

スマートテレビ時代における字幕等の在り方に関する検討会
多言語字幕ワーキンググループ（第2回）議事要旨

1. 日時

平成26年2月26日（水）17時00分～19時00分

2. 場所

総務省8階 第一特別会議室

3. 出席者

(1) 構成員

鈴木主査、藤沢主査代理、石戸構成員、伊藤構成員、高橋構成員代理（岡構成員の代理出席）、河原構成員、菊池構成員、黒橋構成員、幸山構成員、関構成員、中村構成員、福井構成員、堀木構成員、本間構成員、益子構成員、安木構成員

(2) 総務省

福岡情報流通行政局長、南大臣官房審議官、長塩地上放送課長、徳光地域メディア室長、野崎放送技術課長、湯本情報通信作品振興課長、荻原研究推進室長、小林地域メディア室長補佐

4. 議事要旨

(1) 事務局説明

事務局から資料2-1に基づき、第1回会合の議論に関する事項について説明が行われた。

(2) プレゼンテーション

福井構成員から資料2-2、幸山構成員から資料2-3、堀木構成員から資料2-4、株式会社インフォシティ 岩浪代表取締役から資料2-5に基づき、それぞれプレゼンテーションが行われた。

(3) 意見交換

○菊池構成員

- ・実証実験などとともに規定を整備して、障害者団体などが、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。

○石戸構成員

- ・サードパーティによる多言語字幕表示については、実証実験を重ねて検証していくことが必要。

- ・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と。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けばよい。

○河原構成員

- ・言語や分野をある程度想定した方がよい。一番想定されるのが、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、韓国語というのは妥当なところではないか。
- ・翻訳システムの入力が音声なのかテキストなのかも、技術的に難しさが全然違う。
- ・テレビの映像の上に字幕をオーバーレイしてしまうと、それは誰が出しているのか区別がつかないので、字幕は映像とは別の下のウィンドウに出るようにした方が、これは放送ではないと視聴者にも明示できてよいのではないかと。

○福井構成員

- ・既にハイブリッドキャストの中では、画面とは別のところに字幕が出るような見方をしている。IPTVフォーラムとして、通信から来たものを放送とあわせて見ている人にどのようにリテラシーとして納得してもらおうかが、今後の普及を考えていく上では非常に重要な検討課題。

○河原構成員

- ・多言語というからには、日本語も英語も中国語も、同時に少なくとも2つぐらいは出た方がよい。日本語と英語とか、日本語と中国語というケースはあると思うが、その辺の仕様はそこまで想定されているのか。

○福井構成員

- ・幾つかのアプリケーションを同時に出すことは今後、バージョン2.0で検討されていくことになる。

○益子構成員

- ・言語の数の見通しについては、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、日英中韓がある程度の精度を持って使えるぐらいまでは来ている。領域を増やすことと、精度を上げることは、コーパスをどれだけ充実させられるかということなので、今後頑張っていきたい。

○中村構成員

- ・日本にいる外国人をハンディキャップだとすると、そういうサービスとしてあり得

るだろうし、NHKのコンテンツも使える可能性はある。民放の場合には、実際に収益を得るということで、実際のようなコンテンツに対して多言語字幕を付ける可能性もあるので、その辺をうまく切り分けて設計していけば、できるのではないかと。

- ・ユーザーとして見ると、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。このサービスのクオリティーはこれぐらいという数字化できる仕組みがあると、ユーザー側としてはどれを使えばよいかわかりやすい。

○堀木構成員

- ・技術の開発や精度の向上をどれくらい見込んでいるのか不明であり、どれくらい間違っているかという点と情報提供者に聞かれても、答えようがない。本体画面へのオーバーレイではなく、例えばL字画面のフィールドやタブレット端末などセカンドスクリーンであれば受容されるのではないかと意見もあるが、情報の受け取り方は人さまさまなので、今ここで予想もできないし、スクリーンを変えれば大丈夫ということについてもコメントできない。
- ・放送はあまねく、ほとんど全数リーチのメディアであり、いろいろな人を相手にしている。仮にその放送局が責任を持たないフィールドであっても、それが社会に受容されるのかどうかはよく考えて、実験してみないとわからないので、予断を持つてはいけないのではないかと。

○幸山構成員

- ・放送を受信した後で、何らかのアプリケーションで字幕を自動翻訳して付けていくことについては、個人の利用でやる分にはよいのか、ある程度許されるのかという取り決めはあり得るかと思うが、放送が出る前に放送事業者以外が内容の部分にアクセスするような話がプレゼンの中であったと思うが、それは何を意味しているのか。

○河原構成員

- ・音声認識や機械翻訳は、ここ数年でよくなっているので、東京オリンピックが開催される2020年まであと6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と。
- ・バラエティー番組は、あまりに話がくだけ過ぎて、スラングもあり難しい。ニュースに限っても、放送事業者の立場からすると、ニュース番組で出てくる全ての発言者に対して字幕を付けようとするのが難しい。アナウンサーに限って行えば、音声認識は95%なので、ほとんどできるのではないかとと思うし、機械翻訳にしても、か

なり定型的な表現が多いので、それほど無理なレベルではない。

- ・発音や内容はかなりクオリティーがコントロールされるので、それほど技術的に絶望的だと思わないし、実際諸外国の研究事例を見ても、90%あるいは95%目標にするということをこの提言に書いても、それほど無茶なことではない。
- ・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ことはあまりないのではないか。

○中村構成員

- ・いろいろなアプリとビジネスが広がるのがポイントで、全部人間がやってもよいし、それが一部自動化されてもよい。自動化の部分が広がればより効率的になるし、コストも下がるので、私も今やっている、医療の現場で自動翻訳と人間の医療通訳者をうまく組み合わせて100%にするような、サービスの組み合わせをうまくやることで、実際のクオリティーは担保できるのではないか。

○伊藤構成員

- ・生放送の字幕は技術的な観点から、日本語の字幕であっても10秒程度の遅れがあるが、さらに翻訳がかかって、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はなかなか難しいのではないか。
- ・手元で映像を遅らせるのも1つの方法で、例えばIPTVの規格でどのくらい遅らせるかを信号として送るということを検討してもらうこともあるかもしれない。
- ・英語とか中国語しかわからない人がテレビを見て、これは何だろうと思ってサイトを開くと、1分前にやっていたものが流れていて、それを見て何をやっているかわかるというような、テレビを見ながらそのコンテンツをその場で理解するのではなく、放送として流れたものを利用して別ところでそれを利用して使うということもあり得るのではないか。

○関構成員

- ・テレビの放送そのものを見るのではなくて、例えば1分後に映像も含めてスマホで再生するサービスもあり得る。番組に入っている字幕から自動翻訳にかけてしまう方法や、多言語の字幕をそのまま映像を見ながらつくっていく方法などのケースが大体決まってくるのではないか。そのイメージも早目につくった方がよいのではないか。

○福井構成員

- ・IPTVフォーラムの検討の中で、録画した映像に対してアプリケーションを乗せるという規格検討を今やっている。例えば外国人がよく泊まるホテルや立ち寄る場所にそういう端末を置く、あるいは一般の市販のテレビにその機能を乗せるといったことも考えられるが、実際に全ての受信機にそれを乗せなければいけないのか考える必要がある。

○藤沢主査代理

- ・放送が「あまねく」広く対象としているのに対して、ハイブリッドキャストの通信経由で端末に向けて行われるサービスは、不特定の少数の人を対象としてもよいし、個人でもよいかもしれない。そういうサービスを通信経由で提供できるというところがメリット。
- ・しかし、その人のためだけの受信端末でよいのではなくて、搭載する機能はなるべく共通化した上で、ネットワーク経由での様々な個別のサービスができるところがこの放送通信連携の基本的な考え方。
- ・例えばフォントや映像音声を遅延させるために必要なキャッシュメモリーなどを多言語字幕サービスのためだけに受信端末に実装するようなことは、問題が出てくる可能性があるので、その辺は考えておく必要がある。

○安木構成員

- ・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。
- ・放送系と通信系の同期を取ることは非常に難しく、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機に持たせるというのは経済的に成り立っていない。
- ・フォントの問題は深刻で、日本語、英語以外のフォントを出そうと思うと、ハードウェアとして積まないといけない。
- ・録画については、スマートフォンやタブレットで見るようなユースケースが想定されるが、日本にいる外国人がどういう使い方をするのかという要件整理のようなものが先にないと、この議論は収束しないのではないかと。

○高橋構成員代理

- ・どのようなユーザー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、サービス設計や端末設計ができない。
- ・ハンディキャップを持った人用に出すのであれば、少し時間は遅れてもニュースだけきちんと届けばよいということもあるし、外国人に対しては、ニュースや天気予報をある程度時間を置かずに見せた方がよいといったユースケースの整理がないと、フローティングしてしまう。
- ・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することになり、同期の問題やユーザーのクレームの問題が気になってしまう。技術の立場から、自動翻訳や音声認識は進歩しているので、それを単に受信機という立場だけから切り捨ててよいものではないと思う。シチュエーションを決めて、何かトライした方が、それが10年後に受信機に乗るとか、そういったところで使いやすくなるのではないかな。

○鈴木主査

- ・ソフトパワーがこの多言語翻訳のとても重要な視点。日本の文化に触れられるようにして、日本に対する親しみや敬愛の念を持ってもらうことが重要。