

世界の「言葉の壁」をなくす 多言語音声翻訳技術



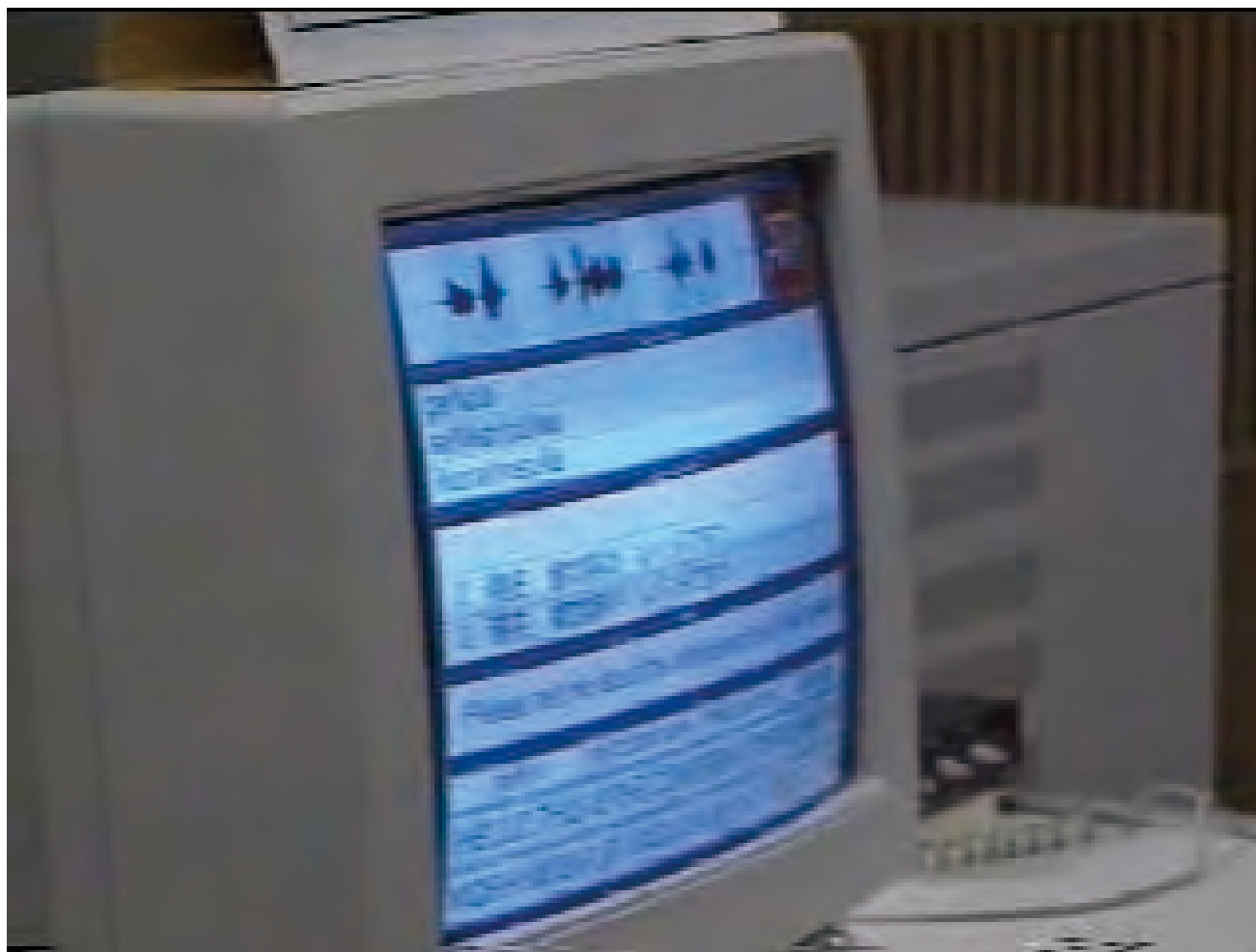
(独)情報通信研究機構
ユニバーサルコミュニケーション研究所
先進的音声翻訳研究開発推進センター
内元清貴

音声翻訳研究の歴史



自動音声翻訳システムの第一号

ATRの第一期プロジェクト(1986-1992)の成果

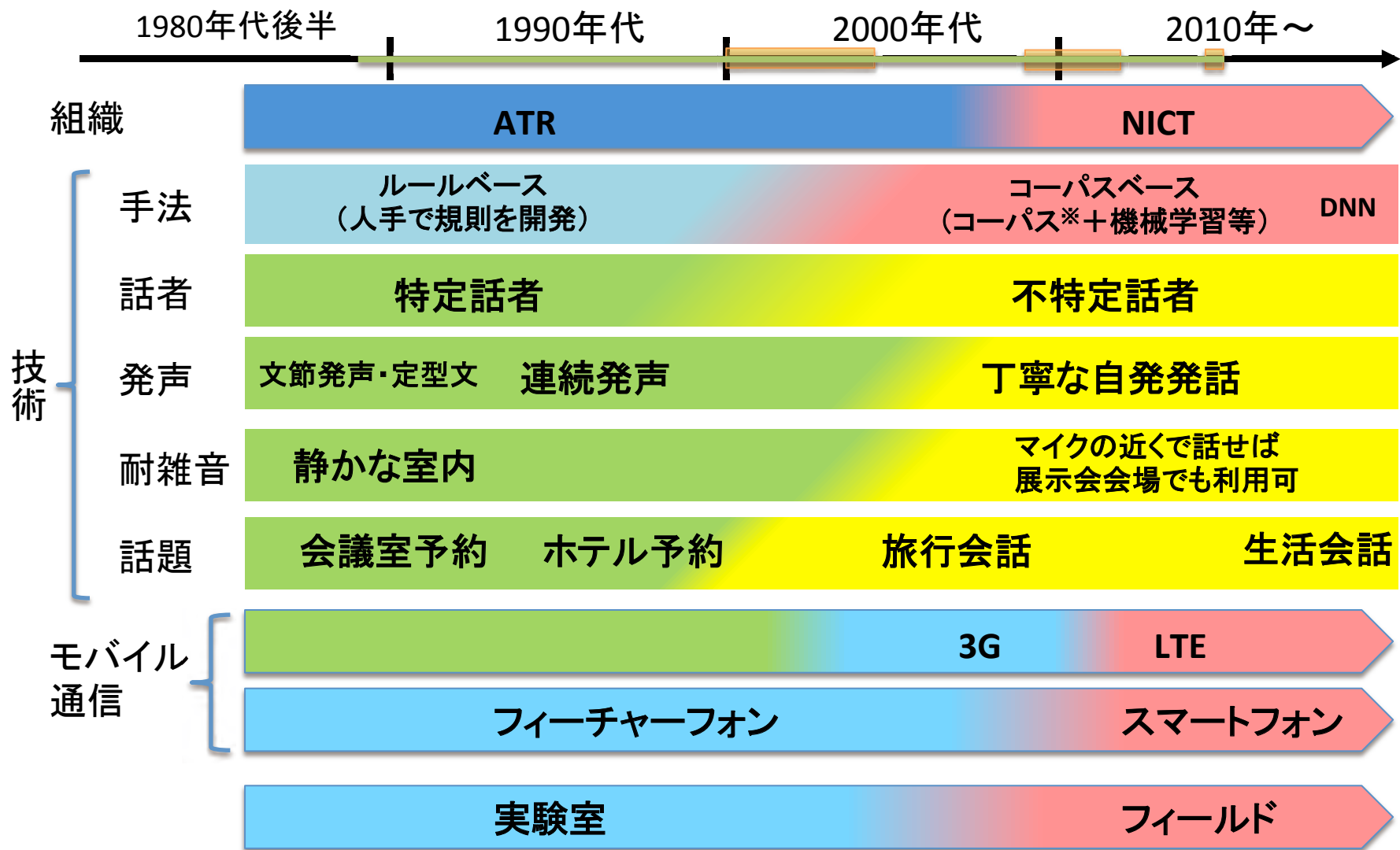


音声翻訳の計算機・通信の進化





ATR/NICTにおける 多言語音声翻訳の研究の歴史



※ 電子化された自然言語の文章から成る巨大な音声・テキストデータ及びそれに言語的な情報が付加されたもの。



近年の多言語音声翻訳

(VoiceTra 2010年)



VoiceTra開発への道のり

社会還元加速プロジェクト (実用化が加速)

2007 2008 2009 2010 2011 2012 2013



総合科学技術
会議デモ

社会還元加速
プロジェクトスタート

2009年度補正予算
「総務省 地域の観光
振興に貢献する自動
音声翻訳技術の実証
実験」にて全国5カ所で
実証実験実施

得られた大規模実験
データによって音声認
識・翻訳精度が向上

北海道: 札幌/富良野/知床
12/28~2/22, 98,830発話
ドコモ/NEC/ソフトバンク/北海道/NICT

中部: 金沢/能登/高山/伊勢/鳥羽
1/5~2/29, 37,692発話
JTB中部/NTT/NICT

関西: 大阪/奈良/広島
12/14~2/29, 40,703発話
ATR-Trel/NICT

関東: 山中湖/山中湖/石和/甲府
1/25~2/21, 40,301発話
JTB GMT/NEC/NICT

九州: 阿蘇/博多/天草/長崎
12/20~2/29, 58,263発話
熊本ソフトウェア/九州産交
/JTB/慶応大学ながさき/NICT

VoiceTra 約86万ダウンロード
1万発話/日の利用

VoiceTra
(2010. 7- 2013.3)
音声翻訳可能言語
(6カ国語)
テキスト翻訳可能言語
(21カ国語)

成田国際空港
NariTra
(2011.12)

KDDI
おはなし
アシスタント
(2013.7)

NTTドコモ
しゃべって
コンシェル
(2012.3)

通信プロトコ
ルの国際標
準化の実現

ITU標準化
(2010.10)

国際共同研究体U-STARを組
織し世界展開開始(2010.6)

VoiceTra技
術を世界標
準にすべく研
究を推進中



多言語音声翻訳技術の発展と社会還元



U-STAR(The Universal Speech Translation Advanced Research Consortium)

<http://www.ustar-consortium.com/index.html>



・24カ国・1地域、30研究機関・大学が参加する研究コンソーシアムにてグローバルに多言語音声翻訳技術の研究開発を推進中

世界標準システムとして発展(NICT/U-STAR)



VoiceTra(2010.7)
6ヶ国語の翻訳を実現
2013年3月までに
857,257件のダウンロードを達成



利用されている
VoiceTraの技術

- 認 音声認識
- 翻 多言語翻訳
- 合 音声合成

NICTの技術によって誕生した商用システム



成田空港
NariTra
(2011.12)



NTTドコモ
しゃべって
コンシェル
(2012.3)



FEAT
VoiceTra+
(2012.12)



KDDI
おはなし
アシスタント
(2013.7)



VoiceTra4U(2012.7)



VoiceTra4U

機能

- ・27言語間の翻訳、17言語(*)の音声入力
14言語の音声出力が可能

(*) 英語・米国語を区別



アジア言語
赤字は
ASEAN諸国

中東言語

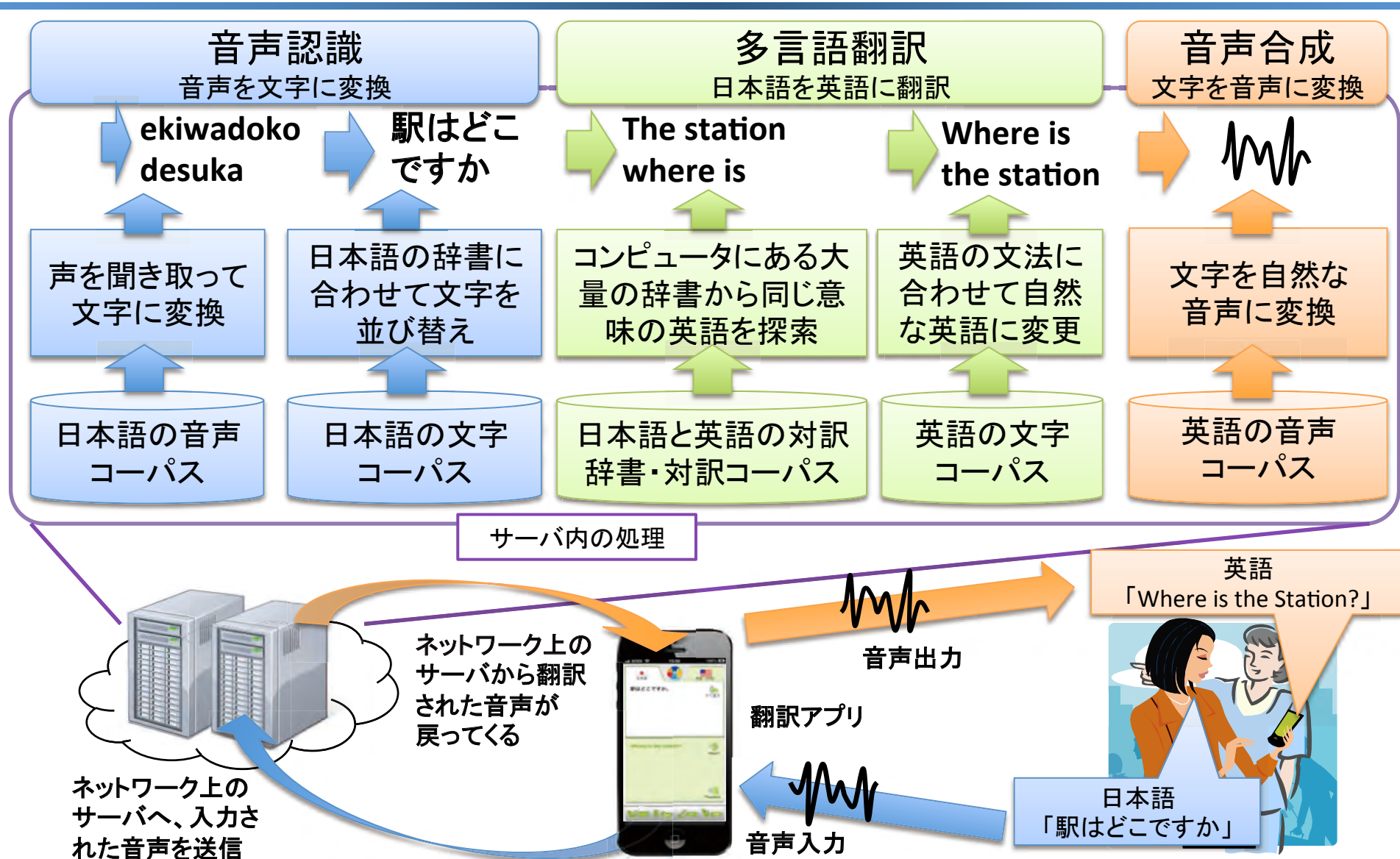
欧米露言語

音声翻訳対応状況

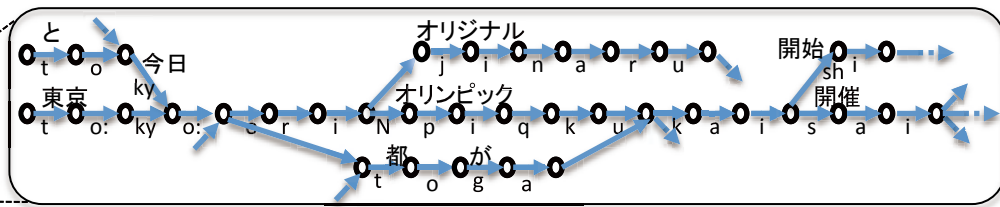
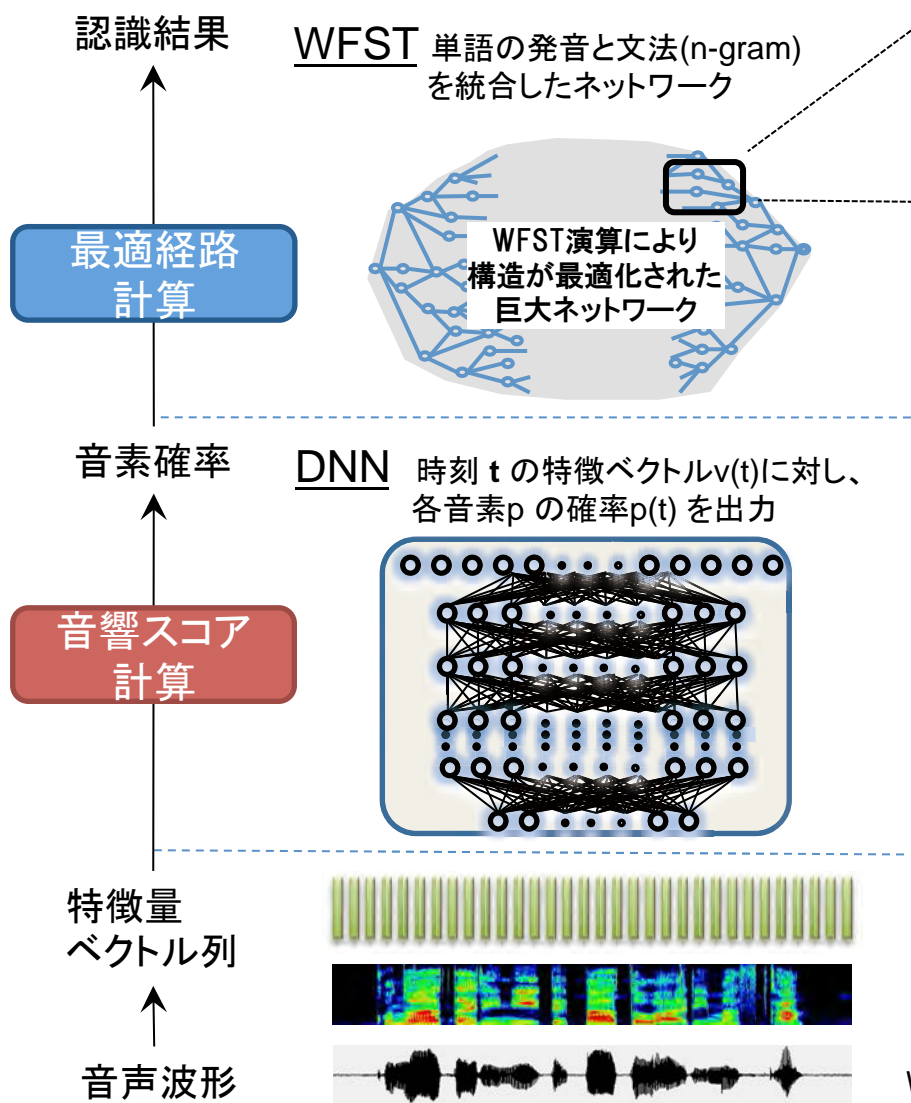
言語	入力		出力	
	音声	テキスト	音声	テキスト
日本語	✓	✓	✓	✓
中国語	✓	✓	✓	✓
韓国語	✓	✓	✓	✓
ウルドゥ語		✓		✓
シンハラ語		✓		✓
ゾンカ語		✓		✓
トルコ語	✓	✓	✓	✓
ネパール語		✓		✓
ヒンディ語	✓	✓	✓	✓
モンゴル語		✓	✓	✓
インドネシア語	✓	✓	✓	✓
タイ語	✓	✓	✓	✓
フィリピン語		✓		✓
ベトナム語	✓	✓	✓	✓
マレー語	✓	✓	✓	✓
アラビア語		✓		✓
英語	✓	✓	✓	✓
イタリア語		✓		✓
オランダ語	✓	✓		✓
スペイン語		✓		✓
デンマーク語		✓		✓
ドイツ語	✓	✓		✓
ハンガリー語	✓	✓	✓	✓
フランス語	✓	✓		✓
ポーランド語	✓	✓	✓	✓
ポルトガル語	✓	✓	✓	✓
ロシア語		✓		✓

音声翻訳の仕組み

NICT多言語音声翻訳の仕組み



音声認識の仕組み

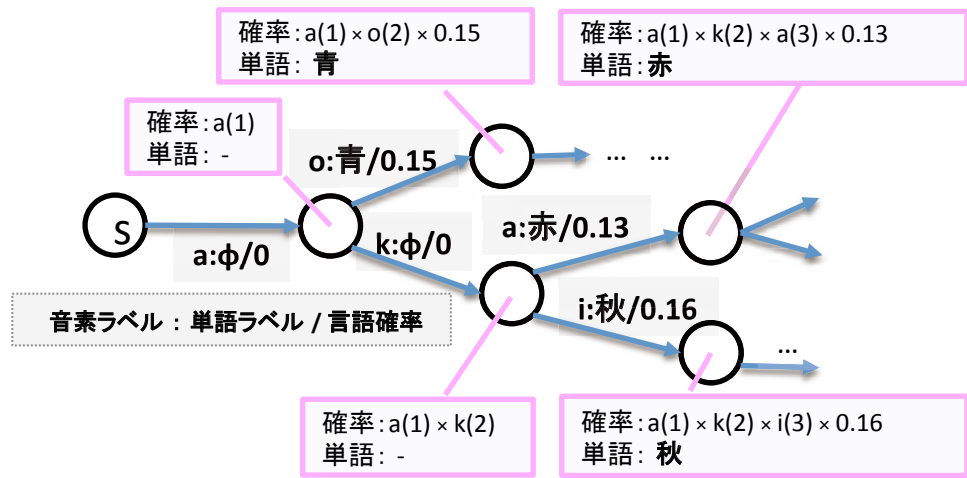


n-gram

単語列 W_1, W_2, \dots, W_{n-1} に続く n 番目の単語に単語 W が出現する確率。

- $P(\text{駅}|\text{東京})$... “東京”の後に“駅”が来る確率(2-gram)
- $P(\text{から}|\text{東京-駅})$... “東京駅”に続き“から”が来る確率(3-gram)
- $P(\text{から}|\text{東京-駅}) = 0.0001$
- $P(\text{の}|\text{東京-駅}) = 0.0001$
- $P(\text{八重洲口}|\text{東京-駅}) = 0.00002$ など膨大な組み合わせ

WFSTの探索 累積確率最大の経路探索



WFST: Weighted Finite State Transducer
DNN: Deep Neural Network



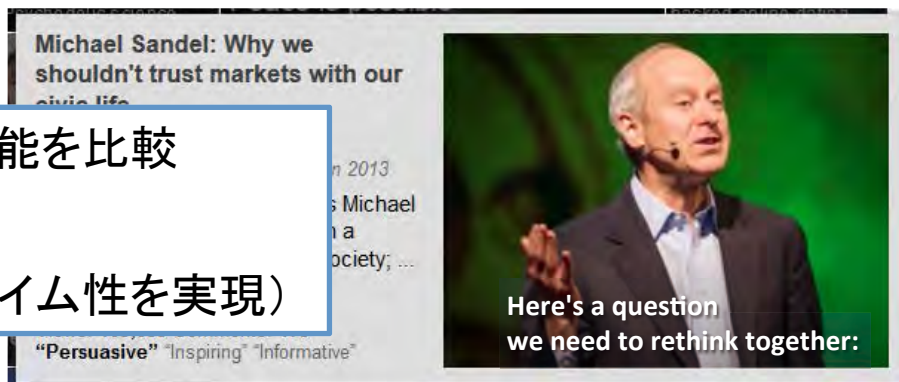
世界一の認識精度を誇る NICTの音声認識技術SprinTra

国際コンペで世界最高性能であることを証明

対象：英語音声講演データのインデキシング性能を比較

結果：NICTは最高の音声認識精度を示した

(他機関のシステムを遥かに凌ぐリアルタイム性を実現)



(画像は<http://www.youtube.com/user/TEDtalksDirector>より引用)

<単語誤り率(%)による評価>

参加組織	2012年	参加組織	2013年	参加組織	2014年
NICT	12.0	NICT	13.5	NICT	8.4
MITLL	12.4	KIT	14.4	EU-BRIDGE	9.8
KIT-NAIST	12.7	MITLL-AFRL	15.9	MITLL-AFRL	9.9
KIT-NAIST2	13.3	RWTH	16.0	KIT	11.4
RWTH	13.6	NAIST	16.2	FBK	11.4
UEDIN	14.4	UEDIN	22.1	LIUM	12.3
FBK	16.8	FBK	23.2	UEDIN	12.7
				IOIT	19.7

評価型国際ワークショップIWSLTで3年連続世界一

IWSLT 2014, International Workshop on Spoken Language Translation

EU-BRIDGE: RWTH & UEDIN & KIT & FBK
MITLL-AFRL: マサチューセッツ工科大学
 リンカーン研究所空軍研究所(アメリカ)
KIT: カールスルーエ工科大学(ドイツ)
 (Facebookに採用されたCMUのエンジンと同等)
RWTH: アーヘン工科大学(ドイツ)
FBK: ブルーノ・ケスラー財団-研究所(イタリア)
LIUM: メーヌ大学(フランス)
NAIST: 奈良先端科学技術大学院大学(日本)
UEDIN: エディンバラ大学(イギリス)
IOIT: ベトナム科学技術アカデミー情報技術研究所(ベトナム)

(HPの画像は<http://workshop2014.iwslt.org/>より引用)

英語音声認識デモビデオ

リアルタイム音声認識字幕付与システム

日 英 中

Select URL
CNN's exclus

自動

Input URL
http://www.youtub

Submit



情報通信研究機構
National Institute of Information and
Communications Technology

日本語音声認識デモビデオ

リアルタイム音声認識字幕付与システム 

日 英 中
Select URL
冬の味覚「干せ」
 自動
Input URL
http://www.youtube.com
Submit



情報通信研究機構
National Institute of Information and
Communications Technology

統計翻訳の仕組み

- | | |
|-----------------|--|
| 1. 京都駅はどこですか | 1. Could you direct me to Kyoto station? |
| 2. 駅はどこですか | 2. Where is the station? |
| 3. トイレはどこですか | 3. Where is the rest room? |
| 4. タクシー乗場はどこですか | 4. Where is the taxi stand? |
| 5. ここはどこですか | 5. Where am I? |

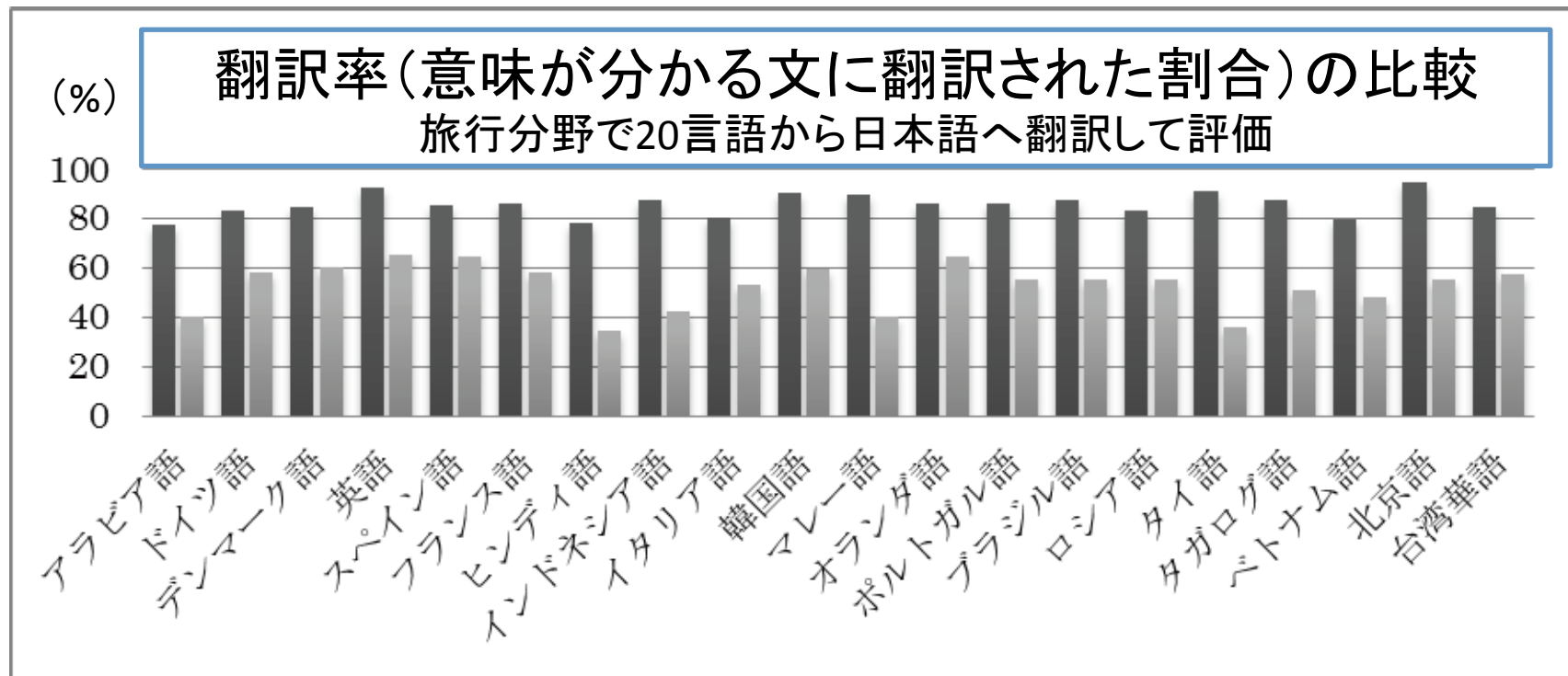
対訳
コーパス

確率付き対訳辞書の自動学習

どこですか	→	Where is	3/5=60%
どこですか	→	Could you direct me to	1/5=20%
どこですか	→	Where am	1/5=20%

NICT統計翻訳と従来手法の比較

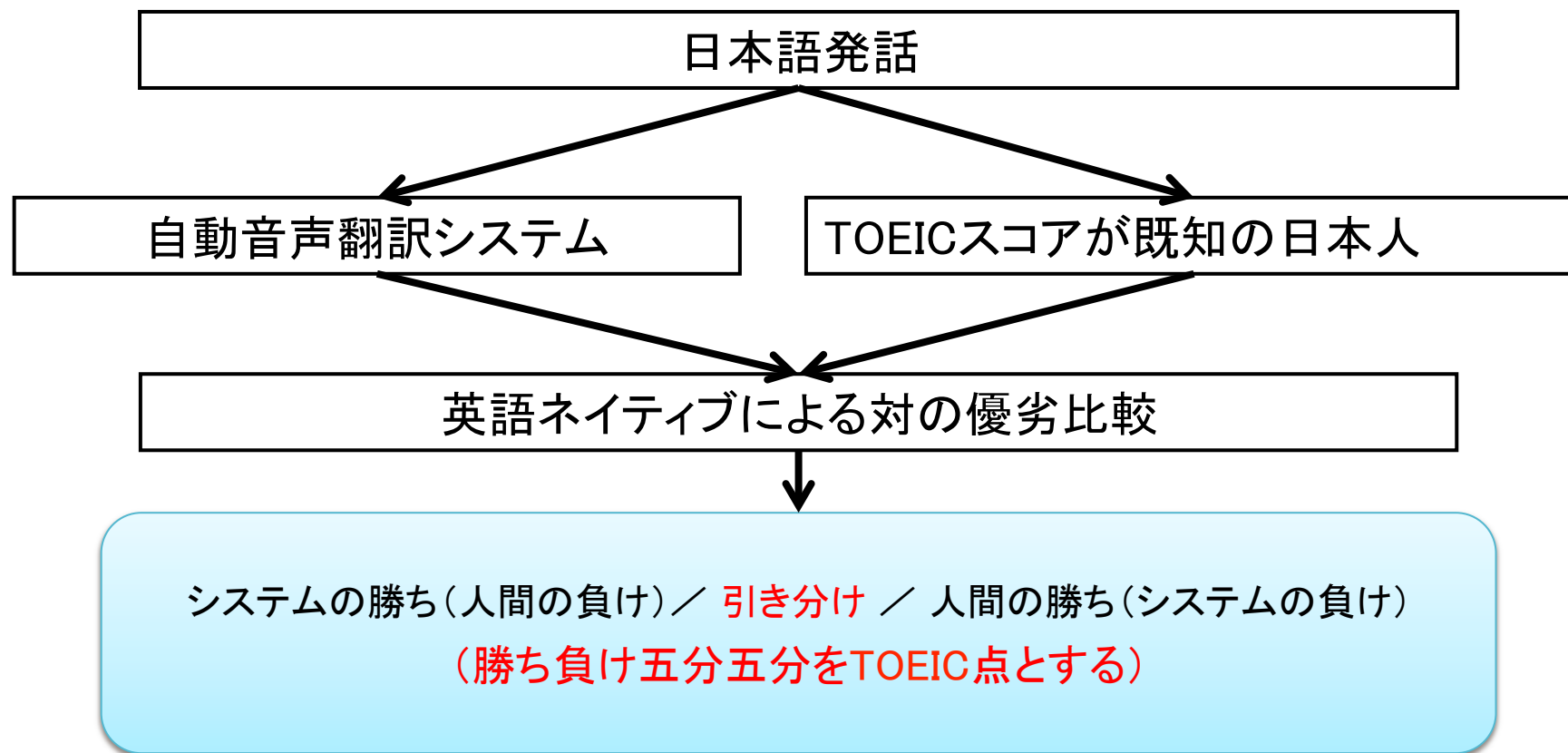
- 統計翻訳は多言語化を高精度で実現可能



NICT多言語翻訳技術
 他社手法

VoiceTraの翻訳性能はTOEIC600点の日本人に相当

いろいろなTOEICスコアの人間の音声翻訳能力と比較し、音声翻訳システムの能力がTOEICスコアでどのくらいの人間に相当するかを評価。





VoiceTra4Uは自由にお使い頂けます。

- iPhoneをお使いの方
 - AppStoreからダウンロードして下さい。
- Andoroidをお使いの方
 - GooglePlayダウンロードして下さい。



検索ウィンドウに「VoiceTra4U」を入力して下さい

NICTより公開しているアプリ

AppStore



NICTの技術を使ったアプリケーション
VoiceTra+,NariTra,こえとら,AssisTra,京のおすすめ
,SpeechCanvas



VoiceTra技術を使ったアプリケーション
VoiceTra+,VoiceTra4U,こえとら,SpeechCanvas

Googleplay



無料で
ダウンロードして、
ご自由に
お使い頂けます!!

NICTより公開しているサイトの一例

「みんなの自動翻訳@TexTra® (https://mt-auto-minhon-mlt.ucri.jgn-x.jp/)」

みんなの自動翻訳@TexTra®
日本語 ⇅

「みんなの自動翻訳@TexTra®」は、自動翻訳をみんなで育てるサイトです。

ユーザー名

パスワード

ログイン

新規登録

パスワード再発行



自動翻訳を使ってみよう!

「みんなの自動翻訳@TexTra®」にユーザー登録(無料)を行うと、あらかじめ登録された自動翻訳を試したり、自動でファイルを翻訳できたり、サイト上で翻訳エディタを使用して自分で翻訳することができます。



自動翻訳をカスタマイズしてみよう!

自分で用意した翻訳データを登録し、使用することで自分だけの自動翻訳にカスタマイズすることも可能です。また、みんなの対訳をあつめて、みんなの自動翻訳を定期的に更新しています。



Web APIを使ってみよう!

「みんなの自動翻訳@TexTra®」では自動翻訳、辞書引き、用語検索、類似文検索の機能をWeb APIとして公開しています。

ニュース

2014年7月28日 NICTと特許庁が多言語特許文献の高精度自動翻訳の実現に向けて協力合意
<http://www.nict.go.jp/press/2014/07/28-1.html>

2014年6月19日 「みんなの自動翻訳@TexTra®」を一般公開
<http://www.nict.go.jp/info/topics/2014/06/140619-1.html>

2014年6月19日 第9回AAMT長尾賞を受賞

Tweets Follow


minhonMT @minhonMT 31 Jan
 自動翻訳「みんなの(英語 - 日本語)」が更新されました。
mt-auto-minhon-mlt.ucri.jgn-x.jp

要素技術の応用：障がい者支援

音声認識技術の応用 聴覚障がい者支援アプリ『こえとら』

開発のきっかけ: VoiceTraの音声翻訳システムのデモをみたユーザからの問い合わせ



VoiceTraというNICTの音声翻訳機を使ってみたら日本語の音声認識性能がとても高かった。翻訳しないモードを作ってもらえないか
それができると聴覚障がい者と健聴者のコミュニケーションに役立つ

熊本聾学校の山田先生
(2011年11月)

2011

2012

2013

2014

2015

聾学校での実証実験

2013年6月
“こえとら”を
App Storeに公開

2014年6月
英語版 “KoeTra”を
App Storeに公開

2015年1月
事業会社に技術移転し、
2月にAndroid版公開。

聾学校、情報提供センター、企業などでの実証実験

『こえとら』の主な機能



音声認識技術と音声合成技術の活用

→ 音声から文字、文字から音声への変換により、
聴覚／発声機能を代替

- 聴覚障がい者（聴障者）が普段の生活の中で使うスマートフォンアプリ
- 健聴者とのコミュニケーションを円滑に行うことが目標
- 音声、文字、絵、地図などを使って情報のやり取りをすることができる

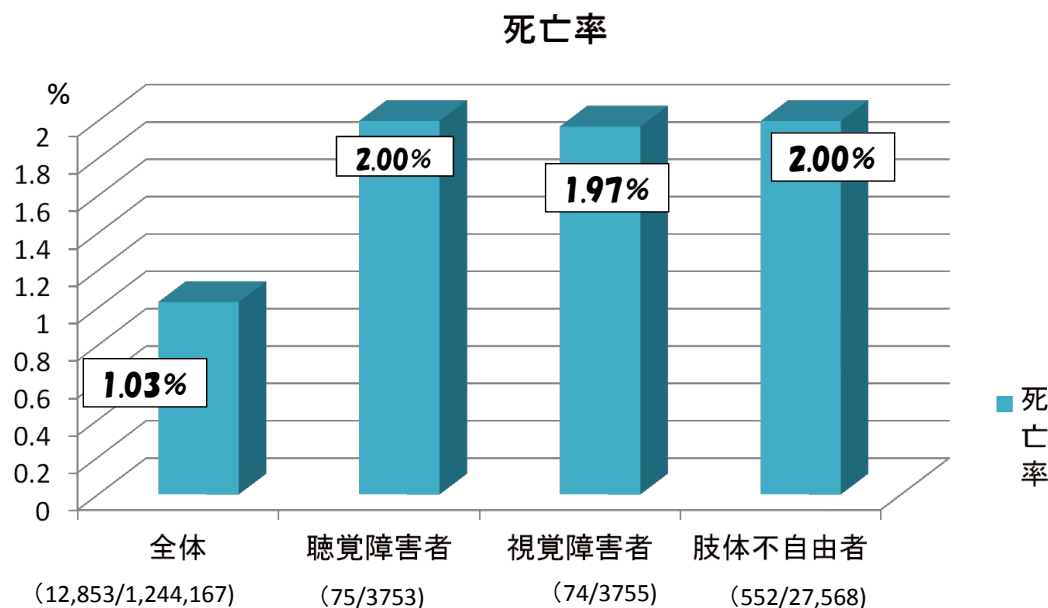
2分でわかる「こえとら」



サポートページ: <http://www2.nict.go.jp/univ-com/plan/applications/koetra/>

災害時の対応

東日本大震災における障がい者の死亡率は被災地全体の死亡率の**2倍**



出典：
災害時要援護者の避難支援に関する検討会（第2回）
資料8-2 p.9より抜粋

- 日常生活から**災害時**まで使える。
 - ▶ ネットがなくても必須部分のサポート



- iPhone上で稼働する音声認識・合成エンジンの開発
- サーバと端末内エンジンを組み合わせて使用する(ハイブリッド化)



先生のお話



健聴者のサポート

SpeechCanvasの開発



・困っている聴障者をサポート

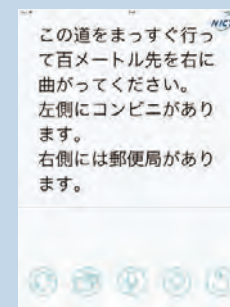


交番



役所などの窓口

地方自治体との連携



・聴障者と接する機会が多い健聴者も困っている
→困っている健聴者をサポートするツール



職場での会議、業務指示

一般企業への導入



学校での授業

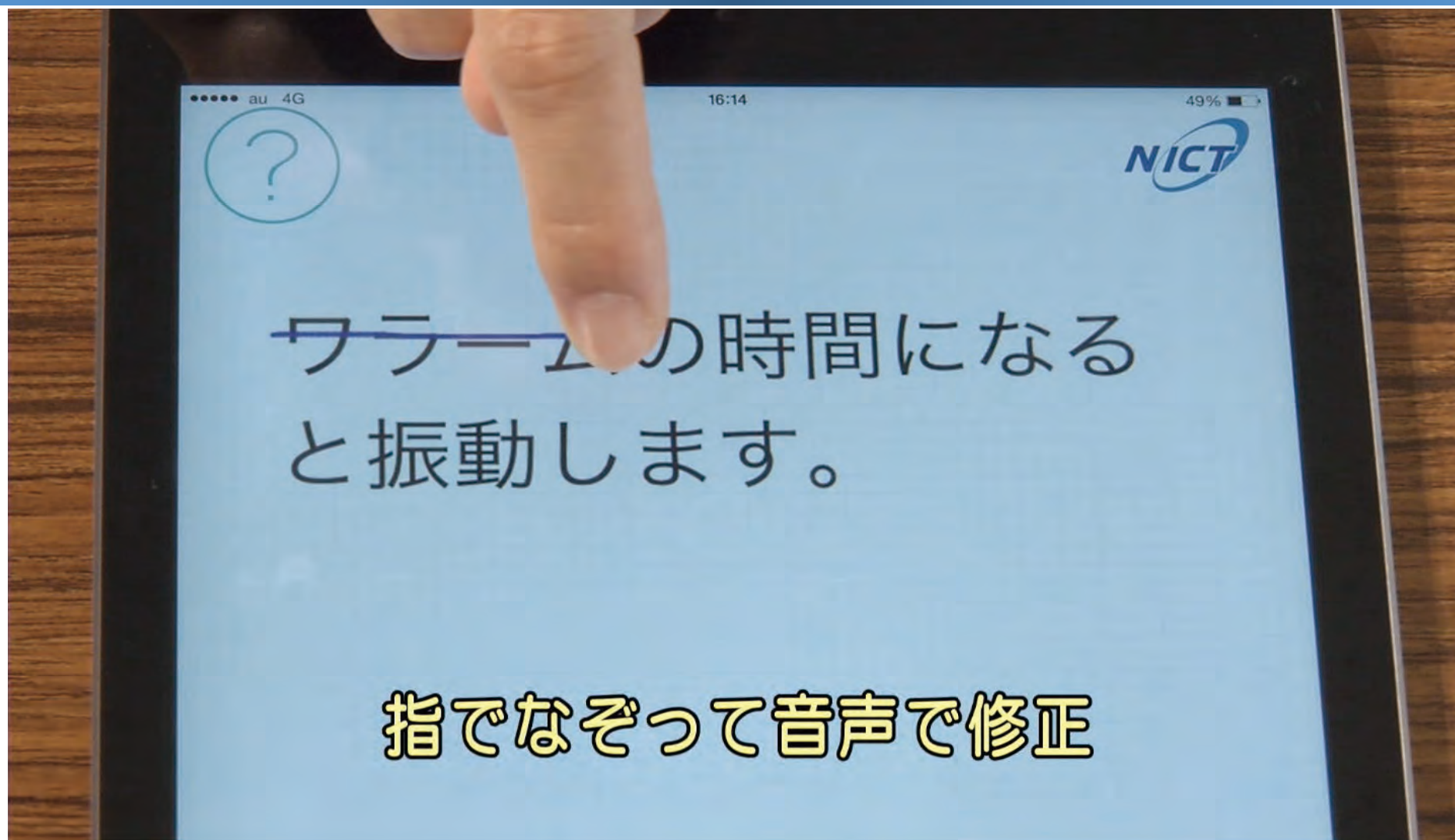
教員の補助



高齢者と同居する家族

家族の補助

2分でわかる「SpeechCanvas」



サポートページ: <http://speechcanvas.nict.go.jp/>

聴覚障がい者と健聴者の円滑なコミュニケーションの支援

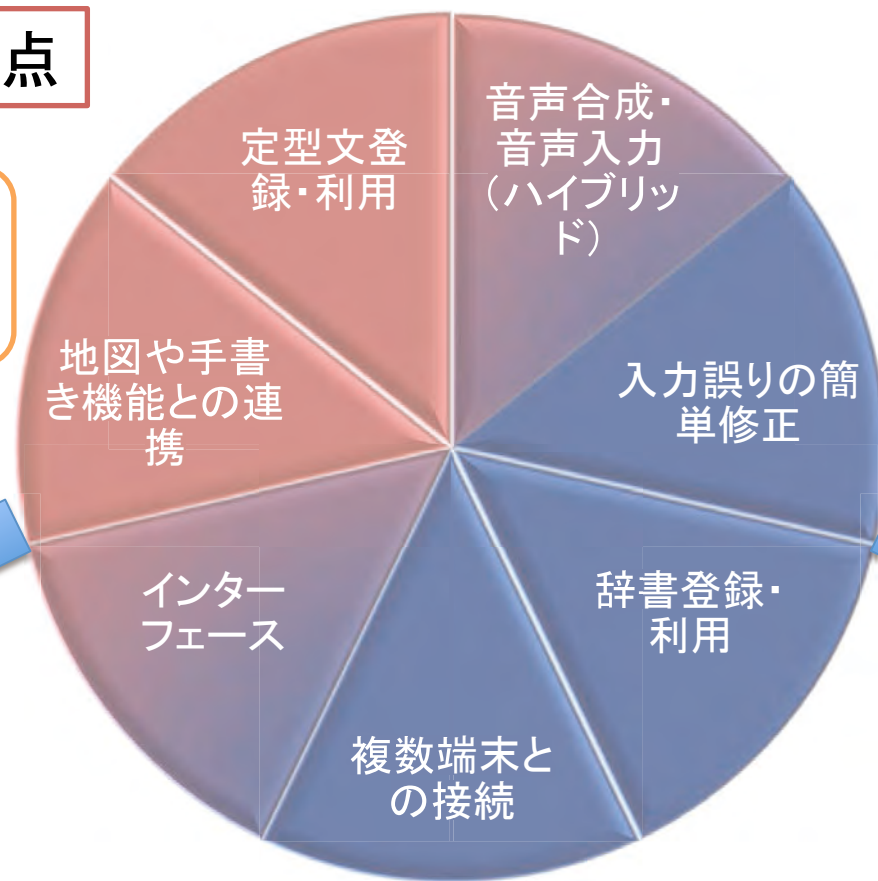
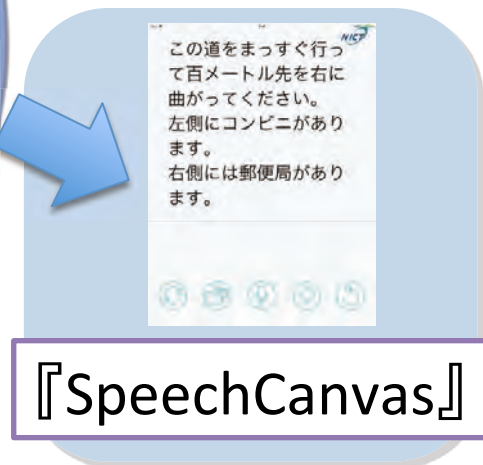
聴覚障がい者の視点

聴覚障がい者の情報送受信の効率化(手段の多様化)を重視



健聴者の視点

健聴者の音声入力の効率化を重視



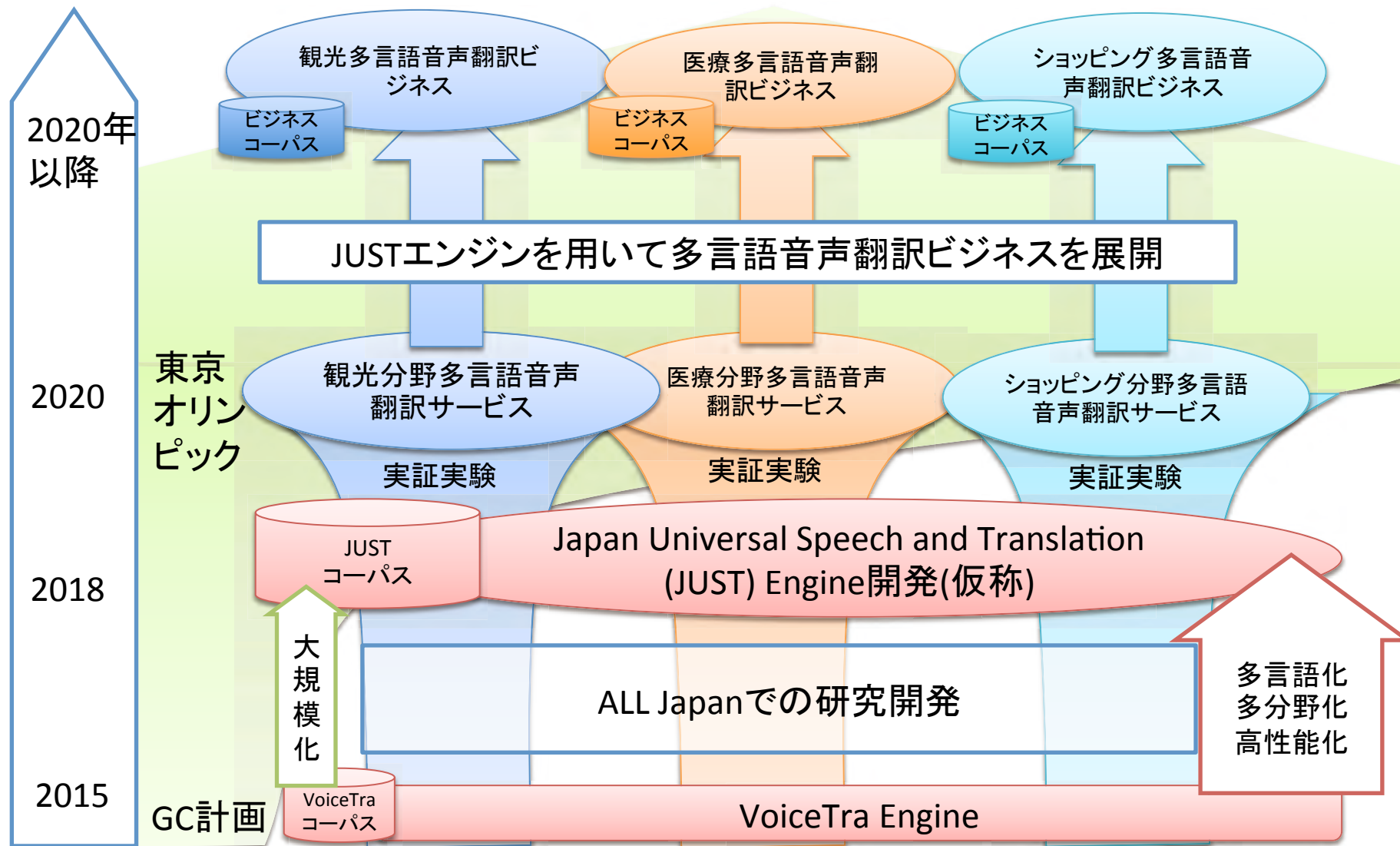
各々のニーズ(求められる機能)の違いに対応

グローバルコミュニケーション計画

- 平成26年4月11日に新藤大臣が記者会見にて、グローバルコミュニケーション計画を発表
- Mission
 - 世界の「言葉の壁」をなくす
- Vision
 - (1) グローバルで自由な交流の実現
 - (2) 日本のプレゼンス向上
 - (3) 東京オリンピック・パラリンピックでの「おもてなし」
- Action
 - 関係する企業や関係省庁等と連携、協力しながら、まずは6年間のロードマップを共有して取り組む
 - (プロジェクト1) 病院、商業施設、観光地等における社会実証
 - (プロジェクト2) 多言語音声翻訳の対応領域、対応言語を拡大するための集中的な研究開発投資
 - (プロジェクト3) 2020年東京オリンピックにおける社会実装



(画像はhttp://www.soumu.go.jp/menu_news/kaiken/01koho01_02000275.htmlより引用)

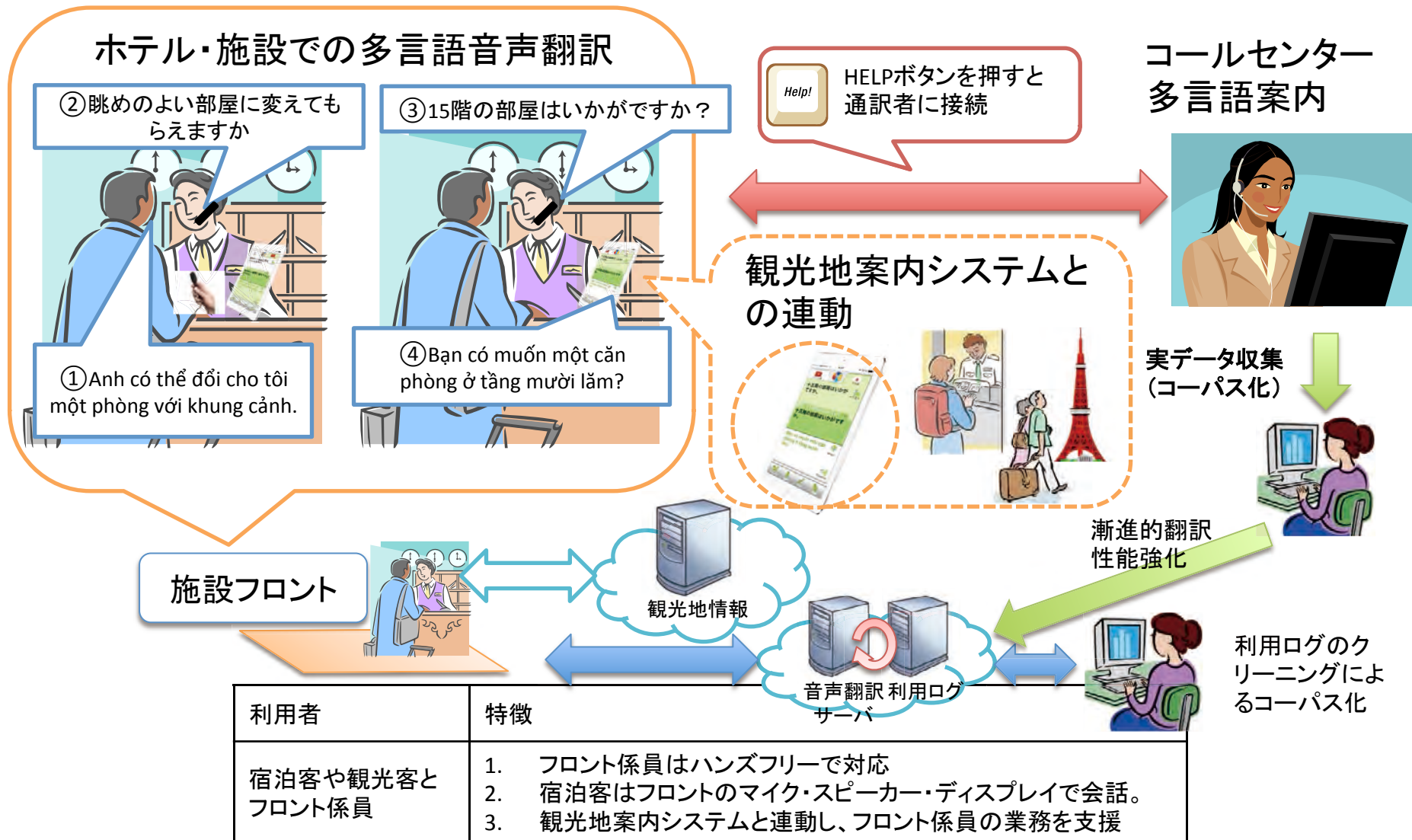


2020年の利用イメージ

目標：2020年にショッピング、交通、医療、ホテルなどで「普通」のICT機器として活用される



NICT サービスイメージ：観光分野（フロント）



利用者	特徴
宿泊客や観光客とフロント係員	<ol style="list-style-type: none"> 1. フロント係員はハンズフリーで対応 2. 宿泊客はフロントのマイク・スピーカー・ディスプレイで会話。 3. 観光地案内システムと連動し、フロント係員の業務を支援

サービスイメージ：観光分野（交通）

車載ディスプレイ表示

〇〇ホテルまで行きます。



分散型ハイブリッド（ネットワーク&ローカルサーバ）で常時サービス提供

駅構内等での騒音環境下での利用



ノイズ除去技術と雑音に頑健な音声認識技術の併用による騒音対策

券売機や据え置き型端末と対話による情報取得

利用ログのクリーニングによるコーパス化



音声翻訳利用ログサーバ



音声翻訳利用ログサーバ

公共交通機関



タクシー



利用者	特徴	利用者	特徴
乗客と運転手	<ol style="list-style-type: none"> 1. 運転手はヘッドセット等を用いてハンズフリーで対応 2. 分散型ハイブリッド（ネットワーク&ローカルサーバ）によって、ネットワークがないところでもサービス 3. 多言語観光案内コンテンツ・対話機能も搭載 	駅員（と一般外国人）	<ol style="list-style-type: none"> 1. ノイズ除去技術と雑音に頑健な音声認識技術の併用による騒音対策 2. 電子透かし技術の活用による効率的な情報伝達 3. 利用ログ収集・コーパス化による翻訳性能強化 4. 駅構内でのUI改良（ハンズフリー化）

サービスイメージ: ショッピング分野

ハンズフリーでの対応

① Apakah anda memiliki warna lain?



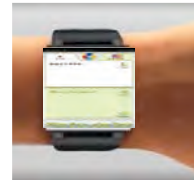
② 色違いはありますか。

③ お調べします。赤色があります。



④ Mari saya periksa. Ada merah.

接客のための
デバイス例



翻訳機能付きレジ

① Berapa harganya?



② いくらですか？

③ 画面をご覧ください。



④ Silahkan lihat di layar.

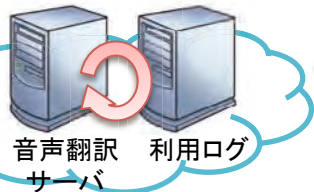
売り場



レジ

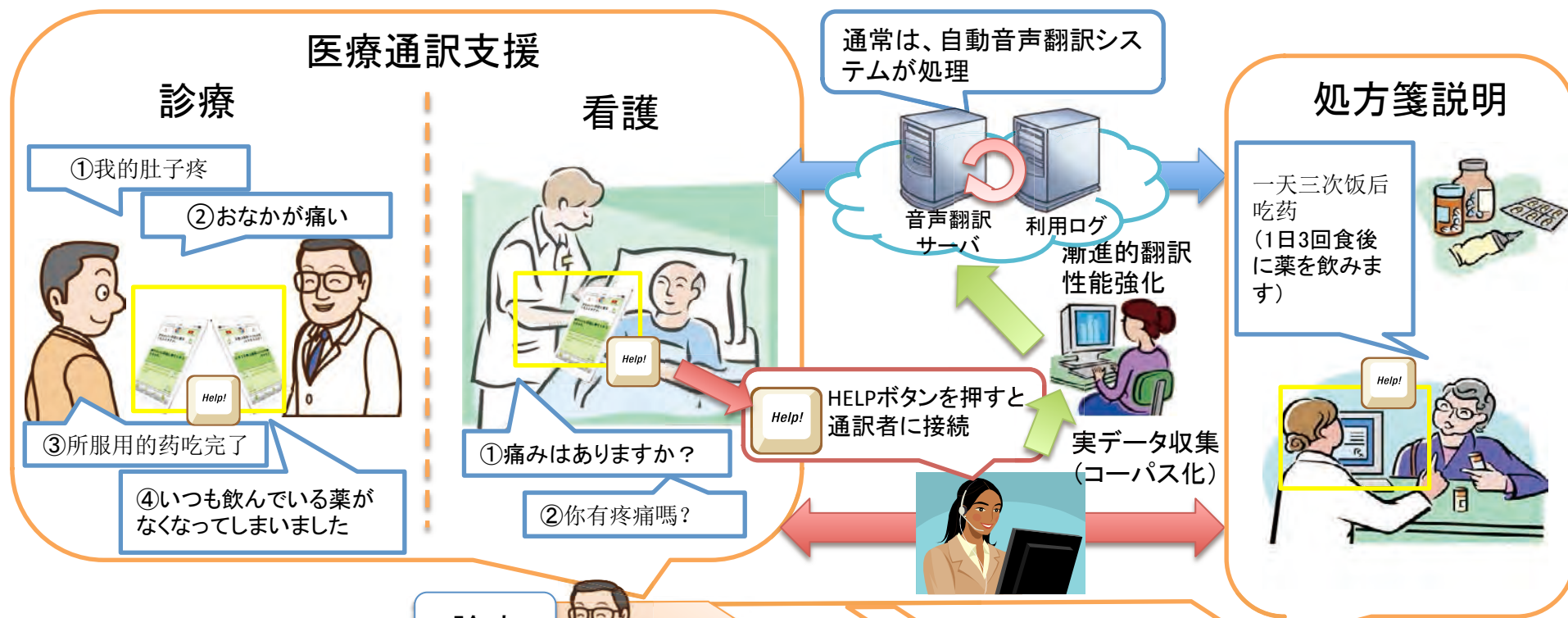


利用ログのクリーニングによるコーパス化

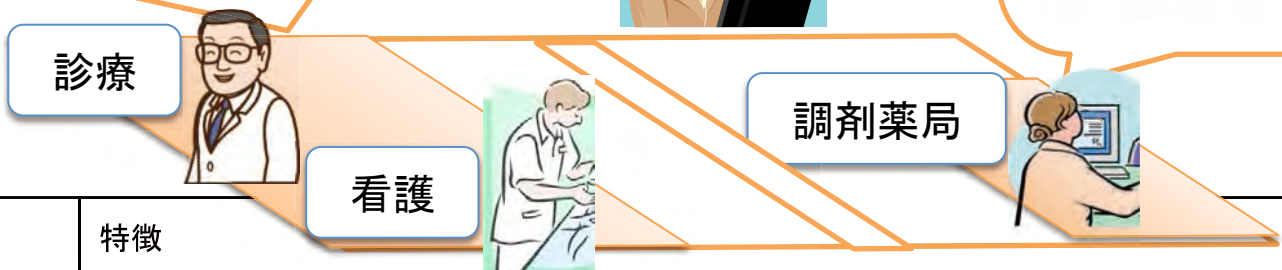


利用者	特徴
来客と販売員	<ol style="list-style-type: none"> 販売員はハンズフリー（ヘッドセットやHMD等）で対応 顧客は店内のマイク・スピーカ・ディスプレイで情報を得る。 商品データベースなどと連携して利用 翻訳機能付きレジでデパートから個人店までカバー。

サービスイメージ：医療分野



スマホやタブレットに加え、医療現場により適したUIの開発 (ハンズフリー化)



利用者	特徴
医療者(と一般外国人)	<ol style="list-style-type: none"> 1. コールセンターとして事業化済の人間系(通訳)と機械系(自動通訳)の連携 2. 実データ収集による漸進的翻訳性能強化 3. 医療現場でのUI改良(ハンズフリー化)

内閣府も重視

- 内閣府の動き
 - 科学技術イノベーション総合戦略2014 (H26.6.24)
 - 多言語翻訳技術はコア技術として記載



<http://www8.cao.go.jp/cstp/sogosenryaku/>



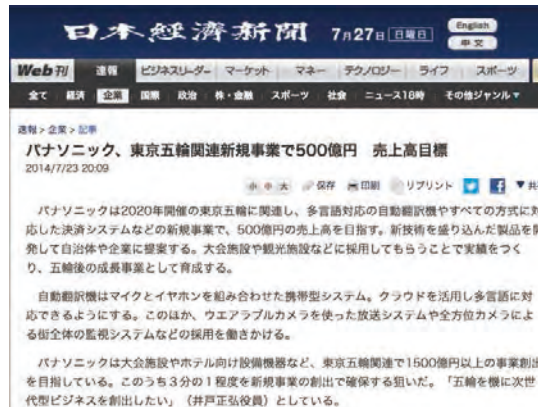
概要版: <http://www8.cao.go.jp/cstp/sogosenryaku/2014/gaiyo2014.pdf>

社会の動き

- グローバルコミュニケーション計画の発表によって、企業は動き出している。
- NICTとの共同による実証実験
- 2020年のオリンピック・パラリンピック東京大会を見据え、多言語音声翻訳技術の開発・実用化を推進する「グローバルコミュニケーション開発推進協議会」を設立



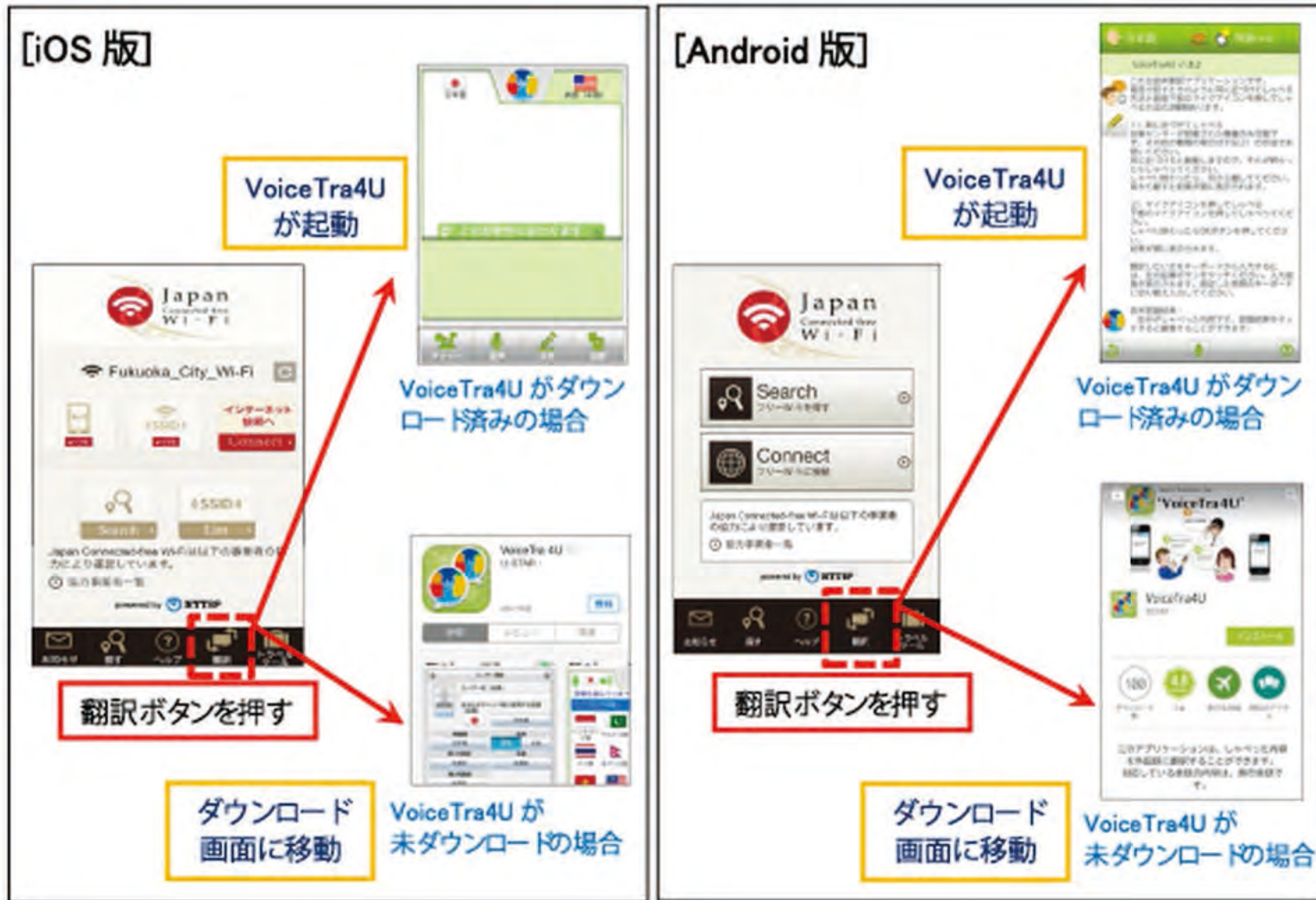
京浜急行電鉄での活用
朝日新聞(H26.7.16)



パナソニックの計画 日経新聞(H26.7.23)



NTTドコモの機械翻訳サービス
合併会社設立(H26.9.29)



(NICTプレスリリース資料(<http://www.nict.go.jp/press/2014/09/01-1.html>)より引用)

Japan Connected-free Wi-Fiの翻訳ボタンを押すと、VoiceTra4Uが起動又はダウンロードされる。

共同実証実験の例

- 京浜急行電鉄
 - 駅改札等での外国語対応にVoiceTra4Uを活用
- 東京都
 - 東京マラソンでボランティアがVoiceTra4Uを活用
- 東京大学附属病院
 - 窓口で専用アプリを設置



品川駅(改札)

(NICTプレスリリース資料
(<http://www.nict.go.jp/press/2014/07/15-1.html>)
より引用)

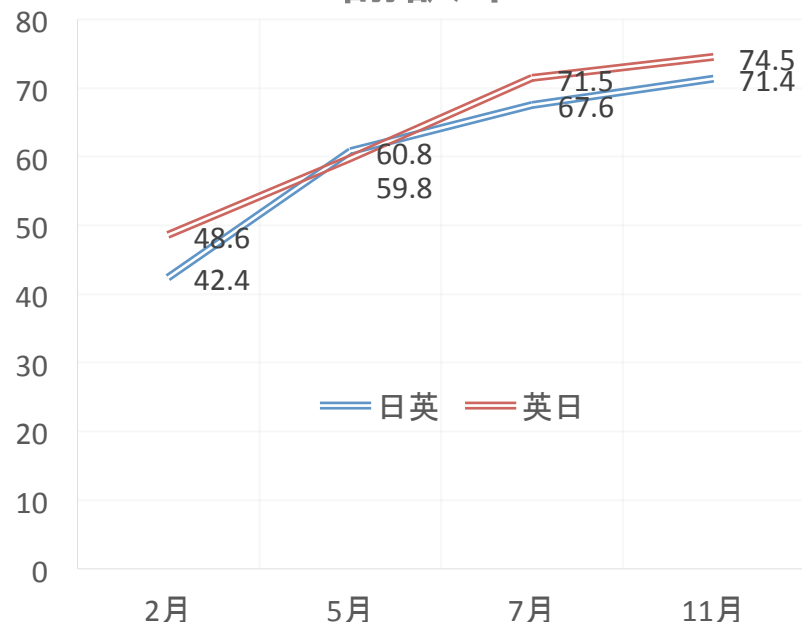


総合受付

(日経新聞(H27.1.25)でも紹介)

医療分野の翻訳性能向上

翻訳率



	入力文	従来翻訳	改良後の翻訳
英⇒日	My gums are swollen and painful.	歯ぐきが腫れと痛みがあります。	歯茎が腫れて痛いです
	My left knee hurts when I try to sit down or stand up.	左の頃立ったり坐ったりするひざが痛みます。	立ったり坐ったりするたびに左のひざが痛むのですが。
日⇒英	右手のどのあたりがしびれますか。	Do you have numbness in which part of your hand?	Which part of your hand is numb?
	入院生活についての説明はよくわかりましたか。	Did you understand the explanation about your hospitalization life often?	Did you understand the explanation about your hospitalization life?

- 目的
 - NICTを中心に産学官の力を結集
 - 多言語音声翻訳技術をアプリケーションに適用して社会実装
 - 世界の「言葉の壁」をなくし、グローバルで自由な交流を実現
- 主な活動内容
 - 研究開発及び標準化の推進
 - 社会実装及び実用化の促進
 - 情報の収集、交換及び提供 等
- 会員募集中
 - <http://gcp.nict.go.jp>



グローバルコミュニケーション
開発推進協議会

ホーム | 協会概要 | 会員募集

OUTLINE

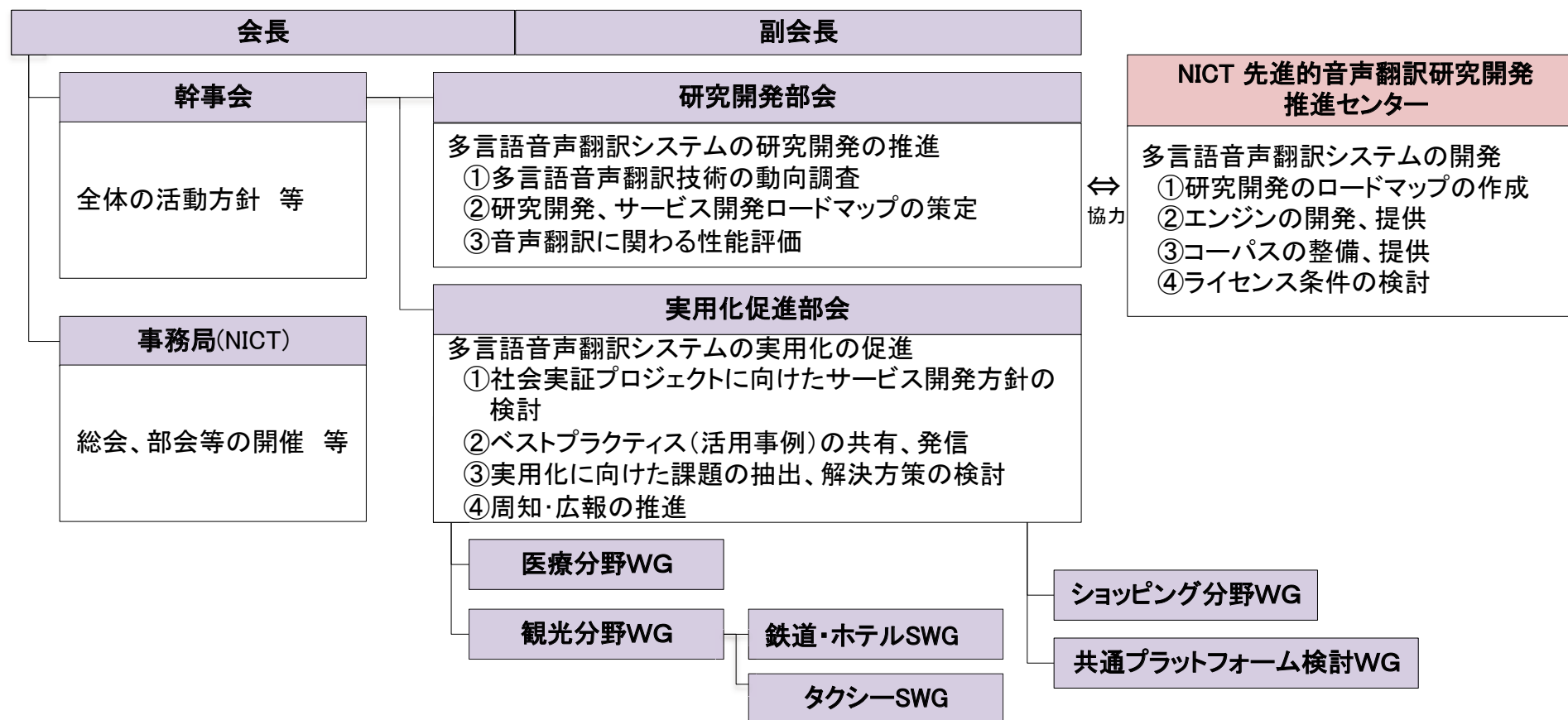
独立行政法人情報通信研究機構(NICT)、関連分野の企業及び有識者は、2020年のオリンピック・パラリンピック東京大会を見据え、多言語音声翻訳技術の開発・実用化を推進する「グローバルコミュニケーション開発推進協議会」を設立いたしました。
会員を随時募集しておりますので詳しくは[会員募集](#)ページをご覧ください。

本推進協議会に関する問い合わせ先
グローバルコミュニケーション開発推進協議会 事務局
白土 保、信岡 道明、泉田 晋男、中村 憲治
Tel: 042-327-7340 Fax: 042-327-5837 E-mail: gcp-imb@nict.go.jp

Copyright 2014 グローバルコミュニケーション開発推進協議会 All Rights Reserved.



グローバルコミュニケーション開発推進協議会 (体制図)



設立発起人(順不同、敬称略)

東京大学大学院 情報学環長 須藤 修 (会長)、奈良先端科学技術大学院大学 教授 中村 哲 (研究開発部会長)、ATR-Trek、KDDI (実用化促進部会長)、ソニー、東芝、凸版印刷、日本電気、NTT (副会長)、NHK、パナソニック (副会長)、日立製作所、フィート、富士通、NICT (副会長)



2020年の音声翻訳のイメージ



まとめ

私たちは、世界の「言語の壁」を越える技術としての多言語翻訳技術のレベルを高め、いつでも、どこでも、誰にでも価値ある情報を伝えられるグローバルコミュニケーション社会の実現を目指します。

