

多言語対応推進フォーラムの舞台裏 ～リアルタイム翻訳字幕～

ヤマハ株式会社 クラウドビジネス推進部 SoundUDグループ 森口 翔太

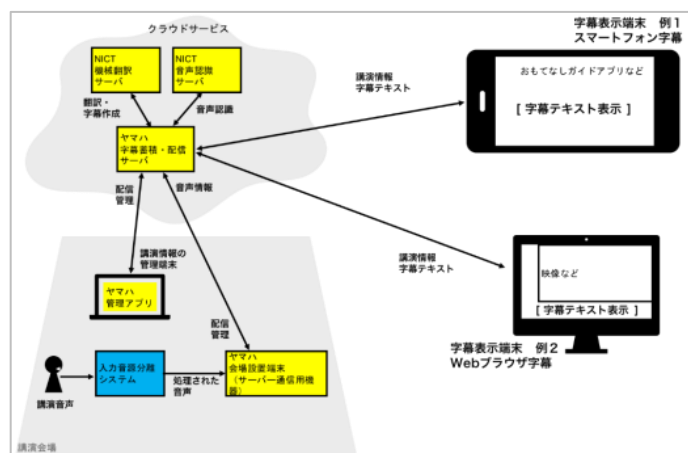
令和2年12月23日、一年延期となった2020年オリンピック・パラリンピック大会の開催を見据え、withコロナ時代の多言語対応の取組を紹介するため、「多言語対応推進フォーラム」が開催されました。

フォーラムでは登壇者の発言をリアルタイムで日本語から英語に翻訳し、配信画面へと字幕を表示する「リアルタイム翻訳字幕」技術が4社で分担して運用されました。ヤマハ株式会社は、総務省からの委託事業として開発中である、自動で認識し翻訳を行う「多言語情報配信サービス（仮称）」を運用し、リアルタイム翻訳字幕配信を行いました。

「多言語情報配信サービス（仮称）」は、講演やアナウンスなどの音声を自動で認識・翻訳し、スマートフォンやPCに多言語字幕情報として配信するサービスです。

ヤマハ株式会社ではこれまでも公共施設における多言語対応の支援を行い、東京都や総務省の委託事業として様々なプログラムに取り組んできました。

この「多言語情報配信サービス（仮称）」は、国立開発法人 情報通信研究機構（NICT）の持つ音声認識・機械翻訳サーバーが利用されています。パネルディスカッションのように複数の話者が混在する場合に、音声を適切に認識・翻訳できるよう、入力音源分離システムの研究も並行して行われました。



ヤマハ株式会社 クラウドビジネス推進部 SoundUDグループの森口氏は、「今回開催された『多言語対応推進フォーラム』では、複数の話者が同時に話をすることが想定されていたため、話者を特定するための音声経路の工夫や表示方法の処理が難しかった。」と、準備段階における苦労を振り返りました。話者ごとにラベルをつける工夫を加え、字幕と連動して自動で話者の名前が表示されるようにしたことで、話し手の切り替わりをわかりやすく表現することができました。表示の際、遅延はあったものの、全体として読みやすい字幕になったと評価しました。また、字幕の表示に関する監修を外国人に依頼し、文字の表示量やスクロールスピードを細かく調整したことで読みやすい字幕になり、外国人の参加者からも字幕が見やすい、という評価を得ています。一方で、森口氏は、「外国人の方に評価いただいたことで、この技術・サービスの優れている点が明確になったものの、課題も浮き上がってきた。」と話しました。「字幕の遅延は話をわかりにくくし、進行に与える影響が大きいので、今後はよりスピーディに表示する工夫が必要である。さらに字幕が途切れたり、会場内に音声反響することによって意図しないような結果が表示されたりすることがあり、認識・翻訳精度の部分に課題が残されている。」と述べました。

森口氏は最後に、「『多言語対応推進フォーラム』では、複数の登壇者がマイクを使用したことで、複数話者が存在する場合の音声サンプルを数多く集めることができたので、今後の研究に役立てることができる。翻訳事業に関わる多くの方に自動認識・翻訳を実際に見ていただけたことで、社会全体で多言語情報サービスがさらに加速していくことを願っている。」と話しました。

(令和3年2月作成)

※フォーラムの様子は、こちらのURLからご覧いただけます <https://tokyodouga.jp/nf68gam9q1a.html>

