

国際シンポジウムのための多言語支援システム

中口 孝雄^{1,a)} 大谷 雅之¹ 高崎 俊之^{1,2} 石田 亨¹

受付日 2016年4月10日, 採録日 2016年10月4日

概要: 本研究では, 国際シンポジウムのための多言語支援システムを開発し実際に適用した. 様々な言語話者が参加する国際シンポジウムでは, 翻訳が必要な言語ごとに複数名通訳者を確保する必要があり, 非常に高コストとなる. 機械翻訳は低コストで多言語に展開できるが, 機械翻訳の単純な適用では, 必要な翻訳精度が得られない. そこで我々は, サービスの価値共創の観点から, 発話者と同じ言語話者が入力者となり, 機械翻訳と協働して多言語支援を行う方式を提案し, それを実現するためのインタフェースを備えたシステムを開発した. 開発したシステムを 20 名規模のプレゼンテーションおよび 1,000 名規模の国際シンポジウムでの講演に適用して評価した. その結果, 単純に機械翻訳した場合の品質が平均 Fluency で 1.84, 平均 Adequacy で 2.89 に対し, 講演への適用では同 3.76, 3.85 となることを確認した.

キーワード: CSCW, 多言語コミュニケーション, 機械翻訳, 情報保障

Multi-language Support System for International Symposiums

TAKAO NAKAGUCHI^{1,a)} MASAYUKI OTANI¹ TOSHIYUKI TAKASAKI^{1,2} TORU ISHIDA¹

Received: April 10, 2016, Accepted: October 4, 2016

Abstract: In this research, we developed the multi-Language Support System for international symposium and apply it. When the participant's language passes to a variety, the simultaneous translation becomes very high cost. However the simple application of machine translation doesn't bring enough quality that is needed for international symposium. Thus, we propose the way that persons who speaks in a same language to speaker cooperate with machine translation for supporting multilingualization in the view point of co-creation of value. And we developed the system and applied it to actual presentations. As the result, while the translation quality of simple application of machine translation is 1.84 of fluency and 2.89 of adequacy, we confirm the quality of our method is 3.76 and 3.85.

Keywords: CSCW, multilingual communication, machine translation, information support

1. まえがき

言語の異なる参加者で構成される国際シンポジウムでは, 言語バリアを低減するため, 多言語支援が行われている. 講演における多言語支援としては, 人が発話内容を翻訳する, 同時通訳が行われる. 同時通訳は非常に高度なタスクであり, 特に特徴が異なる言語間の場合, 通訳者への負担が大きくなることが指摘されている [1]. 同時通訳で

は通訳者が (1) 話者の発話の聞き取りと意味理解, (2) 訳文の構築, (3) 訳文の発話を, 話者の発話と並行に行い, 話者が話し終わる前に訳文の発話を開始する. 高度な技術力と集中力が必要であり, 複数の通訳者が交代しながら行うため, 複数の言語に対応する場合, 人的コストが増大する.

機械による多言語支援も試みられている. 通訳者の 3 つのタスクを機械で置き換えることを目的としており, タスクに応じて音声認識, 機械翻訳, 音声合成の研究や, VoiceTra [2] などの, 全タスクを統合した音声翻訳の研究も行われている. 音声認識では, 人の発話音声进行を解析し, テキストを生成する. ノイズや話者の声の特性などに影響を受けるが, 適切に録音環境の整備や認識モデルの最適化

¹ 京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University, Kyoto
606-8501, Japan

² 特定非営利活動法人パンゲア
NPO Pangaea, Kyoto 600-8035, Japan

^{a)} nakaguchi@i.kyoto-u.ac.jp

を行った場合、80~95%の認識精度を得られる [3]。機械翻訳は多言語に対応したものが入手可能であり、Google 翻訳*1や Bing 翻訳*2などインターネット上に公開されているものも存在する。様々な言語サービスを集積し提供している言語サービス基盤である言語グリッド [4] では、7つの翻訳サービスが提供されており、翻訳可能な言語の組合せは 2,980 言語対である。しかし人の翻訳精度には及ばないため、精度が必要な場合はドメインを限定して利用される。

人による同時通訳は精度が高いが、コストが高くなる。機械による通訳は低コストで多言語に展開しやすいが、精度が低い。このトレードオフにより、異文化コラボレーション活動における多言語支援を容易に拡大させることが難しくなっている。本研究では、価値共創 [5] の観点から、人と翻訳サービスが協働し、低コストで高品質でかつ多言語への展開が容易な多言語支援方式を提案する。具体的には、人が発話の聞き取りおよび翻訳前編集と入力を行い、機械翻訳が翻訳を行う。さらに実際にシステムを実装して実用に供し、結果を評価する。以下本稿では、本方式の課題を述べ、システムの実装とフィールドへの適用、および結果の評価と考察を述べ、最後に結論を述べる。

2. 多言語支援システムの課題

通訳者のタスクを置き換える際に、まず最初に問題となるのが、発話を解析して機械処理可能なデータに変換する音声認識である。音声認識の精度が低ければ機械翻訳に正常なデータが入力できず、翻訳の品質が低下する。そのため本稿では、音声認識タスクを人間で置き換えることを考える。人間による音声のテキスト化の手法としては、主に聴覚障害者への情報保障のために用いられている、PC テイク [6] があげられる。PC テイクでは、パソコンを利用して複数人で協調して入力することで、発話をほぼそのままテキスト化できる。各入力者の担当部分は通常文字数で決まり、また入力が発話に追いつけなくなった場合にも文の途中で担当部分に変更される。そのため個々の入力は文の構造とは無関係に区切られることになり、そもそも翻訳に適さない。また個々の入力を連結して翻訳したとしても、発話文は長く複雑な構造になりがちであり、翻訳品質の低下を招く。このことを確認するため、予備実験として日本語の講演ビデオに対して PC テイクを実施し、その結果に機械翻訳を適用して、翻訳品質を評価した。PC テイク用のソフトウェアとしては、IPTalk [7]*3を利用した。翻訳は、言語グリッド上で提供されている JServer*4を用いて、英語と中国語に対して行った。結果を表 1 に示す。

表 1 は、入力を連結して文を復元して翻訳し、文献 [8]

表 1 PC テイク結果を翻訳した場合の翻訳品質

Table 1 Translation quality of speech dictation by PC-take.

文作成方法	文数	WA	文意保持度合	翻訳先言語	平均 Flu.	平均 Ade.
PC テイク	22	79.98	4.8	中国語	1.77	2.68
				英語	1.91	3.09

に基づいて流暢さ (Fluency) と適切さ (Adequacy) を評価したものである。Fluency, Adequacy, および後述する文意の保持度合いは最高点が 5、最低点が 1 であり、これは以降の評価においても同様である。また発話文の再現度を示すため、Word Accuracy (以下 WA) と文意の保持度合いを計測した。WA では、音声を別途書き起こしたものと PC テイク結果に対して Mecab*5により形態素解析を行い、手動で単語間の対応をとり、次式を用いて計算した。

$$WA = 1 - \frac{S + D + I}{N}$$

N は書き起こしたテキスト内の単語数、S は置換が行われた単語数、D は削除された単語数、I は挿入された単語数である。

PC テイクでは発話をそのまま入力するが、言い間違いや冗長な部分が削除されたり、短く置き換えられたり、また読みやすさのための挿入が行われる。表 1 の入力では、“どうもありがとうございます”の“どうも”の削除、“わけですけども”の“わけですが”による置き換え、“ご覧なられた”への“ご覧になられた”という“に”の挿入などが見られた。各変更の数は、置換が 56、削除が 126、挿入が 9 である。後述する翻訳前編集後の WA の計算では、置換が 141、削除が 363、挿入が 45 発生していた。

文意保持度合いでは、発話の録音から書き起こした文章と PC テイクの入力結果を比較し、文意が保持されているかどうかを日本語ネイティブの評価者により評価した。評価結果は機械翻訳の単純な適用では良い品質は得られないことを示している。PC テイクの結果は機械翻訳に適しておらず、翻訳品質を改善するには、テキストを機械翻訳に適した形に変形する、翻訳前編集が必要である。

また本稿では、提案手法とそれを実装した多言語支援システムを、実際に多言語支援を必要としている現場での実用に供することを目的としている。システムは提案手法による入力を支援するユーザ・インタフェース (以降 UI) を備える必要があり、具体的には、協調入力のために他者の入力状況を確認するタイプ・モニタ機能、これまでの入力内容を確認するためのログ表示機能、翻訳前編集時に翻訳品質を確認するための折返し翻訳機能を想定する。

すなわち本稿で取り組む課題は、以下のものである。

- 人と機械が協働する方式による翻訳品質改善
- 翻訳前編集と協調入力を支援する UI の実現

*5 <http://taku910.github.io/mecab/>

*1 <https://translate.google.co.jp/>

*2 <http://www.bing.com/translator/>

*3 http://www.geocities.jp/shigeaki_kurita/

*4 <http://www.kodensha.jp/platform/rent/>

3. 発話内容の翻訳前編集

機械翻訳の翻訳品質を改善するために、翻訳前編集が有効であることが知られている。稲葉ら [9] は日本の農業専門家とベトナム人との機械翻訳を介したコミュニケーションの改善のために、I. 日本語表記上の書き換えを行う、II. 曖昧な表現を避ける、III. より英語にしやすい表現にする、の3つに分類される12のルールを導入している。I. ではなく漢字を使う、文章を短く切るなど、II. でも適切な助詞を使う、係り受けを明確にするなどの、翻訳先言語の知識を必要としないルールがあげられている。III. では、分類としては英語にしやすいという表現が使われているが、内容は複合名詞を避ける、無理な造語を避ける、無生物主語を使うなど、これも英語に対する知識は必要としない。

前章のPCテイク結果に対して翻訳前編集を行った結果を表2に、また表3に編集例の抜粋として、翻訳前の日本語原文および機械翻訳による翻訳結果と、翻訳前編集適用後の日本語文と翻訳結果を2例示す。2例とも、編集前の日本語文の状態でも意味はとれるが、それを翻訳したものは良い品質ではない。このように、人間にとって意味がとれても、機械翻訳にとって翻訳しにくい文章が存在し、容易に品質低下を招く。最初の例の編集では、語順を入れ替えて意味を明確にし、2番目の例では字面はほとんど同

表2 PCテイクに翻訳前編集を適用した際の翻訳品質
Table 2 Translation quality of pre-edited PC take.

文作成方法	文数	WA	文意保持度合	翻訳先言語	平均 Flu.	平均 Ade.
翻訳前編集	40	42.63	4	中国語	3.13	3.70
				英語	3.18	3.25

表3 翻訳前編集と翻訳の例
Table 3 Example of pre-edit and translation.

	日本語文	英文 (翻訳結果)
編集前	科学技術、文化が高いレベルにいる日本が、人類に対してどのようなことができるのか。	What kind of thing can do Japan where science and technology and culture are necessary for the high level to the mankind?
	日本の科学技術と文化のレベルは高いですが、人類に対して何ができるのでしょうか。	Japanese science and technology and culture are ranked high, but what is made to the mankind?
編集前	今、67億人の10億人は飢餓にも瀕している。	6,700,000,000 1,000,000,000 people are also on the verge of hunger now.
	今、67億人中の10億人は、飢餓に瀕している。	1,000,000,000 people out of 6,700,000,000 people are on the verge of hunger now.

じだが、2文字(“中”, “、”)の挿入と1字(“も”)の削除によって、翻訳結果が改善している。このように翻訳前編集では、単に発話文を単純化するだけでなく、機械翻訳の入力として適した文に変更して入力を行い、それにより翻訳品質を向上させる。またその際、WAのような、単語の並びや同一性を反映した評価は大きく下がり、ただし文意は保たれるはずである。実際表2では、表1に比べて、WAは大きく下がり、文意の保持度合いの低下は限定的であり、かつ翻訳品質が向上することが確認された。

よって本稿では、PCテイクと翻訳前編集の組合せによる協調翻訳前編集を採用する。これにより発話が聴衆へ届けられるまでのプロセスは、(a) 話者の発話の聞き取りと意味理解、(b) 担当部分の決定、(c) 翻訳前編集、(d) 機械による翻訳、(e) 翻訳結果の表示、となる。1章に示した同時通訳のタスクと比較すると、(1)と(a)が対応し、ただし(1)での意味理解がその後の翻訳を前提とするものであるのに対し、(a)での意味理解は(b)担当部分の決定と(c)翻訳前編集を前提としたものになる。(b)および(c)は独自のタスクであり、(2)と(d)が、(3)と(e)が対応する。同時通訳においては(1)~(3)を言語ごとの通訳者がそれぞれ独立して行うが、提案手法では(a)~(c)のタスクは入力者達が協調して行い、(d)以降は言語ごとの機械翻訳器や本システムのUIが行う。

4. 多言語支援システムの構成と実装

本システムでは、複数の入力者による協調翻訳前編集結果に対し、翻訳・表示を行う機能を実現する。そのために、協調翻訳前編集入力を行うための入力画面と、翻訳結果を表示するための字幕表示画面を備える。翻訳は言語グリッドの翻訳サービスインタフェースに準拠した翻訳サービスを用いる。図1に、システムの構成を示す。

入力者は図中の入力用インタフェースという枠で示したコンポーネントのどちらかを用いて入力を行う。連携入力ソフト (IPTalk) はPCテイクで一般的に利用されているWindows用ソフトウェアであり、協調入力のためのタイプモニタ機能と入力ログの表示機能を備えている。一方本システムの入力画面では、それらの機能に加え折返し翻訳機能を備えている。入力画面と字幕表示画面はWebブラウザ上で動作する。

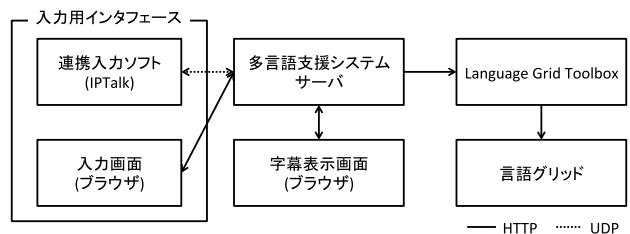


図1 システム構成

Fig. 1 System structure.

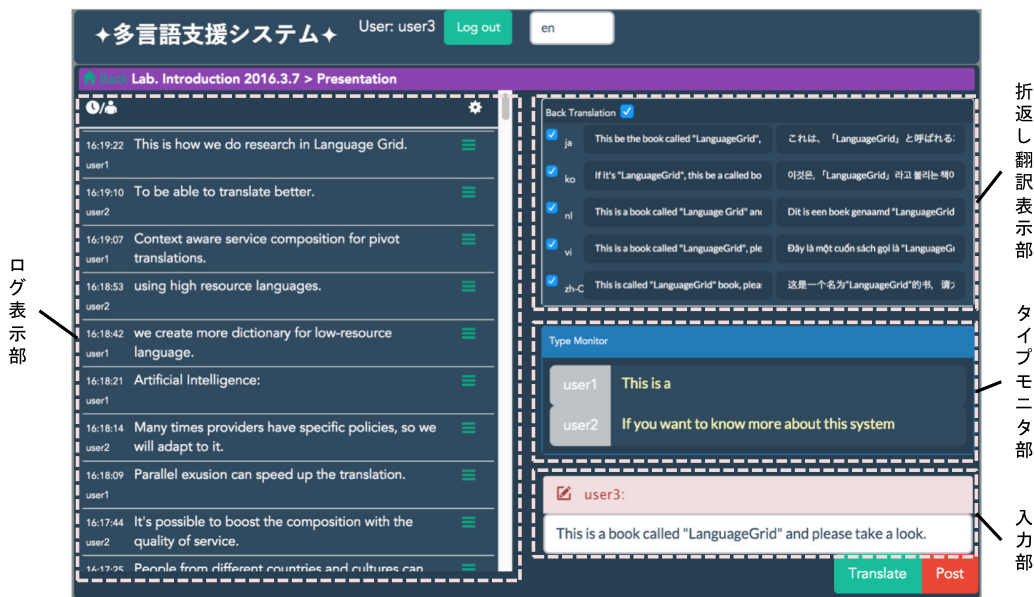


図 2 入力画面
Fig. 2 Input screen.

Lab. Introduction 2016.3.7	
また、言語グリッドに多くの 他の学術研究があります。	또, 언어 그리드에 많은 다른 학술연 구가 있습니다.
これは私達の大型プロジェク トのうちの1つです。	이것은 우리들의 대형 프로젝트 가 운데 1개입니다.
これは、どのようにベトナム語 を学ばずにエキスパートが子供 を助けることができますかです。 있을지입니다.	이것은, 어떻게 베트남어를 배우지 않고 엑스퍼트가 어린이를 도울 수 있을지입니다.
子供はここからRecipeCardを 見ることができます。	어린이는 여기에서 RecipeCard를 볼 수 있습니다.

図 3 字幕表示画面
Fig. 3 Translation screen.

本システムと各画面の間は、画面表示時には HTTP で通信し、その後のメッセージ送信や翻訳結果の受信には WebSocket *6 を使用している。これにより、入力内容や翻訳結果を画面間で即時に送受信している。IPTalk と本システムは、UDP ベースのプロトコルで通信する。IPTalk は 7 台まで接続可能であり、各画面は接続数に制限はない。多言語支援の実施者は、入力者の数に応じて IPTalk や入力画面を追加し、翻訳・表示する言語数に応じて字幕表示画面を追加できる。翻訳サービスには、Language Grid Toolbox [10] を用いる。これにより言語グリッドが対応している複数の翻訳サービスや辞書サービスを、発話言語や対象分野に応じて組み合わせて使うことができる。

図 2 に示した本システムの入力画面は、ログ表示部、折返し翻訳表示部、タイプモニタ部、入力部を備える。入力者はタイプモニタによって他者の入力内容の確認を行いつつ、自身が入力する文を決定し、入力する。また、ログの

閲覧や折返し翻訳機能を利用し、これまでの入力状況の確認や、翻訳結果の確認を行うことができる。

字幕表示画面を図 3 に示す。画面内の内容は、後述する研究室室内プレゼンテーションで実際に表示されたものである。字幕表示画面は翻訳結果を提示することを目的としており、画面のほとんどを翻訳結果が占めている。字幕表示画面には入力者が入力したテキストを表示することも可能であり、その場合、図中“(translated)”と表示されている部分は“(original)”と表示される。

5. 多言語支援の実施

多言語支援システムの入力方式の実現性の検証、および問題点の分析を行うため、2 種類の講演に対して実際に適用した。一方は国際シンポジウムであり、発話の言語は日本語である。もう一方は研究室室内プレゼンテーションであり、発話の言語は英語である。

5.1 多言語支援 1：国際シンポジウム

国際シンポジウムでの日本語講演に対して多言語支援を実施した。入力者は 2 名（入力者 A、入力者 B）であり、母語は日本語である。このシンポジウムでは、可能な限り多くの発話内容を入力すること、かつ翻訳品質を確保することが求められたため、日本語データ入力業務の経験がある者を募集し（ただし PC テイクの経験はなかった）、事前にトレーニングを行った。その際、最初に発話を聞き取って翻訳前編集して入力し、翻訳結果をチェックする作業を 1 時間 30 分の間、3 回繰り返した。この際、稲葉ら [9] の書き換え基本ルールを入力者へ提示した。その後他の入力者と協調しつつ、リアルタイムに翻訳前編集入力と結果の

*6 <https://tools.ietf.org/html/rfc6455>



図 4 シンポジウム会場の様子

Fig. 4 The look of conference hall.

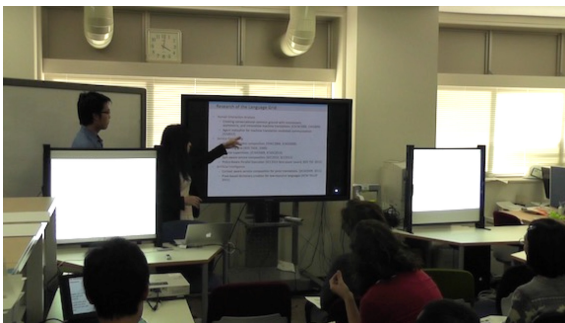


図 5 プレゼンテーションの様子

Fig. 5 The look of presentation.

チェックを行う作業を5時間の間、5回繰り返した。トレーニングを通じて入力された文数は、入力者 A が 495 文、入力者 B が 481 文である。

実際に多言語支援を実施したのは、ワイズメンズクラブ国際協会第 26 回アジア地域大会である。アジアの各地から合計 962 名が参加し、3 日間行われた。図 4 に会場の様子を示す。日本語での講演、1 時間 20 分に対して、多言語支援が実施された。会場にはスクリーンが 3 つあり、中央のスクリーンはつねにライブ映像または登壇者が用意したスライドが、左に日本語と英語が、右に日本語と中国語が表示された。多言語支援を通じて 738 文が入力、翻訳され、表示された。

マシン構成は、本システムのサーバ機能を稼働させるマシン 1 台 (MacBook) と、字幕表示画面用マシン 2 台 (Windows 機)、加えて IPTalk 用マシン 2 台 (Windows 機) で、すべてノート PC を用いた。入力に IPTalk を用いたのは、入力者が IPTalk に慣れていたのである。

5.2 多言語支援 2：研究室内プレゼンテーション

24 名が参加する約 30 分間のプレゼンテーションでの多言語支援を行った。聴衆は 20 名であり、話者の言語は英語、入力者は 2 名 (入力者 C, 入力者 D) である。会場に字幕用スクリーンを 2 つ設置し、一方にベトナム語、中国語、オランダ語を、もう一方に日本語と韓国語を表示した。図 5 に、会場の様子を示す。

表 4 多言語支援における入力状況

Table 4 Status of input of multilingual support.

項目	実験 1 (日)		実験 2 (英)	
	A	B	C	D
入力文数	290	448	94	76
平均文字数	12.8	10.8		
平均単語数			6.68	7.7
入力誤り (%)	4 (1.38)	1 (0.22)	4 (4.26)	2 (2.63)
重複 (%)	7 (2.41)	11 (2.46)	0 (0)	1 (1.32)
連続入力 (%)	32 (11.03)	191 (42.63)	36 (38.3)	19 (25)

表 5 入力者アンケート (1: 最悪-5: 最良)

Table 5 The questionnaire of inputter (1: worst-5: best).

設問	入力者		実験 2 (英)	
	A	B	C	D
入力画面は見やすかったですか?			4	4
折返し翻訳機能は役に立ちましたか?			4	4
タイポモニタ機能は役に立ちましたか?			5	5
入力機能は使いやすかったですか?			4	3
入力ツール (IPTalk) は使いやすかったですか?	4	4		
もう一人の入力者とはうまく連携できましたか?	5	4	4	5
スピーチ内容は全て入力できましたか?	4	4	3	3

マシン構成は、サーバ用 1 台 (MacBook) と入力者用 2 台 (MacBook)、字幕表示用 2 台 (Windows 機) である。入力は本システムの入力画面を用いた。入力者は両名とも PC テイクの経験はなく、事前に 2 時間 (単独入力 30 分、協調入力 1 時間 30 分) の練習を行った。多言語支援を通じて計 170 文が入力、翻訳され、表示された。

5.3 結果

多言語支援を通じて行われた入力の内訳を、表 4 に示す。多言語支援 1 (表中実験 1) では英語で、多言語支援 2 (表中実験 2) では日本語で入力が行われた。平均単語数は、入力された英語文の単語数であり、平均文字数は日本語の文字数である。入力誤りは、“j これは何でしょうか” や “shrae” (正しくは “share”) など、明らかな入力ミスを含む文の数である。重複は、ある入力者が “from germany,” と入力したのに続けて別の入力者が “From Germany and Denmark.” と入力するなど、直前の入力と重複する内容を含む文の数である。連続入力は、一方の入力者が、他方の入力を待たずに連続して入力した文の数である。

実験終了後に、両方の多言語支援の入力者と、多言語支援 2 の聴衆に対してアンケートを実施した。表 5 と表 6 に入力者のアンケート結果と感想を示す。表 5 では、入力

表 6 入力者の感想 (抜粋)

Table 6 Comments from inputters (extraction).

実験 1 (日)
<ul style="list-style-type: none"> 他の入力者の入力状況を見ることが出来るのは便利でした。 同じ部分を入力していることに気がついても、それを最終的に確定してその入力文が表示されるか/されないかまでは全くわからない。最終的にどちらも諦めた結果、その部分が表示されないということも結構ありました。 大体のところ (80%ぐらい) は入力できたとは思いますが、早口のところや、例示で色々出していたところなどはやはり抜けはあったかと思えます。
実験 2 (英)
<ul style="list-style-type: none"> 後半の講演者のスピーチは間を取って比較的ゆっくり話していたためスピーチ内容を前半の講演者の時より入力できたと思えます。 スピーチ部分とスライド説明部分が混ざっていたため、スライドについての記述をするかを迷う場面がありました。 ポストするには Ctrl + Enter を押さなければならぬので、時間がかかりすぎました。

表 7 多言語支援 2 の聴衆アンケート (1: 最悪-5: 最良)

Table 7 The questionnaire of audience (exp.2)

(1: worst-5: best).

設問	評価毎の人数				
	1	2	3	4	5
翻訳画面は見やすかったですか?	1	3	2	7	3
翻訳の表示タイミングは適切でしたか?	0	6	7	3	0
翻訳結果は分かりやすかったですか?	1	9	5	1	0
字幕は内容の理解に役立ちましたか?	0	5	5	5	1
今後もこのシステムを利用したいと思えますか?	0	2	6	5	3

表 8 多言語支援 2 の聴衆の感想 (抜粋)

Table 8 Comments from audience (exp.2) (extraction).

<ul style="list-style-type: none"> 翻訳画面のモニタは十分明るく、文字も一番後ろで立っていてもはっきり読めました。 翻訳を出す画面とプレゼンで示す画面が異なっており、視線の移動が非効率であった 少し遅いが適切だったと思えます。 内容のパーツをひろいながら見ていくことができたので理解には役立ちました。 There are some mis-translation. I think the original language should also be shown. 英語だけでは理解しにくい部分を日本語や韓国語と一緒に見えるので理解に役立ちました。 Sometimes if I try to read the translation, I miss listening.

ツールの機能や使い勝手に関するものと、連携について、およびカバー率について 5 段階評価を実施し、各入力者の評価値を示している。入力ツールに関しては、実験 2 では本システムの画面を使用し、画面内の各機能についても設

表 9 入力者の翻訳結果の品質

Table 9 The quality of translations of inputs.

文作成方法	文数	翻訳先言語	平均 Flu.	平均 Ade.
協調翻訳前編集	50	中国語	3.58	3.82
		英語	3.94	3.88

問を設けている。表 7 と表 8 に聴衆のアンケート結果と感想を示す。表 7 では、5 つの設問に関して 5 段階評価を実施し、評価ごとの人数を集計した。聴衆 20 名のうち 17 名からアンケートの提出を受け、うち 1 名は前に座った人で翻訳画面が隠れたことにより回答が不完全であったため結果から除外している。表 9 に、多言語支援 1 で入力された文から無作為に 50 文を抽出し、翻訳品質を評価した結果を示す。

6. 考察

6.1 入力方式

本システムにおける入力は、講演者の発話を複数の入力者が協調し、機械翻訳に適した文章を交互に入力することを想定しており、そのためのトレーニングを行ってから多言語支援実験を実施している。しかし、表 4 に示したとおり、全入力者で連続入力があり、特に入力者 B と C は全体の 4 割程度にのぼった。入力文数は B と C の方が A, D よりも多く、入力文字数は逆に A, D の方が多かった。これは練習での入力方法の提示が十分でなかったか、あるいは発話内容や入力者の適正により翻訳前編集に必要な時間に差が生じた可能性も考えられる。

入力者へのアンケート結果では、入力者 D の入力機能へのものを除いて 4 以上の評価が得られた。入力者 D に対して追加インタビューを行ったところ、キーアサインのカスタマイズにより Ctrl+Enter が押しにくいことが原因だった。もう一方の入力者は問題なかったと回答しており、本質的な問題はないと考える。

本方式では参加者が入力者となり、機械翻訳と協働して多言語支援を実施することを想定しているため、入力者の練習時間は短いほうが良い。しかし文字数などの分担基準を設けやすい PC テイクと異なり、複数名での協調翻訳前編集入力では、他者が発話のどこを担当するかを予測する必要があるため、ある程度の練習は必要である。多言語支援 1 では 6 時間 30 分、多言語支援 2 では 3 時間の練習時間を設けた。表 9 では多言語支援 1 において本方式による品質の一定の向上が示されたが、表 7 では多言語支援 2 の聴衆の翻訳結果に対する評価は低かった。両実験では入力言語と翻訳先言語が異なり、発話内容も異なるため単純な比較はできないが、少なくとも 6 時間のトレーニングを行うことで一定の品質を達成できたことは示され、同時通訳者ほどの高い技能や長い習得時間は必要ないと思われる。ただし多言語支援 1 の入力者においても戸惑いは見られ、よ

り練習時間を短縮し、かつ入力効率を上げるためには、入力を支援するための機能やルール of 改善を検討する必要がある。

6.2 内容理解

表 7 と表 8 は多言語支援 2 の聴衆へ行ったアンケートの結果である。画面に関しては、画面が大きい、文字がくっきりとして見やすいことがあげられたが、反対に画面が小さいという意見もあり、評価がばらけた。内容の理解についても、評価 2, 3, 4 がともに 5 人ずつ、5 が 1 人という結果だった。プレゼン画面と字幕画面が異なっており、視線の移動が非効率だった、少し遅いが適切だったという意見もあった。本システムでは発話を翻訳して表示するが、講演者のジェスチャやスライド、映像など発話以外の情報とタイミングがずれると内容理解に影響が出ると考えられる。同時通訳における話者の発話から通訳者の訳出までの遅延時間は、高木ら [11] の報告では日英で平均 4.13 秒、英日で平均 3.19 秒である。提案手法は人間によるタイプ入力であること、および翻訳可能な長さまで入力してから翻訳する必要があること、さらに翻訳処理時間も必要（予備実験における平均翻訳時間は 269 ミリ秒である）なことから、同時通訳の速度を超えることは難しく、可能な限り入力者の負担を軽減し、理解の妨げとならないように遅延を改善するための検討が必要である。

6.3 翻訳品質

表 1 と表 9 を比較すると、PC テイクに単純に機械翻訳を適用した場合の平均 Fluency は 1.84, 2.89 に対し、提案手法を適用した場合は 3.76, 3.85 となることを確認した。両者とも日本語を英語と中国語に翻訳しており、言語別では日中で同 3.58, 3.82, 日英で同 3.94, 3.88 であった。両者では元となる発話が異なるため単純な比較はできないが、話者は同一人物であることから、本方式によって一定の翻訳品質が実現されたことを示していると考えられる。

7. 関連研究

機械翻訳による多言語支援では、人間が行う通訳タスクを機械によって置き換える。VoiceTra [2] のような音声を直接翻訳する試みでは、タスク全体を機械で置き換える。この場合音声認識により発話内容が機械処理可能なデータに変換され、次に翻訳が、最後に音声合成が行われる。しかし人間の自然な発話には言い間違いや言いよどみが含まれているため、話者の音声をそのまま音声翻訳に用いると認識率が低下し、結果翻訳間違いも起こりやすくなる。そのため収録環境の限定や、認識機能で用いるモデルを事前に学習させるなどの最適化を行う必要がある。提案方式では人間が聞き取りを行うため認識機能の最適化は必要ないが、入力者が発話を聞き取りやすい環境は必要である。

近年、短文を即時に多人数へ配信できる Twitter などのサービスに翻訳機能が追加されることにより、国際シンポジウムなどの講演の場で発話内容の概要が配信され、またそれが翻訳されることも増えている。ただし要約筆記や PC テイクを想定したサービスではなく、協調入力のためのインターフェースも持たないため、1 人で入力を行うことが基本となり、要点の抜き出しが前提となる。この作業は、i. 話者の発話の聞き取りと意味理解、ii. 要点の抜き出し、iii. 入力に整理できるが、提案方式では 1 文単位の処理が基本となるのに対し、i. と ii. は複数文にまたがって行われ、大きく情報量が落ちるため、本稿の目的である通訳タスクの置き換えには適さない。

Aiken ら [12] は、参加者が文章を入力し送受信する電子的な会議において、機械翻訳を用いて多言語コミュニケーション支援を試みている。この方式では、参加者は折返し翻訳結果を確認しつつ文を調整して入力することにより、翻訳品質が保たれている。ただし参加者が入力する必要があり、話者が発話のみを行う形態には適用できない。

PC テイク [6] では、人間が発話を聞き取り入力するため、音声認識のような認識誤りは起こりにくい。入力者は話者の発話を可能な限り正確に入力するが、本方式では、翻訳前編集を行い、文意を保ったまま機械翻訳に適した形へ変形して入力するところが異なっている。

PC テイクは要約筆記を出発点としており、近年は全文の入力を目的とすることが多いが、依然として要約を前提とする場合もある。要約における発話文の変更方法に関しては福島ら [13] に詳しいが、主として話者の発話に追随するために、入力者が入力可能な文字数にまで発話文を短縮する。その際、語句や文節の削除、言い換えなどによる簡略化が行われる。一方翻訳前編集では、機械翻訳品質の向上を目的とした修正を行う。複雑な文章よりも簡潔な文章の方が品質が上がる傾向にあるため、要約でもある程度の効果は考えられるが、3 章で触れたように、稲葉ら [9] は 3 つの分類の翻訳前編集ルールを提案しており、分類 II 内の係り受けの明確化が最も翻訳品質の向上に寄与したと報告している。また表 3 の 2 番目に示した、文字数がほぼ変わらない修正でも翻訳品質が向上することが確認されており、要約とは違う観点での編集も必要となる。

8. あとがき

本稿では、低コストで高品質な多言語支援を実現するため、以下の課題に取り組んだ。

- 人と機械が協働する方式による翻訳品質改善
- 翻訳前編集と協調入力を支援する UI の実現

人と機械が共同する方式では、情報保障のために行われている協調入力に翻訳前編集を組み合わせ、その結果に対して機械翻訳を適用することで、発話内容の機械翻訳において、文意を保ったまま翻訳品質を向上させた。また実際

に行われた国際シンポジウムでの実用に供し、翻訳品質において、日中で Fluency 3.58, Adequacy 3.82, 日英で同 3.94, 3.88 というスコアを得られることを示した。同一人物の別の講演に対して PC テイクを行い単純に機械翻訳を適用した表 1 では日中で同 1.77, 2.68, 日英で同 1.91, 3.09 であり、本方式による一定の効果を示している。

UI の実現では、ログ表示機能、タイプモニタ機能、折返し翻訳確認機能を備える入力画面と、翻訳結果の表示画面を備えた多言語支援システムを開発した。実際に研究室室内プレゼンテーションにおいて英語を 5 つの言語に、国際シンポジウムにおいて日本語を英語と中国語に翻訳する多言語支援を実施し、利用者の評価を得て、実装した UI が一定の実用性を持つことを示した。

謝辞 本研究に多大な協力をいただいたワイズメンズクラブ国際協会アジア地域の方々およびアジア地域会長岡野様に感謝いたします。本研究は JSPS 基盤研究 (S) 24220002 の助成を受けたものです。

参考文献

- [1] 成田 一：翻訳と通訳の脳内処理メカニズム（機械翻訳技術の向上），*Japio year book*, pp.294-302 (2015).
- [2] Matsuda, S. et al.: Multilingual speech-to-speech translation system: VoiceTra, *14th International Conference on Mobile Data Management (MDM)*, Vol.2, IEEE (2013).
- [3] 河原達也：ICT・音声認識の活用による講演・講義の字幕付与，*情報処理*, Vol.56, No 6, pp.543-546 (2015).
- [4] Ishida, T. (Ed.): *The Language Grid: Service-Oriented Collective Intelligence for Language Resource Interoperability*, Springer. ISBN 978-3-642-21177-5 (2011).
- [5] Lusch, R.P. and Vargo, S.L. (Eds.): *The service dominant logic of marketing: Dialog, debate and directions*, Armonk, NY: M.I. Sharpe (2006)
- [6] 吉川あゆみ，太田晴康，白澤麻弓：大学ノートテイク入門，人間社 (2001).
- [7] 栗田茂明，河野純大，近藤恵子：運用コスト低減を目指した遠隔パソコン文字通訳システム，*ヒューマンインタフェース学会研究報告集*, 15(8), SIG-ACI-10, pp.13-20 (2013).
- [8] Linguistic Data Annotation Specification: Assessment of Fluency and Adequacy in Arabic-English and Chinese-English Translations (2002).
- [9] 稲葉利江子，北川大輔，菱山玲子：翻訳前編集パターンの分析：ベトナム農業支援プロジェクトを事例として（言語グリッドと異文化コラボレーション），*電子情報通信学会技術研究報告*, AI, 人工知能と知識処理, 112.435, pp.109-114 (2013).
- [10] Tanaka, M. et al.: Language Grid Toolbox: Open source multi-language community site, *4th International Conference on Universal Communication Symposium (IUCS)*, IEEE (2010).
- [11] 高木 亮，松原茂樹，稲垣康善：同時通訳コーパスを用いた通訳者発声タイミングの分析，*言語処理学会第 8 回年次大会発表論文集*, pp.383-386 (2002).
- [12] Aiken, M. and Vanjani, M.: Polyglot: A multilingual group support system, *Issues in Information Systems*, Vol.10, No.2, pp.101-106 (2009).
- [13] 福島孝博，和田裕二，浦谷則好：講演の要約筆記における

要約の手法，*電子情報通信学会技術研究報告*. WIT, 福祉情報工学, Vol.101, No.264, pp.55-59 (2001).



中口 孝雄 (正会員)

1996 年京都コンピュータ学院鴨川校卒業。ATR, アントラッド, NTT-AT を経て、現在、京都大学大学院情報学研究科特定研究員。2006 年京都情報大学院大学応用情報技術研究科修了。情報技術修士（専門職）。サービスコンピューティングと異文化コラボレーションの研究に従事。



大谷 雅之

2009 年電気通信大学電気通信学部卒業。2013 年同大学大学院電気通信学研究科博士課程修了。博士（工学）。2013 年 4 月より京都大学大学院情報学研究科特定研究員。ソフトウェアエージェント、異文化コラボレーションの研究に従事。



高崎 俊之

1999 年東京大学工学部精密機械工学科卒業。2001 年同大学大学院新領域創成科学研究科修士課程修了。修士（環境学）。特定非営利活動法人パンゲア副理事長。現在、異文化コラボレーションのシステム開発とフィールド実践研究に従事。



石田 亨 (正会員)

1978 年京都大学大学院修士課程修了。同年日本電信電話公社電気通信研究所入所。1993 年より京都大学教授。現在、情報学研究科社会情報学専攻。この間、ミュンヘン工科大学、パリ第六大学、メリーランド大学、上海交通大学、清華大学客員教授等を経験。工学博士。情報処理学会、電子情報通信学会、IEEE フェロー。人工知能、特にマルチエージェントシステムを研究。デジタルシティ、言語グリッド等情報技術と社会をつなぐ研究プロジェクトを推進している。本会フェロー。