必要なのは分かっていると思いますが、声の品揃えとなると、全くといって良いほど調査されていないのが現状なのです。

そこで大変恐縮ですが、少しでも皆様のお時間をいただいて、貴重なデータを取らせていただきたく、アンケート記入にご協力ください。

## 5. 実証試験の方法

様々な声質の音声メッセージをそれぞれ数秒ずつお聞きいただき、それぞれの声質が自分の声メッセージとしてふさわしいどうか、あるいは違和感があるかないかを、アンケート用紙にご記入いただきます。

声質は男性の場合9種類、女性の場合5種類です。全部で40分程度です。

また、アンケートは無記名で、個人情報収集いたしません。

## 6. 実証試験に関する資料の開示について

あなたのご希望があれば、研究の独創性の確保に支障がない範囲で、この研究の研究計画および研究方法についての資料を開示いたします。また、この研究に関するご質問がありましたらいつでも担当者にお尋ね下さい。

## 7. 実証試験への参加の任意性

このアンケート調査への参加は任意です。あなたの自由な意思が尊重されます。参加しないことによって、不利益な対応を受けることはありません。

この試験は無記名の質問紙調査です。質問紙にご記入のうえ提出いただくことによって、この試験への参加に同意いただいたものとさせていただきます。

## 8. この実証試験への参加をお願いする理由、代諾手続きの場合の参加が不可欠である理由

本製品のような会話補助支援装置に正当な評価が出来る方、さらにその利用に興味をもたれる多くの方にアンケート調査に参加いただき、良い製品開発につなげたいと思います。

## 9. この実証試験への参加を中断する場合

気分が悪くなったり、気が乗らなかつたりした場合はいつでも中断が可能です。その場合、記入いただいた質問紙は破棄します

なお、いったんご提出いただいた質問紙は、無記名ですので撤回することはできません。この点ご了承ください

## 10. この実証試験への参加に伴う危害の可能性、有害事象発生の際の補償について

アンケート調査は音声を聞いて質問紙に記入するだけで、かつ、途中でいつでも中断が可能ですから、危害、有害事象の発生はほとんどありません。

万一、体調不調の場合は、医療機関へ連絡するなど、必要な手段をとります。

## 11. 実証試験により期待される便益

多くの方のアンケート調査を通じて、本製品の適用となる方（発話障害を持つ方）全般のコミュニケーション支援機能開発に役立てていきます。

## 12. 個人情報の取り扱い

本調査では個人情報は収集いたしません。

## 13. 実証試験終了後の対応・研究成果の公表

学会発表・報告等に使われる場合は、統計的データの形で使用します。

## 14. 実証試験のための費用

実証試験のための費用は、厚生労働省 社会・援護局障害保健福祉部の「障害者自立支援機器等開発促進事業助成金」よりまかなわれます。被験者の方の負担はありません。

## 15. 実証試験に伴う対象者謝金等

モニター利用終了時（アンケートを含む）：1回あたり500円程度

## 16. 知的財産権の帰属

実証試験研究代表者 株式会社ウォンツに帰属します。

## 問い合わせ先・苦情等の連絡先

この実証試験に関する問い合わせ先

研究代表者 株式会社ウォンツ 川島 紳  
名古屋市中区栄三丁目10番22号  
TEL:052-251-0571

この研究に関する苦情等の連絡先

以上の内容をよくお読みになってご理解いただき、このアンケート調査に参加される場合は参加希望日時を担当者までお知らせ下さい。

## 9. 代用音声に関する Web アンケート

### 9.1 Web アンケート 説明、回答者情報取得ページ

<厚生労働省>

平成22年度障害者自立支援機器等開発促進事業

言語障害者向けに  
人間味のある声で会話補助する  
支援機器の低価格化開発に関するアンケート調査  
(会話支援装置に関する無記名式アンケート)

【この実証試験研究について】

この実証試験研究について

1. 試験テーマ  
言語障害者向けに人間味のある声で会話補助する支援機器の低価格化に向けた開発
2. 実証試験実施者  
実証試験研究代表者:川島 紳 株式会社ウオッツソリューション事業部  
実験担当責任者 :本田 武司 株式会社ウオッツソリューション事業部  
分担研究者 :三木 敬 合同会社テクノマインズ  
総括責任者 :川島 紳 株式会社ウオッツソリューション事業部

このアンケートは、20種類程度の音声をお聞きいただき、それぞれの音声についての感想をお答えいただくものです。そのため、パソコンにはスピーカーやイヤホンなどの音の出る機器を接続し、そこから流れ出る音声をお聞きください。所要時間は、およそ4分程度となります。どうぞご協力お願いします。

【問1】

アンケートにご回答いただくご本人様にお尋ねします。

性別	<input type="radio"/> 男性	<input type="radio"/> 女性	
年齢	<input type="radio"/> 30代以下	<input type="radio"/> 40代	<input type="radio"/> 50代
	<input type="radio"/> 60代	<input type="radio"/> 70代	<input type="radio"/> 80代以上
発話障害の有無	<input type="radio"/> あり	<input type="radio"/> なし	

以下の質問は、発話障害をお持ちの方への質問です。

発話障害の程度	<input type="radio"/> 重度(喉頭全摘出、障害等級3級、失声状態) <input type="radio"/> 中度(障害等級4級、日常会話に支障あり) <input type="radio"/> 軽度(日常会話に多少支障あり) <input type="radio"/> 軽微(日常会話にほとんど支障なし) <input type="radio"/> 不明		
発話障害の期間	<input type="radio"/> 1年未満	<input type="radio"/> 1年～3年	<input type="radio"/> 3年～5年
	<input type="radio"/> 5年～10年	<input type="radio"/> 10年以上	
日常会話で使用する コミュニケーション手段 (複数可)	<input type="checkbox"/> 食道発声 <input type="checkbox"/> 会話支援装置 <input type="checkbox"/> その他	<input type="checkbox"/> シャント発声 <input type="checkbox"/> 筆談	<input type="checkbox"/> 人工声帯(人工喉頭) (複数回答可)

次へ進む

## 9.2 Web アンケート 代替音声に関する感想取得ページ

### 【問2】

再生ボタンをクリックし、サンプル音声をお聞きいただいた後、ご質問にお答えください。

問①:この声質(ご自身の発話音声としてふさわしい声質ですか。)(択一)

問②:ふさわしいとしたら、使っても良いと思われる場面をお選びください。(複数選択)

途中いつでも聞き直すことができます。

最後に、ページ最下部の『アンケートに回答する』ボタンをクリックしてください。

サンプル音声 1 / 19

音声	問①	問②
	<input type="radio"/> ふさわしい <input type="radio"/> まあふさわしい <input type="radio"/> どちらともいえない <input type="radio"/> あまりふさわしくない <input type="radio"/> ふさわしくない	<input type="checkbox"/> 家族、近親者、知人、介護者等との日常会話 <input type="checkbox"/> 知人を囲んだ話し合いの場 <input type="checkbox"/> 知らない人もいる話し合いの場 <input type="checkbox"/> 知らない人もいる集会のスピーチ等 <input type="checkbox"/> 知らない人もいる改まった集会のスピーチ等(結婚式、講演会、議会 等) <input type="checkbox"/> 電話

サンプル音声 2 / 19

音声	問①	問②
	<input type="radio"/> ふさわしい <input type="radio"/> まあふさわしい <input type="radio"/> どちらともいえない <input type="radio"/> あまりふさわしくない <input type="radio"/> ふさわしくない	<input type="checkbox"/> 家族、近親者、知人、介護者等との日常会話 <input type="checkbox"/> 知人を囲んだ話し合いの場 <input type="checkbox"/> 知らない人もいる話し合いの場 <input type="checkbox"/> 知らない人もいる集会のスピーチ等 <input type="checkbox"/> 知らない人もいる改まった集会のスピーチ等(結婚式、講演会、議会 等) <input type="checkbox"/> 電話

サンプル音声 3 / 19

音声	問①	問②
	<input type="radio"/> ふさわしい <input type="radio"/> まあふさわしい <input type="radio"/> どちらともいえない	<input type="checkbox"/> 家族、近親者、知人、介護者等との日常会話 <input type="checkbox"/> 知人を囲んだ話し合いの場 <input type="checkbox"/> 知らない人もいる話し合いの場 <input type="checkbox"/> 知らない人もいる集会のスピーチ等

音声	問①	問②
	<input type="radio"/> ふさわしい <input type="radio"/> まあふさわしい <input type="radio"/> どちらともいえない <input type="radio"/> あまりふさわしくない <input type="radio"/> ふさわしくない	<input type="checkbox"/> 家族、近親者、知人、介護者等との日常会話 <input type="checkbox"/> 知人を囲んだ話し合いの場 <input type="checkbox"/> 知らない人もいる話し合いの場 <input type="checkbox"/> 知らない人もいる集会のスピーチ等 <input type="checkbox"/> 知らない人もいる改まった集会のスピーチ等(結婚式、講演会、議会 等) <input type="checkbox"/> 電話

サンプル音声 19 / 19

音声	問①	問②
	<input type="radio"/> ふさわしい <input type="radio"/> まあふさわしい <input type="radio"/> どちらともいえない <input type="radio"/> あまりふさわしくない <input type="radio"/> ふさわしくない	<input type="checkbox"/> 家族、近親者、知人、介護者等との日常会話 <input type="checkbox"/> 知人を囲んだ話し合いの場 <input type="checkbox"/> 知らない人もいる話し合いの場 <input type="checkbox"/> 知らない人もいる集会のスピーチ等 <input type="checkbox"/> 知らない人もいる改まった集会のスピーチ等(結婚式、講演会、議会 等) <input type="checkbox"/> 電話

### 【問3】

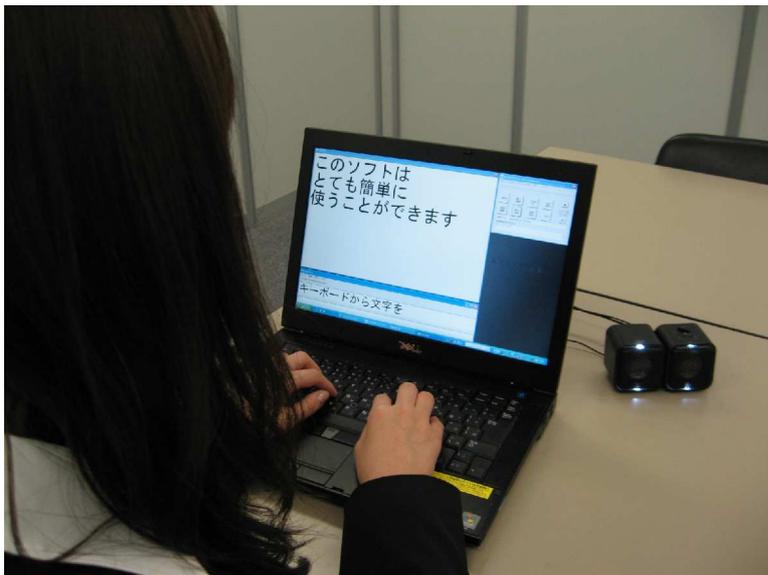
会話支援装置の音声メッセージの声質に関して、ご意見がありましたら、ご自由にご記入下さい。

アンケートに回答する

### 9 . 3 Web アンケート クロージングページ



## 10. 支援機器、音声収録ツール 使用の様子

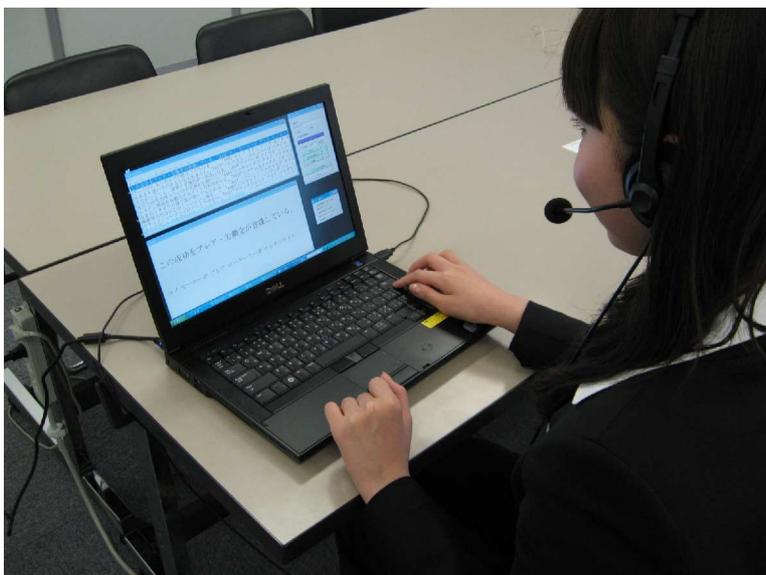


自分の声ソフト（ボルックスター）の利用形態

PCのキーボードから発声したい言葉を入力すれば、スピーカから、あなたの音声が出ます。

自動収録ライブラリ化ソフトの利用形態

ソフトのガイダンスに従って、PCに接続されたマイクに向かって発声し録音をおこなう。所定量の録音完了後に音声データベースの構築をおこなうことができる。



ユーザ録音の様子

録音機器セット、録音作業ガイド、読み上げ原稿に基づき、自分のペースで音声収録をおこないます。

