

NHK放送技術研究所の ユニバーサル放送への取り組み

NHK放送技術研究所
ヒューマンインターフェース研究部

研究開発を進めている技術

1. 生放送に字幕を付与するための音声認識
2. スポーツ実況を補完する音声ガイド
3. 気象情報を伝える手話CG

生放送に字幕を付与するための音声認識

1. 復唱音声を認識して字幕に

- ・番組の制約なく字幕付与可能
- ・復唱者の確保が課題
- ・認識誤りは人手で修正



2. 番組音声を直接認識して字幕に

- ・発話スタイルや話題に制約
- ・ニュース原稿データベースを学習
- ・認識誤りは人手で修正



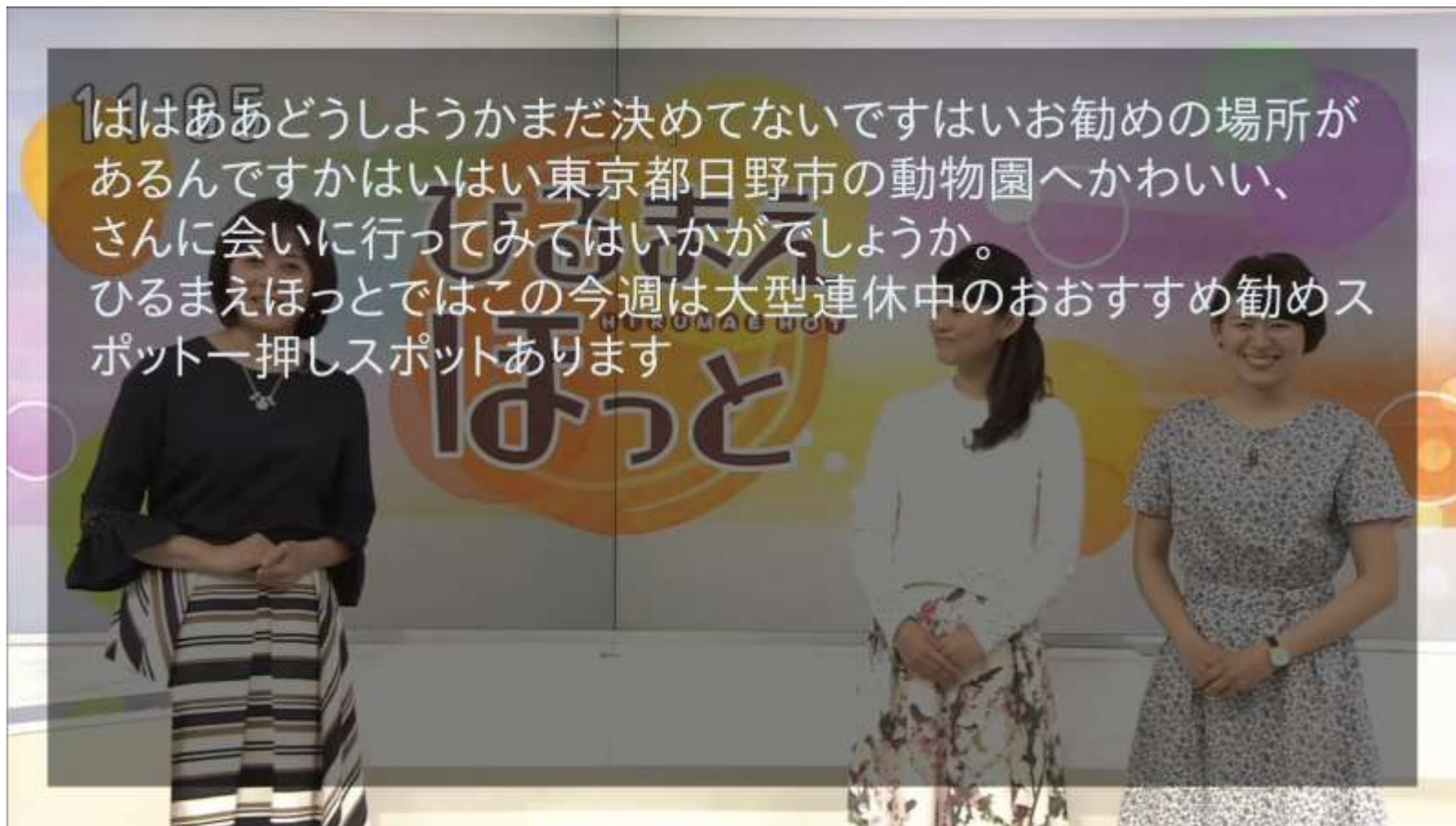
3. 音声認識を用いて読み原稿を選定して字幕に

- ・あらかじめ用意できる原稿だけが字幕化可能

音声認識の課題

1. 認識精度は100%にならない
 - ・統計的な手法を用いる限界
 - ・95%の認識精度ならば、修正する単語は6秒に1回
2. 復唱、言い換えが必要 ⇒ 解決の方向へ
 - ・様々な発話スタイルや話題を認識
 - ・深層学習によるブレークスルー
3. 原稿がない部分も認識 ⇒ 解決の方向へ
 - ・大規模な学習データと大規模なモデル（辞書）

最新の音声認識



残された課題

1. さらなる認識精度の改善

- ・ 一般話者の対話やインタビュー
- ・ 学習データの増強とともに音声認識のブレークスルー

2. 認識誤りの修正

- ・ 原稿が利用できれば原稿で修正
- ・ 原稿が利用できない部分の誤りを簡易に修正

スポーツ実況を補完する音声ガイド

■ テレビのアナウンサー実況では
画面に表示されている情報の言及は少なく
音声だけでは状況把握がしにくくなることもある

■ 表示されているテロップなどは
主催者などがリアルタイムに
配信するデータから制作

⇒ このデータから
音声ガイドを自動生成可能

音声ガイド

配信データ

今の試合状況

試合の説明

JPN 1-0 ENG
前半 12:34

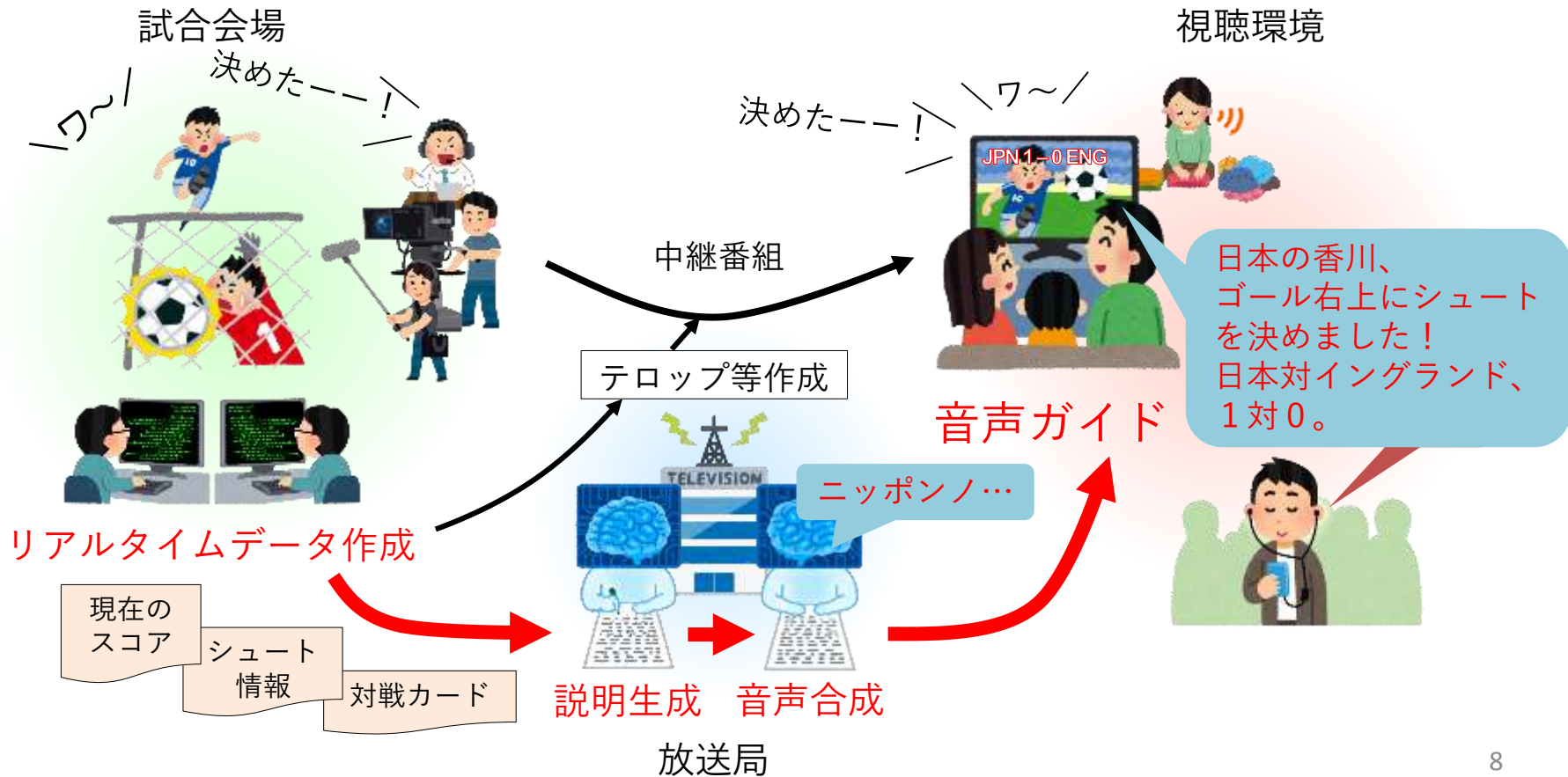
サッカー男子決勝

10 香川真司

選手名

決めた——！

音声ガイド自動制作の流れ



アナウンサー／解説者

バックハンド。

アナウンサー／解説者

フォア！

アナウンサー／解説者

連続攻撃！

アナウンサー／解説者

(はい、) 決まったー！ (いいですね！)

アナウンサー／解説者

やはり落ち着いてますね。
はい、攻撃をカウンターする、あの、
そこまでの余裕が出てきました。

福原のポイント。

現在の挑戦と残された課題

挑戦

- ・ 自動実況 （ネット配信競技に実況を付与）
- ・ アナウンサー実況との共存（発話予測、文の最適化）
- ・ オリンピック以外のスポーツへの適用

課題

- ・ スポーツ番組以外の番組への一般化

気象情報を伝える手話CG

■ 手話放送の増加しない現状

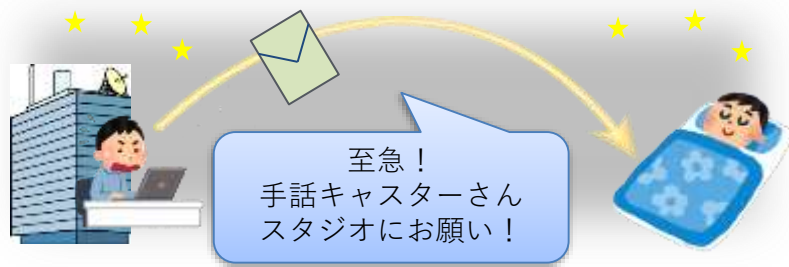
- 手話通訳士の絶対数が少ない
- 手話通訳士の放送局常駐は困難



緊急の気象情報第一報を
CGアニメーションの手話で

■ 日本語から手話の翻訳には多くの課題

- データから手話を生成して誤りのない情報を提供



ネット配信で検証

2017年2月20日 気象情報手話CGサイト一般公開



- ・ 府県天気予報電文に対応
関東7都県 県庁所在地
- ・ 原則1日3回（5時、11時、17時）
自動で動画を作成して配信



聾者がわかりやすいレイアウト

聾者がいつでも手話で
最新の天気予報を確認できる

天気

きょう

あす



ここに、公共放送



残された課題

1. 任意の日本語文を手話に翻訳

- ・ 直接翻訳ではなく、さまざまな意識が必要
- ・ 手話通訳士のスキルを持つ翻訳機を実現するブレークスルー

2. 翻訳、動作の誤りの発見と修正

- ・ 音声認識のように、翻訳誤りを発見するのは困難
（発見は手話通訳士しかできない）
- ・ 誤りを発見できたとしても、修正する手段がさらに必要

今後の展望

AI技術の進展

- ・ AI技術の進展が音声認識や合成などの様々なブレークスルーを生み出してきた
- ・ 今後もしばらくは不可能が可能になってくる
- ・ 一方で、基礎技術の確立には従来通り数年の時間を要している
- ・ 困っている人の声をよく聞いた研究推進が必要である

AI技術の限界

- ・ 機械学習の枠組みの中で、絶対誤らないAIの実現は困難
- ・ 誤りの修正手段が必要
- ・ 重大な誤りを起こさないAIの実現に向けた研究開発が必要