

点線枠内は、構成員による発言

1 実現を想定する多言語字幕サービスモデル

(1) 想定されるサービスモデル・字幕作成プロセス

<親回第1回>

- どのような形の翻訳システムをきちっとした放送にのせるか、ハイブリッドキャスト、さらに、使い勝手は悪いがないよりはましとして、スマホで字幕が出るといったケースが考えられる。フレキシブルにロードマップを2020年に向けて書いていくことが重要。
- どこまでの不正確性を許容するのかというのが1つの観点。放送事業者が直接提供するとなると正確性が求められるが、サードパーティが提供するなど、様々な運用上の仕組みも併せてWGで検討することが、物事を進めるうえで重要。不正確だけど物事が進むということは、一種のトレードオフの関係。
- 受信機への多言語フォントの搭載が課題。
- 受信機への多言語フォントの搭載といった技術的な問題もあるので検討してほしい。

<WG第1回>

- 言語をどうするかを含め、ユーザーは誰か、どういうニーズがあるかを踏まえながら議論することが必要。
- ハイブリッドキャストの受信機は、現在家庭向けに開発されている。多言語字幕のホテルでの利用、サイネージでの利用等まで含めて考えると、ユーザー（対象）を誰と見なして議論するかが重要。

<WG第2回>

- マネージされた中でどれだけ自由を担保するのが、オリンピックの後も続くサービスになれるのかどうかの大きな分かれ目。
- 日本発のいいサービスとして世界に広げていく源にするためには、いかに盛り上げていけるか、みんなに使われるものになっていくかということが重要。そのときに放送事業者の責任と自由さのせめぎ合い、この辺についてその運用とか実証実験があり得るのではないか。
- スマートテレビにおける通信・放送の連携サービスが真に新たな価値を生み出すには、それによってまず地上基幹放送の特質を損なうことがないことが一番大事であり、むしろ、これを補強、発展させたい。
- 災害放送時にやはり英語での字幕が大事なことだと思う。そういう地上基幹放送の特質がむしろ発展するという視点であれば民間放送事業者としても検討に持っていきやすい。
- 正確性と迅速性は放送の根幹。多言語字幕サービスで想定しているベストエフォートの正確性では、放送事業者以外の実施主体を検討することが現実的。
- 翻訳の入力信号を何にするのか、誰がやるのか、機械がやるのか人間がやるのか、人間と機械がどれぐらいの割合でやるのか。幾つか典型的なパターンがあるのではないか。それを考えていかないと、この後実証というイメージもできていかないのではないか。
- 音声入力なのかテキスト入力なのかも、結構技術的には難しさが全然違うので、ポイントだと思う。
- 日本にいる外国の方がどういう使われ方をされるのかというのが、先に何か要件整理のようなものがないと、この議論はなかなか収束しないのではないか。
- ハイブリッドキャストで通信経路で端末に向けて行われるサービスは、不特定の少数を対象したサービスも提供できるところがメリット。しかし、フォントやキャッシュメモリーなど、少数派向けの多言語字幕のためだけに受信端末をつくることになる、果たしてメーカーがつくってくれるのかという問題は考える必要がある。
- ある特定目的だけのために開発するというのは、メーカーにとってはあまり効率的なことではない。
- どのようなユーザにどのようなシチュエーションでメリットがあるのかを整理しないと、サービス設計とか端末設計ができない。
- 受信機メーカーの立場から言うと、どうしても安全サイドに設計してしまわなくていけなくて、同期の問題であったり、最後はユーザのクレームの問題だったりというのがどうしても気になってしまう。一方、技術の立場から考えてみると、自動翻訳とか音声認識というのはものすごく進歩しているので、それを単に受信機という立場だけから切り捨てていいものではないと思う。

(2) 想定する言語・分野、研究開発を推進する言語・分野

<親回第1回>

- 五輪も控え、さらなる翻訳技術の向上に期待。優先的に実施する分野、言語を決めることが早期の実現につながるのではないか。
- 対象とする番組、言語の選択が課題。
- 対象となる言語は、英語等はもちろん重要。さらに日本語も外国語として考える必要がある。自国で働いている人にいかに自国の言葉を教えるか。

<WG第1回>

- 訪日外国人等数のデータを踏まえる必要等もあるのではないか。
- 英中韓は技術開発の人手を確保するのは比較的容易と思われるが、少数言語は難しいのではないか。
- 在日外国人向けに分かりやすい日本語というのも選択肢の一つではないか。

<WG第2回>

- 技術開発の観点からも、言語や分野を絞っていく必要があるのではないか。言語としては英語から、分野としては災害、避難情報などの安全安心情報からと思うが、いずれにせよ、実証実験などで検証していくことが必要ではないか。
- 言語や分野をある程度想定したほうがよい。一番想定されそうなのが、やはりニュース。日本にいる外国人を考えると英語と中国語、韓国語というのは多分妥当なところではないか。
- バラエティー番組とかなら気軽に始められるのではないか。
- ニュース番組のアナウンサーに限ってやれば、機械翻訳にしても、かなり定型的な表現が多いので、そんなに無理な内容ではない。

(3) 翻訳内容の正確性や表示の遅延

<親回第1回>

- オンライン字幕（生放送番組の字幕）での誤り削減と遅れ軽減が課題。（人手での修正は人材の確保と時間遅れが課題。直訳は原文よりも長くなる傾向、時間遅れが蓄積。）
- 提供者によってビジネスモデルが変わり、水準が変わるのではないか。
- ハイブリッドキャストの場合その運用主体・事業主体をどうするか、放送の場合は完全でないといけませんが、ハイブリッドキャストの場合はどの程度のレベルでやるか、誰がどういう基準でやるか、それもある程度ビジネスにしていかななくてはならない、WGで検討してほしい。

<WG第1回>

- 非リアルタイム翻訳の場合は、放送のストリームと通信で送られる字幕の同期をとることはできるが、ネットワーク環境にも依存して遅延が生じる。通信の遅延に併せて放送表示を遅らせるということも技術的には考えられるが、少しでも早く提供するという放送の使命からはどうか（ほとんどあり得ないのでないか。）リアルタイム翻訳の場合、人が修正するとその分さらに遅延が生じる。
- どの程度までの遅延が、実用サービスとして許容されるかは、利用者にもよるのではないか。
- 最近、タイムシフト視聴も広がっており、録画した番組にハイブリッドキャストで字幕を提供するというのもあるのではないか。
なお、録画番組とハイブリッドキャストの連携についての技術仕様は、IPTVフォーラムで検討中。

<WG第2回>

- 自動認識というのは永久に100%にはいかないと思うが、東京オリンピックが開催される2020年まで、あと5年というスパンを考えてその方向で頑張っていけば、それなりのものができるのではないか。
- 通信は自宅の環境にかなり依存していることもあり、そもそもテレビ自体が通信系を必ずつなぐものでもないこともあるので、テレビ受信機で放送と通信を確実に同期したサービスを提供するというのは極めて難しい。
- 何分間の映像を遅延させるというバッファを受信に持たせるというのは経済的にはあまり成り立たない。
- 生放送の字幕は、技術的な観点からどうしても遅れがある。例えば30秒遅れて出る場合、字幕を見ながら映像を理解するというアプリケーションの場合にはなかなか難しいのではないか。
- ハンディキャップの方に出すのであれば、少し時間はおくれてもニュースだけきちんと届けばいいとか、外国人であれば、ニュースや天気予報をある程度時間を置かずに見せたほういいとか、そういったようなユースケースの整理すべき。
- 正確性を求めても誰もこの程度いいとは言えない。結構高次元で決めてしまい、いいですと言うしかないと思う。

(4) 視聴者から見て外形的に、字幕と放送を区別できる工夫・形態

<WG第1回>

- （責任関係に関し、）放送事業者が自ら多言語字幕サービスを実施するのではなく、放送外マネージドアプリによりサードパーティがサービス提供主体になる場合も、（利用者からみた場合）放送とは全然別の全くのフリーのサービスとは言えないのではないか。
- （早い段階での実現を目指す場合は）テレビ端末ですべて実施するのではなく、スマートフォンやタブレット端末等を使いながらのサービスも考えるべきではないか。

<WG第2回>

- ハイブリッドキャストを使って多言語字幕をテレビ画面にオーバーレイ表示することになる場合に、多言語字幕も放送局から送出されていると視聴者に受け取られる可能性が一層高まるので、翻訳精度などの正確性の大幅な向上が必要になる。
- まずは対象番組を限定した上で、サードパーティによるスマホやタブレットへの表示について実証実験を行うなどによって検証を積み重ねることからスタートしてはどうか。
- 視聴者から見ると、テレビの映像の上にオーバーレイしてしまうと区別がつかないと思う。字幕は映像とは別の下に字幕ウィンドウに出たほうが、これは放送ではないと視聴者にも明示できていいのではないか。
- 別のタブレットとかスマホでサービスを提供する形態にすれば、放送事業者に対して視聴者がクレームを言うということはないのか。
- 放送局が責任を持たないフィールドであっても、ほんとうにそれが社会に受容されるのかどうかはよくよく考えてみて、実験してみないとわからないことだと思うので、あまり予断を持っていくのはいけないのではないかと思う。

(5) 視聴者へのベストエフォートであることやサービス提供主体等の周知(表示)

<WG第1回>

- 相当数ある言語をすべて正確に翻訳することは不可能。不完全性を前提として、例えばどこかが性能を認証して、事業者毎に精度を売りにすることも考えられるのではないか。
- 技術的に精度を上げていく。ただし、限界はある。それを運用上の仕組みでカバーする、視聴者の理解を得ていくことも進めていく必要があるが、(マネージドアプリとして実施する場合についての)分野毎の最低限の正確性に関するガイドラインも必要ではないか。

<WG第2回>

- 多言語字幕を提供しているサードパーティの名前を常時クレジット表示するなどの形で責任の所在を明確にする必要ではないか。
- IPTVフォーラムとして、字幕と放送とあわせて見ている方に、どういうふうのリテラシーとして納得してもらうかが、今後の普及を考えていく上では非常に重要な検討課題。
- ある程度のクオリティーコントロールのようなものを考える必要があると思う。外側からこのサービスのクオリティーはこれぐらいというのがデジタル化できるような仕組みがあると、ユーザ側としてはどれを使っていいのかわかりやすい。

3 多言語翻訳システム精度向上のための取組の推進

<親回第1回>

- 機械翻訳と音声認識の技術は、日本が非常に進んでいる。（その他の機関と）相乗効果のある研究体制がNICTを先頭にしてできることを期待。
- 研究を進めながら実証実験を進めていく、そして東京五輪までには実際に使われる技術まで育てるということが重要。
- 五輪も控え、さらなる翻訳技術の向上に期待。優先的に実施する分野、言語を決めることが早期の実現につながるのではないか。【再掲】

<WG第1回>

- 自動翻訳の精度を上げるためには、大規模なデータ（対訳コーパス）の蓄積が重要。対訳コーパスを集めることは権利関係で難しいが、コーパスを集め共通して使える仕組みの実現が望まれる。
- 多言語の字幕が付いていったときに、そのデータそのものをアーカイブしていけば、それを使って精度がよくなる。そのデータの利用の仕方について、フェアユースとまでは言わずとも、ある種のガイドラインを決めた方がよい。例えば、新たなコンテンツを翻訳するためにだけ使うということであれば、大規模な対訳の多言語のデータを非常に有効に活用できるのではないか。
- NHKのBSプレミアムで、外国のニュース群を日本語にも訳して、2カ国語で提供している一連の番組があるが、多言語字幕放送の高度化のために、それも役に立てることは、1つの検討の課題。

<WG第2回>

- NICTでは、観光という利用目的に特化した研究を今やっており、言語については、日英中韓というのがある程度の精度を持って使える。領域を増やすことと、精度を上げるというのはコーパスをどれだけ充実させられるかにかかっているので、今後頑張っていきたい。

4 ビジネス化促進、その他多言語字幕サービスの実現・普及に向けた措置

(1) 放送事業者とサードパーティ間の許諾・契約締結の円滑化

<WG第1回>

- (放送外マネージドアプリケーションに関して) 放送局の立場からは、放送が実施される前に、番組関連情報(メタデータ、字幕等の情報)をサードパーティに提供することは、一般的には難しいと考える。

<WG第2回>

- 日本語字幕データの事前提供について、番組編集の自由を確保する観点から、NHKは番組の内容にかかわる情報の第三者への事前提供は原則として行わない。
- NHKは営利を目的とした業務ができないことが放送法に規定されているので、サードパーティのビジネスに協力することはできない。しかし、障害者を支援するNPOなど、営利を目的としない団体であれば、検討は必要であるが、可能性は十分ある。
- 予定していた番組の放送が中止になることがあり、仮にサードパーティが事前に多言語字幕を作成していた場合に、もし損害が生じても、NHKは受信料でその損害を補償したりすることはできない。
- 多言語字幕を制作する第三者への情報提供は、地上基幹放送の特質に深く関わる問題であり、ハードルが高い。
- 日本にいる外国人をハンディキャップだとすると、サービスをNPOがやる場合、NHKのコンテンツを使える可能性はあるのではないか。
- インターネットを経由した字幕利用アプリケーションに関しては、無料で許諾なしの原則フリーで提供した方がよいのではないか。

(2) 多言語翻訳システムの他分野への応用

<親回第1回>

- デジタルサイネージ等やデジタル教科書といった分野との連携活用ができるのではないか。
- 民間の力を使いながら、いかに自由に多言語字幕アプリをハイブリッドキャストのアプリとして使っていくかという視点は重要。
- NICTの新しい技術は日本の新しい産業、高齢社会に向けたイノベーションになるのではないか。

<WG第1回>

- デジタルサイネージやデジタル教科書への多言語翻訳システムの応用の他、日本語字幕にスマートフォンをかざすと自国の言葉が出てくるという、字幕を読み取って音声にするというサービスも十分考えられるのではないか。

5 多言語字幕サービスの実現に向けた実証実験・推進体制

<親回第1回>

- 研究を進めながら実証実験を進めていく、そして東京五輪までには実際に使われる技術まで育てることが重要。【再掲】

<WG第1回>

- テレビ受信機に搭載する必要なフォントの数やどのくらいの遅延がゆるされるかの観点を含め、実証実験（トライアル）を実施していくことが必要ではないか。
- 技術的な課題や制度（運用）的な課題が色々あるが、実証実験をしながら、その結果を踏まえて進めることが必要ではないか。

<WG第2回>

- 技術開発の観点からも、言語や分野を絞っていく必要があるのではないか。言語としては英語から、分野としては災害、避難情報などの安全安心情報からと思うが、いずれにせよ、実証実験などで検証していくことが必要ではないか。【再掲】
- 実証実験などとともに規定を整備して、実際上の障害者団体などが最大限こういう機会を利用できるように制度的にも検討していくべきではないか。
- まずは対象番組を限定した上で、サードパーティによるスマホやタブレットへの表示について実証実験を行うなどによって検証を積み重ねることからスタートしてはどうか。【再掲】

6 全般

<親回第1回>

- ハイブリッドキャスト受信機の普及が課題。
- 字幕という言葉はどうするか。もう少しキャッチフレーズ的にいい言葉がないか、代案も普及の方策として考えられるので検討してほしい。
- インターネットを前提とした社会の中の6つのキーワードを共有していきたい。①グローバル（日本の文化といたものが地球全体にどう貢献できるのか等の視点が重要。）②アクセシビリティ（その確立に貢献したのが長野パラリンピックのウェブサイト。大きな目標をもって世界のためにやっていくということはとても重要。）③ソーシャル（人間が持っている力を結集すること。）④マルチステークホルダー（情報社会では、全員がやるべきことをやることが重要）⑤クラウドコンピューティング ⑥縦書き（日本人は縦書きというのを上手に使う。国際標準の基盤として標準化を取っていくことは重要。）

<WG第2回>

- 2020年の日本の社会にとってどういう意味があるのかということも、このワーキングで検討されて、報告書に盛り込んだほうがよいのではないか。
- さまざまな分野への波及を見通されているということが、システムの研究開発を推進する原動力にもつながるので、こうしたことにも本ワーキングとして目配りをしたほうがよいのではないか。