産総研 ともに挑む。つぎを創る。
国立大学法人 筑波技術大学
UTokyo
JST
PRESS RELEASE
研究成果発表

2026 年 9 月 25 日

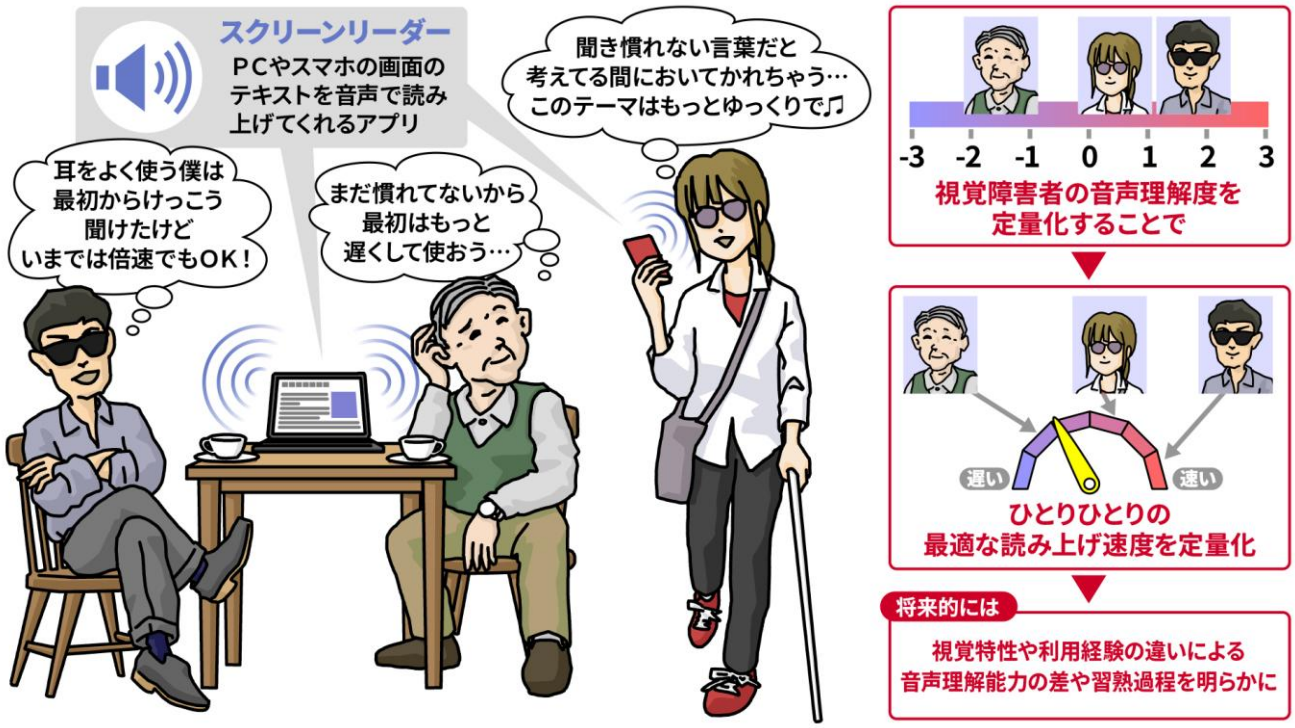
国立研究開発法人産業技術総合研究所／国立大学法人筑波技術大学／国立大学法人東京大学／国立研究開発法人科学技術振興機構（JST）

合成音声読み上げ時の音声情報理解力を定量化

視覚障害者一人ひとりに「ちょうどいい」読み上げ速度を提供するために

ポイント

- 速度調整可能な合成音声読み上げソフト「スクリーンリーダー」使用時に、文章の難易度と利用者の理解能力を分離し、同じ尺度で評価する手法を開発
- 開発した手法を用いて、個々の利用者が効率よく文字情報を理解できる読み上げ速度を定量化
- 視覚障害者個人の視覚特性の違いによる音声理解能力の差や習熟過程を分析するための基盤データに



概 要

国立研究開発法人産業技術総合研究所（以下「産総研」という）人間社会拡張研究部門 三浦 貴大 上級主任研究員と、国立大学法人筑波技術大学 坂尻 正次 教授ら、国立大学法人東京大学先端科学技術研究センター 藪 謙一郎 特任研究員らは共同で、視覚障害者の音声理解能力を定量化する手法を開発しました。

スクリーンリーダーは、パソコンやスマートフォンの文字情報を速度調整可能な合成音声で読み上げるためのソフトウェアです。それは、視覚障害者の就学や就労、社会参加を支える情報アクセシビリティの基盤となっています。スクリーンリーダーを使用して音声を理解できる速度には個人差がありますが、その要因や音声理解の習熟過程はこれまで明らかになっていませんでした。

本研究では項目反応理論を応用し、読み上げられた文章が理解できなかった時、それは文章が難解だったからか、読み上げ速度が速すぎたからかを区別し、文章の難易度と利用者の理解能力を同じ指標で定量化する手法を開発し

ました。開発された評価手法を用いたところ、全盲の利用者は、弱視（ロービジョン）の利用者と比べ、より速い読み上げ速度でも高い理解を保つなどの科学的根拠を持つ知見が得られました。本成果は、視覚障害の状況や利用経験による個人差を分析し、音声理解の習熟過程や能力向上の可能性を評価するための基盤データとなり、「速すぎて聞き取れない」「遅すぎてもどかしい」という不便を減らし、視覚障害のある人がより快適に、より速く情報にアクセスできる環境作りにつながると期待されます。

なお、本研究成果は Interspeech 2026（2026 年 9 月 27 日～10 月 1 日、オーストラリア シドニー）で発表予定です。

下線部は【用語解説】参照

研究の社会的背景

スクリーンリーダーは、パソコンやスマートフォンの画面に表示された文字や操作内容を、速度調整が可能な合成音声で伝え、主に目の見えない・見えにくい人（視覚障害者）のデジタル機器の利用を支援するソフトウェアです。視覚障害者がデジタル情報にアクセスし就学や就労をはじめ社会生活を送るための、重要な基盤となっています。

日本では、視覚障害のある身体障害者手帳の所持者は約 27 万人と推計されています（厚生労働省「令和 4 年 生活のしづらさなどに関する調査」）。手帳を所持していない人や、加齢によって見えにくくなった人まで含めると、実際に「見えにくさ」を抱える人はさらに多いと考えられ、誰にとっても身近な課題です。

スクリーンリーダー利用者の多くは、慣れるにつれて読み上げ速度を徐々に速くします。熟練した利用者の中には、標準の何倍もの速度で聞き取れる人もいます。しかし、情報を最も効率よく理解できる読み上げ速度には個人差があり、その違いには利用経験だけでなく、全盲か弱視か、先天盲か後天盲かといった視覚特性や、読み上げる内容など、さまざまな要因が関係すると考えられています。一方で、こうした個人差がどのように生じるのか、利用経験に伴って音声理解能力がどのように習熟していくのか、また利用者が現在どの程度の速度まで理解できるのかを定量的に評価する方法については、十分に検討されていませんでした。

研究の経緯

産総研は、視覚障害者の音声理解能力の個人差や習熟過程を定量的に評価し、一人ひとりの特性に応じた学習の最適化を実現するための研究を進めています。今回、学力テストや心理検査で広く用いられている項目反応理論を応用し、スクリーンリーダー利用時の音声理解能力を定量的に評価する手法の開発に取り組みました。

なお、本研究は JST 創発的研究支援事業（JPMJFR234Y）、JSPS 科研費（JP23K25693、JP23K17582、JP24K22314 および JP26K02945）の助成を受けたものです。

研究の内容

これまで、スクリーンリーダー利用において「この利用者は、どの速度まで読み上げを理解できているのか」を客観的な数値として示す方法が確立されておらず、読み上げを理解できない要因が文章側と読者側のどちらにあるのかを区別するための知見がありませんでした。従来研究では、特定の速度で読み上げられた音声に対する正答率を比較する方法が主に用いられてきました。これでは、文章が理解されなかった場合に、その原因が「読み上げ速度が

速すぎたから」なのか、「文章そのものが難しかったから」なのか、あるいは「その利用者の音声理解能力がまだ十分に習熟していないから」にあるのかを区別できませんでした。

そこで本研究では、スクリーンリーダーの読み上げ速度に対する「理解のしやすさ」を定量的に測るために、項目反応理論を応用した評価手法を構築しました。項目反応理論は、学力テストや資格試験で用いられる統計理論で、「問題の難しさ」と「回答する人の能力」を同じ測定基準（ものさし）の上で別々に推定できる点に特徴があります。本研究はこの考え方を、音声読み上げの理解しやすさの評価に応用しました。

具体的には、音韻的にバランスの取れた日本語の文章群（研究用に整備された「ITA コーパス」[1,2]）を用い、これらを毎分 150 語から 500 語まで 5 段階の速度で読み上げた音声を作成しました。実測した発話速度は約 6〜27 モーラ/秒で、通常の会話（およそ 7〜8 モーラ/秒）の約 0.8〜3.5 倍に相当します。視覚障害のあるスクリーンリーダー利用者 11 名を対象に、これらの音声聞き取り実験を行い、内容理解度の回答とともに、「聞き取りやすさ」「わかりやすさ」を評価してもらいました。

得られた回答データから、「どの音声がどれだけ理解しにくいか（項目の難易度）」と「どの利用者がどれだけ理解できるか（個人の能力）」を、共通の測定基準上で分離して推定しました。両者が同じものさしの上に並ぶため、「この利用者の能力は、この速度の難易度を上回っているか」という形で、一人ひとりに適した速度を判断できるようになります。

さらに本研究では、ベイズ統計を取り入れた項目反応理論を用いました。従来の項目反応理論は安定した推定に多数の参加者を必要としますが、ベイズ統計を用いることで、参加者が限られる場合でも安定した推定が可能となる他、推定結果の確からしさ（不確実性の幅）も同時に示すことが可能です。視覚障害者のように対象者が限られる研究において、この点は重要な利点となります。

構築した評価手法を用いて分析した結果、まず読み上げ速度が速くなるほど理解の難易度が系統的に高まることが確認されました。内容の理解度は、毎分 150 語ではほぼ全問正答に近い水準でしたが、毎分 300 語、毎分 500 語と速くなるにつれて低下しました（図 1 (a)）。特に毎分 250〜300 語のあたりが理解の境界となり、これを超えると理解度が大きく下がる傾向が見られました。

一方で、最適な速度には大きな個人差があり、すべての利用者に当てはまる唯一の正解は存在しないことも明らかになりました。利用者ごとの能力の違いは、速度による難易度の変化を上回るほど大きく、一律の設定では利用者の能力を十分に活かせないことが定量的に示されました。

また、全盲の利用者は、弱視（ロービジョン）の利用者と比べ、より速い読み上げ速度でも高い理解を保つ傾向が確認されました。実際、普段設定している読み上げ速度も、全盲の利用者では標準の約 2.3 倍、弱視の利用者では約 1.4 倍と明確な差がありました（図 1 (b)）。分析の結果、この違いは「速い読み上げに慣れているから」という利用経験だけでは説明できず、視覚障害の状態そのものと利用経験とが、それぞれ独立して音声理解能力に関係していることが示唆されました。

これらの結果は、読み上げ速度に対する理解のしやすさが、文章の難易度と利用者の音声理解能力を共通の尺度上でそれぞれ独立に推定できること、そして視覚障害の状態や利用経験に応じた個別最適化に科学的な根拠を与えられることを示しています。

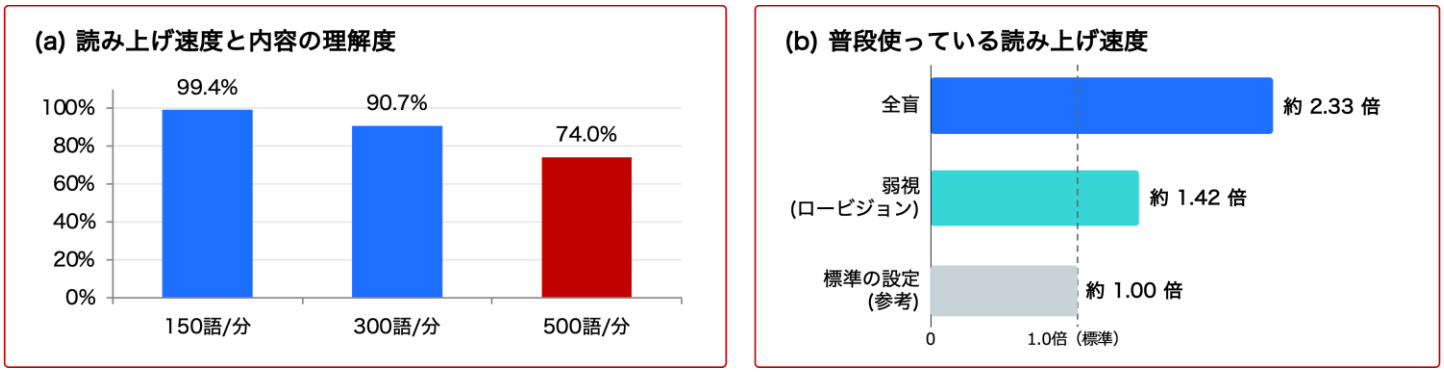


図 1 主な分析結果

（a）読み上げ速度と内容の理解度。速度は音声合成時の設定値（毎分 150～500 語、実測値は約 6～27 モーラ/秒）。
（b）普段使っている読み上げ速度の平均値（標準比）

今後の予定

今後は、より多くの視覚障害者を対象とした大規模な調査を実施し、本手法で得られる推定の精度を高めるとともに、視覚障害の状態や利用経験と音声理解能力との関係の詳細を明らかにします。特に、スクリーンリーダーを使い始めて間もない利用者を対象に含めることで、利用経験に伴って音声理解能力がどのように習熟していくのかを継続的に評価することを目指します。

あわせて、利用者の音声理解能力を短時間で測定し、一人ひとりに適した読み上げ速度をその場で提示できる評価ツールの開発に取り組みます。将来的には、視覚障害者の情報アクセス環境の改善に加え、速い音声を聞き取る能力（速聴）の訓練や評価など、幅広い応用につなげることを目指します。

学会情報

学会名：Interspeech 2026
発表タイトル：Profiling Speech Rate Abilities of Visually Impaired Screen Reader Users by Bayesian Item Response Theory
著者：Takahiro Miura, Masatsugu Sakajiri, Masaki Matsuo, Keiichi Yasu, Junji Onishi, Ken-ichiro Yabu, Tohru Ifukube

研究者情報

産総研
人間社会拡張研究部門 三浦 貴大 上級主任研究員
筑波技術大学
坂尻 正次 教授、松尾 政輝 助教、安 啓一 准教授、大西 淳児 教授
東京大学先端科学技術研究センター
藪 謙一郎 特任研究員、伊福部 達 研究顧問

参考情報

[1] 小口純矢，金井郁也，小田恭央，齊藤剛史，森勢将雅：ITA コーパス：パブリックドメインの音素バランス文からなる日本語テキストコーパスの構築と基礎評価，情報処理学会研究報告，vol. 2021-MUS-131, no. 31, pp. 1-4, 2021.
[2] ITA コーパスの文章リスト公開用リポジトリ, https://github.com/mmorise/ita-corpus

用語解説

スクリーンリーダー

パソコンやスマートフォンなどのコンピューターの画面に表示された文字を合成音声で読み上げるソフトウェア。目の見えない・見えにくい人は、画面を見てマウスを操作する代わりに、この読み上げを聞きながら、キーボードやフリック操作でコンピューターを使う。特に全盲の人たちにとっては情報を得るための必須のツール。Windows や macOS、iOS、Android には、こうした読み上げ機能が標準で搭載されており、日本でも Narrator、PC-Talker、NVDA、VoiceOver などが広く使われている。読み上げ速度の調整の他、声質の変更などもでき、現在ではさまざまな言語に対応している。

情報アクセシビリティ

年齢や障害の有無にかかわらず、誰もが必要な情報やサービスに支障なくアクセスし、利用できること。たとえば、Web サイトや電子文書が音声読み上げに対応していることや、文字の拡大・色の調整ができることなどが含まれる。視覚障害のある人にとっては、スクリーンリーダーで読み上げられる形で情報が提供されているかどうかが、仕事や学び、日常生活のしやすさを大きく左右する。

項目反応理論（項目応答理論ともいう、Item Response Theory：IRT）

テストや調査の結果を分析するための統計理論。「問題の難しさ」と「回答する人の能力」を同じ測定基準の上で別々に推定できる点が特徴。学力テストや資格試験、心理検査などで広く使われている。単純な正答率では「難しい問題だったのか」「その人の能力が高い（低い）のか」が区別できないが、項目反応理論を使うと両者を切り分けて捉えられる。近年では、ベイズ統計を取り入れた項目反応理論も発展しており、参加者が少ない場合でも比較的安定した推定ができ、かつ推定の確からしさ（不確実性）も同時に示せる利点がある。本研究では、このベイズ版の項目反応理論を用いた。

モーラ

日本語の音の長さを表すリズムの単位。一般的に、日本語話者が話す時の「ひらがな 1 文字分」の長さが 1 モーラ（1 拍）に該当。多くの場合、ひらがな 1 文字が 1 モーラに対応し、小さな「っ」、鼻音の「ん」、長音の「ー」も、それぞれ 1 モーラとして数える。一方、「きゃ」「しゅ」のように小さな「ゃ」「ゅ」「ょ」を伴う音は、2 文字で 1 モーラとなる。

ベイズ統計

分析前に得られている知識や仮定と、実際に観測されたデータを組み合わせて、知りたい量の確からしさを推定する統計的な推論方法。分析前の知識や仮定を確率分布で表したものを「事前分布」、観測データが各仮定のもとで得られるもっともらしさを「尤度（ゆうど）」、これらを組み合わせて更新した確率分布を「事後分布」と呼ぶ。推定値だけでなく、その不確実性も併せて示すことができる。本研究では、この方法を取り入れた項目反応理論により、文章の難易度と利用者の音声理解能力をそれぞれ推定した。

機関情報

国立研究開発法人産業技術総合研究所

https://www.aist.go.jp/

ブランディング・広報部　報道室　hodo-ml@aist.go.jp

国立大学法人筑波技術大学

https://www.tsukuba-tech.ac.jp/

広報室　kouhou@ad.tsukuba-tech.ac.jp

国立大学法人東京大学先端科学技術研究センター

https://www.rcast.u-tokyo.ac.jp/

広報広聴・情報支援室　press@rcast.u-tokyo.ac.jp

国立研究開発法人科学技術振興機構（JST）

https://www.jst.go.jp/

広報課　jstkoho@jst.go.jp

（JST 事業に関する問い合わせ）

国立研究開発法人科学技術振興機構（JST）

創発的研究推進部

永井 諭子

〒102-0074　東京都千代田区九段南 3 丁目 3-6 麹町ビル 3 階

03-5214-7276　　souhatsu-inquiry@jst.go.jp