

NICTにおける 自動『同時通訳技術』のご紹介

AAMT年次大会
2022年12月1日

情報通信研究機構
内山将夫

総務省委託研究開発

「多言語翻訳技術の高度化に関する研究開発」に関して
NICTの機械翻訳研究開発にフォーカスしてご紹介

- ① 背景・アウトカム目標・推進体制・実証実験などの概要
- ② 自動同時通訳における機械翻訳(同時翻訳)の適用法
- ③ 同時翻訳を実施するために開発するコーパス
- ④ 同時翻訳のこれまでの成果の概略
- ⑤ 今後の課題

謝辞:

本件は、総務省の「ICT重点技術の研究開発プロジェクト(JPMI00316)」における「多言語翻訳技術の高度化に関する研究開発」による委託を受けて実施した研究開発による成果である。

総務省委託研究開発

「多言語翻訳技術の高度化に関する研究開発」

概要

1. 背景

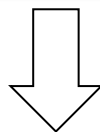
「グローバルコミュニケーション計画」(平成26年)

<成果>

逐次翻訳精度の向上 / 自治体・医療・交通・観光等の様々な分野で活用

<今後の展望>

逐次対話方式に一定の成果が得られたが、今後の更なる利活用拡大のために、発話タイミングに寄らない方式の検討が必要



「グローバルコミュニケーション計画2025」(令和2年)へ

従来のNICTの多言語自動翻訳技術の更なる高度化により、従来の短文への逐次翻訳から、会話の文脈（文章の流れ）や話者の意図も補完した「**同時通訳**」へと飛躍させるとともに、開発成果を通訳サービスの提供又は利活用の主体となる企業等が共通的に利用可能となる仕組みを構築し、社会実装することを目指す。

※ 総務省 多言語翻訳技術の高度化に関する研究開発 基本計画書より抜粋

2. 2025年における社会実装のイメージ

- ・ 実証結果に基づくアジャイルな研究開発と普及活動の両軸を推進し、社会実装を目指す。
- ・ 2025年の大阪・関西万博をショーケースとして、日本の自動同時通訳技術を世界に発信。

令和2年 ~ 令和7年（2025年）

※出典：グローバルコミュニケーション計画2025（2020年3月 総務省）

本委託研究開発

同時通訳技術の社会実装

- 美術館・博物館
- イベント
- 教育
- 自治体
- 応対
- 防災
- 講演
- 国際会議
- 遠隔作業 等々

連携

GCP協議会

第1期研究成果を活用した製品・サービス

2025年における社会実装のイメージ（一例）

ショーケースとしての2025年日本国際博覧会（大阪・関西万博）における利活用例

3. アウトカム目標

国内外の様々な場面での「通訳」ニーズに対してICTの力で応え、

社会経済活動において

「言葉の壁」を感じさせない環境を創出することにより、

企業等の労働生産性の向上や「働き方改革」の推進、

利用者利便の一層の増進に寄与する。

4. 推進体制と役割

研究代表：凸版印刷株式会社

(1) 自動同時通訳基盤技術

- ア) 入力分割・要約・翻訳出力最適化技術 (国立研究開発法人情報通信研究機構(NICT))
- イ) 多様な情報源を活用した通訳精度向上技術 (マインドワード株式会社)
- ウ) 自動通訳性能評価尺度の確立 (株式会社インターグループ)

(2) 自動同時通訳システム技術

ア) 自動同時通訳ユーザインタフェース(UI)技術

a) システム利活用要件に応じた総合検証技術 (凸版印刷株式会社)

[社会実証] アバター対話 (凸版印刷株式会社)

[社会実証] 情報伝達 (東日本電信電話株式会社)

[社会実証] 対面・遠隔対話 (ソースネクスト株式会社)

[社会実証] 遠隔協業 (株式会社KDDIテクノロジー)

b) 入力音源分離技術 (ヤマハ株式会社)

[社会実証] 音源分離 (ヤマハ株式会社)

イ) 自動同時通訳プラットフォーム技術 (フェアリーデバイスズ株式会社)

(3) AI学習、検証・評価用データベース(同時通訳用コーパス)の構築 (NICT)

○様々な実証を行い、検証結果を研究開発機関にフィードバック

5. 昨年度までの実証① ~ヤマハ~

(出典：2020年10月12日 ヤマハ（株）ニュースリリース)



YAMAHA
Make Waves

News Release
ヤマハ株式会社
2020年10月12日

リモート講演会・授業をインタラクティブ化する新機能を「Remote Cheerer」に搭載

「CEATEC 2020 ONLINE」ニューノーマルコンファレンスのオープニングスペシャルをプロスポーツ業界と共に行います

～「おもてなしガイド」自動翻訳字幕機能の検証も実施～

ヤマハ株式会社(以下、当社)は、初のオンライン開催となる「CEATEC 2020 ONLINE」(10月20日(火)開幕)にて、ニューノーマルコンファレンスのオープニングスペシャルを「プロスポーツ業界×最新テクノロジーで「スポーツ観戦」のニューノーマルを創る！」と題し、プロスポーツ業界関係者と共に行います。

本コンファレンスは2部構成で開催します。第1部では、スポーツ専門の定額制動画配信サービスを提供するDAZN(ダズン)、明治安田生命J1リーグ、J2リーグ、J3リーグの試合でリモート応援システム「Remote Cheerer powered by SoundiD」(以下、「Remote Cheerer」)を使用するクラブをそれぞれ代表し、ヤガン鳥嶋、ジュビロ磐田、福岡ユナイテッドFCのスピーカーと共にパネルディスカッションを実施します。第2部では、阪神タイガース(野球)、サンロッカーズ渋谷(バスケットボール)、オービックシーガルズ(アメフト)、新日本プロレスリング(プロレス)、清水エスパルス(サッカー)のスピーカーと共に「Remote Cheerer」の取り組みを紹介するトークセッションを行います。「CEATEC」のコンファレンスにおいて、プロスポーツ業界関係者が一堂に会するのは今回が初となります。

これにあわせ、今後、リモートでの開催増加が予想される講演会や国際会議、授業などでご活用いただけるよう、リモートでも現地でもインタラクティブな参加を可能とする新機能を「Remote Cheerer」に搭載し、本コンファレンスで初公開いたします。より幅広いシーンで利用できるように、スポーツやエンターテインメントで使用される拍手などの音声を会場で鳴らす機能に加え、挙手、コメント、アンケートといった機能を新しく追加しました。詳細な使用方法については当日の放送内で案内します。




「Remote Cheerer powered by SoundiD」講演会・授業向け機能
(左：講師者の利用イメージ、右：受講者の利用イメージ)

また、「CEATEC 2020 ONLINE」では、本コンファレンスを含む一部コンファレンスにて「おもてなしガイド」の自動翻訳字幕機能の検証も行います。自動翻訳字幕は、総務省の「ICT重点技術の研究開発プロジェクト(FWI00316)」における「多言語翻訳技術の高度化に関する研究開発」の委託を受けて実施する実証実験として、当社及び国立研究開発法人情報通信研究機構(NICT)が提供します。

当社は現在、330を超える企業・団体が会員とともに「SoundiD(Sound Universal Design)」の普及推進に取り組んでおり、今後も引き続きインバウンド・バリアフリーサービスの推進と、ニューノーマル社会で必要とされる様々な音のユニバーサルデザイン化に関するソリューション構築に取り組んでいきます。

1

実証名	オンラインイベントにおける音源分離実証
イベント名	CEATEC 2020 ONLINE
開催日時	2020年10月20日～23日 10:00～17:00 (最終10:00～20:15)
概要	オンライン展示イベント、CEATEC 2020 ONLINEのカンファレンス枠(3Channelのみ)の映像配信でリアルタイム翻訳字幕を配信
主な検証内容	<ul style="list-style-type: none"> • 翻訳字幕配信システムの動作検証 • 音声認識端末へのマイクの入力経路による音声認識の精度影響検証 など

5. 昨年度までの実証② ～NTT東日本～

(出典：2020年12月22日 NTT東日本（株）ニュースリリース)

2020年12月22日

(報道発表資料)

入間市
東日本電信電話株式会社 埼玉西支店

総務省事業「多言語翻訳技術の高度化に関する研究開発」における
入間市での実証実験の実施について

東日本電信電話株式会社埼玉西支店（支店長：宅間由美子、以下「NTT 東日本」）は、総務省の情報通信技術の研究開発「多言語音声翻訳の高度化に関する研究開発」コンソーシアム*1による「情報伝達」の実証実験を埼玉入間市（市長：杉島 理一郎、以下「入間市」）で実施致しました。この実証実験では、多言語のコミュニケーション手段が求められる実用シーンを想定して当該研究開発により開発する自動同時通訳プラットフォーム技術を利用した情報伝達システムを用いて、多文化共生時代における在留外国人対応の一助として入間市の防災業務にどのように寄与するか確認し、自動同時通訳システムの改善や社会実装につなげていきます。

*1 自動「同時通訳」技術の研究開発コンソーシアムを設立～官民の協力を挙げて「同時通訳」の自動化を実現～
https://www.ntt-east.co.jp/release/detail/20200828_01.html

※本件は、総務省の「ICT 重点技術の研究開発プロジェクト（JPM100316）」における「多言語翻訳技術の高度化に関する研究開発」の委託を受けて実施する研究開発の計画に係るものです。

1. 実証実験の概要
在留外国人を含む人々が集う公共施設など、災害時に「同時通訳技術を用いた情報伝達システム」（以下、情報伝達システム）の活用が見込まれる入間市役所関連施設において、平常時から災害発生直後、さらには避難後に至るまでシームレスに利用可能な情報伝達システムの提供のあり方について同システムを用いた実証およびアンケートやヒアリングを通して確認します。

【実証する項目】

- ・入間市が住民向けに発行している「防災ガイドブック」を実証に参加する外国人に対し、情報伝達システムを用いて説明を行い内容が理解出来たかの確認を行う。
- ・情報伝達システムのユーザーインターフェース、ユーザーエクスペリエンスの確認を行う。
- ・様々な防災業務シーンにおける情報伝達システムの有用性の確認を行う。

【実証日】
2020年12月21日

実証名	災害・防災を想定した自治体での実証実験
開催日	2020年12月21日
概要	在留外国人を含む人々が集う公共施設などにおいて、自治体における平常の業務（防災啓発を含む）から災害発生後の避難施設運営などの有事の業務でも利用可能な情報伝達システムのあり方についての実証を実施
主な検証内容	<ul style="list-style-type: none"> ・自治体が住民向けに発行している「防災ガイドブック（日本語版）」の説明を職員が日本語で行い、外国語話者が翻訳内容を理解出来たかの確認 ・情報伝達システムのUI/UXの確認 など

当日の様子



5. 昨年度までの実証③ ～凸版印刷～

(出典：2020年10月9日 凸版印刷（株）ニューズリリース)

News Topics

TOPPAN

2020年10月9日
凸版印刷株式会社

凸版印刷、教育機関向け「多言語 WEB 連絡帳システム」を開発
教員と保護者が言語を意識することなく Web 上で定型文を選択し、連絡内容を多言語で通知する「多言語 WEB 連絡帳システム」の実証実験を開始

凸版印刷株式会社(本社:東京都千代田区、代表取締役社長:藤 秀晴、以下 凸版印刷)は、総務省の委託研究「多言語翻訳技術の高度化に関する研究開発」を2020年6月9日より実施しています。このたび凸版印刷は、WEB上で定型文を選択することで連絡内容を多言語で通知する「多言語 WEB 連絡帳システム」(以下「本システム」)を教育機関向けに開発。2021年1月から2021年2月末まで実証実験(以下 本実証)を行います。

本システムは、教員が保護者への連絡を行う際、WEB上で送信先と定型文を選択すると保護者があらかじめ登録した言語で連絡内容を一斉にメールで送信します。これにより、教員と外国人児童の保護者とのコミュニケーションを支援することを目指します。

なお、本実証の実施に伴い、本実証に参加を希望する教育機関を2020年10月から2020年12月まで、募集します。



システムの概要

© Toppan Printing Co., Ltd.

■ 開発の背景

日本社会における急速なグローバル化の進展に伴い、教育機関では在籍する外国人児童は増加傾向にあり、その国籍も多様化しています。教員は外国人児童およびその保護者との意思疎通が難しいことが大きな課題となっています。

凸版印刷はこれまで「VoiceBiz」をはじめ、在留外国人との多言語コミュニケーションを支援する多言語音声翻訳アプリを展開してきました。このたび、教育機関向けにWEB上で連絡内容を多言語で通知することが可能な「多言語 WEB 連絡帳システム」を開発。教育現場において在留外国人児童の保護者と教員の間に、言葉の壁をなくした教育環境の実現を目指します。

実証名	教育現場における実証
開催日	2020年12月～2021年2月
概要	教員から親に対する連絡において多言語 WEB連絡帳システムのIDを発行。一定期間多言語システムを活用してもらった後ヒアリングを実施
主な検証内容	<ul style="list-style-type: none"> ・教育現場での利便性確認 ・将来に向けたニーズ確認 など

当日の様子



②自動同時通訳における機械翻訳（同時翻訳）の適用法

本プロジェクトにおける
NICTの機械翻訳（同時翻訳）の研究開発の概要のご紹介

要素技術を統合した同時通訳エンジン

- 同時通訳では、話者が発話を停止しないので、チャンクへの入力分割が必要
- 深層学習で、チャンク分割とチャンク単位でのAI翻訳が可能

1. 音声認識結果の単語列を
分割器がチャンクに

音声認識

結果の

単語列: w1 w2 w3 w4 w5 w6 w7 w8 w9 w10



チャンク:w1 w2 w3

チャンク:w4 w5 w6

チャンク:w7 w8 w9

チャンク:w10 w11 w12



Φ+チャンク翻訳: ||| w1 w2 w3

→ ||| t1 t2 t3

文脈+チャンク翻訳:w1 w2 w3 ||| w4 w5 w6

→ t1 t2 t3 ||| t4 t5 t6

文脈+チャンク翻訳:w1 w2 ... w6 ||| w7 w8 w9

→ t1 t2 ... t6 ||| t7 t8 t9

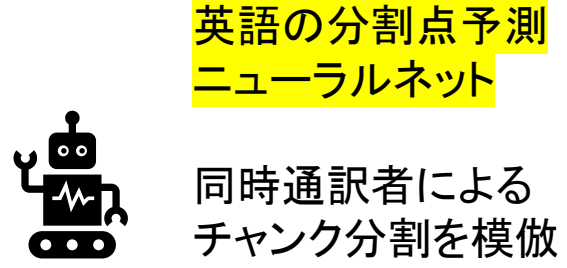
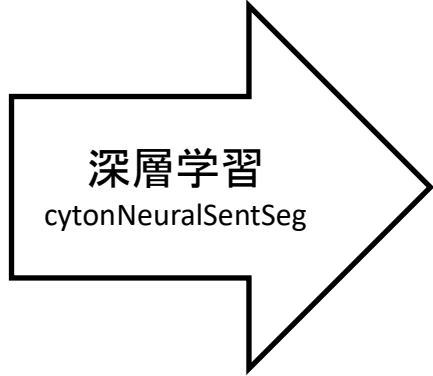
2. 文脈(タグ、履歴等)とチャンク
を入力として、AI翻訳で
チャンク単位に翻訳

→アノテーション付きデータを
翻訳エンジンで学習

同時通訳用コーパスの概要

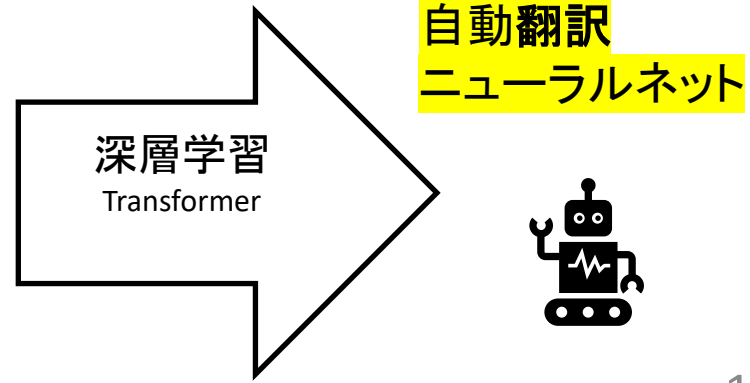
- AI翻訳において、チャンク分割点予測や、役割等のタグ付与、省略データの補完等を深層学習していくうえでコーパス設計が最重要となるため、本事業で構築を鋭意進めている。

同時通訳の分割点データ



役割等のタグ、省略補完データ

役割	発話	
患者	友人が転んで頭を打ったようです。	It seems my friend fell down and hit his head.
医者	意識もないんですか？ => 彼は 意識もないんですか？	Is he also unconscious?

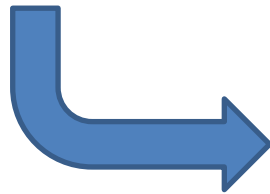


- 機械翻訳自体
Transformerベース
みんなの自動翻訳@TexTraをファインチューニング
- 文脈利用翻訳
タグを利用したり、前の文脈を利用する
- 翻訳単位分割技術(文・チャンク)
同時翻訳用に研究開発⇒やや詳しく説明

五月雨式の文分割の例

even cats were watching this video cats were watching other cats watch this video but what 's important here is the creativity that it inspired amongst this techie geeky internet culture there were remixes someone made an old timey version and then it went international there were remixes someone made an old timey version

Table 1: Illustration of Input for Sentence Segmentation



- even cats were watching this video
- cats were watching other cats watch this video
- but what 's important here is the creativity that it inspired amongst this techie geeky internet culture
- there were remixes
- someone made an old timey version

翻訳単位分割技術 (文・チャンク分割)

- Xiaolin Wang, Masao Utiyama, Eiichiro Sumita: Online Sentence Segmentation for Simultaneous Interpretation using Multi-Shifted Recurrent Neural Network. MTSummit (1) 2019: 1-11
CytonNSS: <https://github.com/arthurxlw/cytonNss>
- 入力: 五月雨式の単語列
- 出力: 多項分布 (i 前の単語が文末のとき i が 1)

教師信号

Time Stamp	1	2	3	4	5	6	7	8	9	10	11	12	13	...
Input	i	'd	like	some	tea	and	cake [†]	that	will	be	a	very	nice	...
Target	0	0	0	0	0	0	1	0	0	0	0	0	0	...
Shift by 1	0	0	0	0	0	0	0	1	0	0	0	0	0	...
Shift by 2	0	0	0	0	0	0	0	0	1	0	0	0	0	...
Shift by 3	0	0	0	0	0	0	0	0	0	1	0	0	0	...
Shift by 4	0	0	0	0	0	0	0	0	0	0	1	0	0	...
残余	1	1	1	1	1	1	1	0	0	0	0	1	1	

先読み

Table 2: Illustration of Multi-Shifted Target Signals for Sentence Segmentation. The input is a sequence of words. The target signals are 0's and 1's where 1 means a sentence boundary after the current time stamp. The last four rows shift the target signals by 1 to 4 time units. [†] Suppose the sentence ends here.

ニューラルネットワーク構造

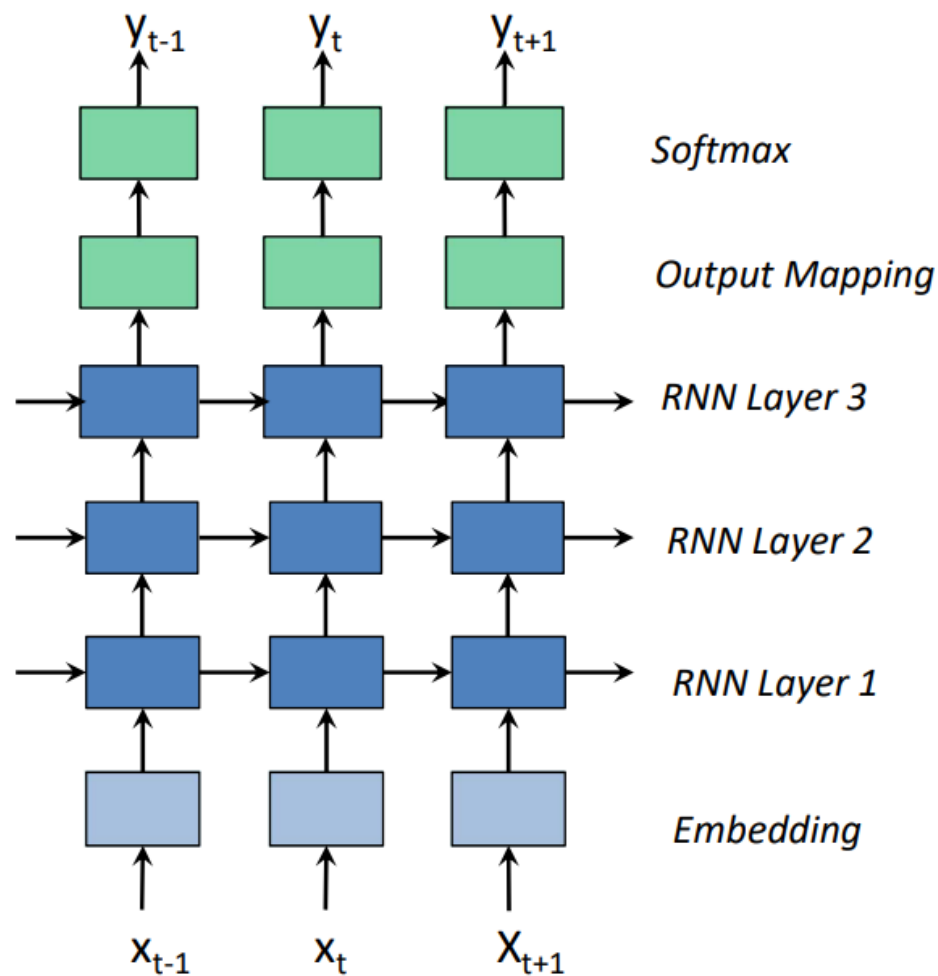


Figure 1: Network Architecture of Multi-shifted RNN Sentence Segmentor

訓練事例

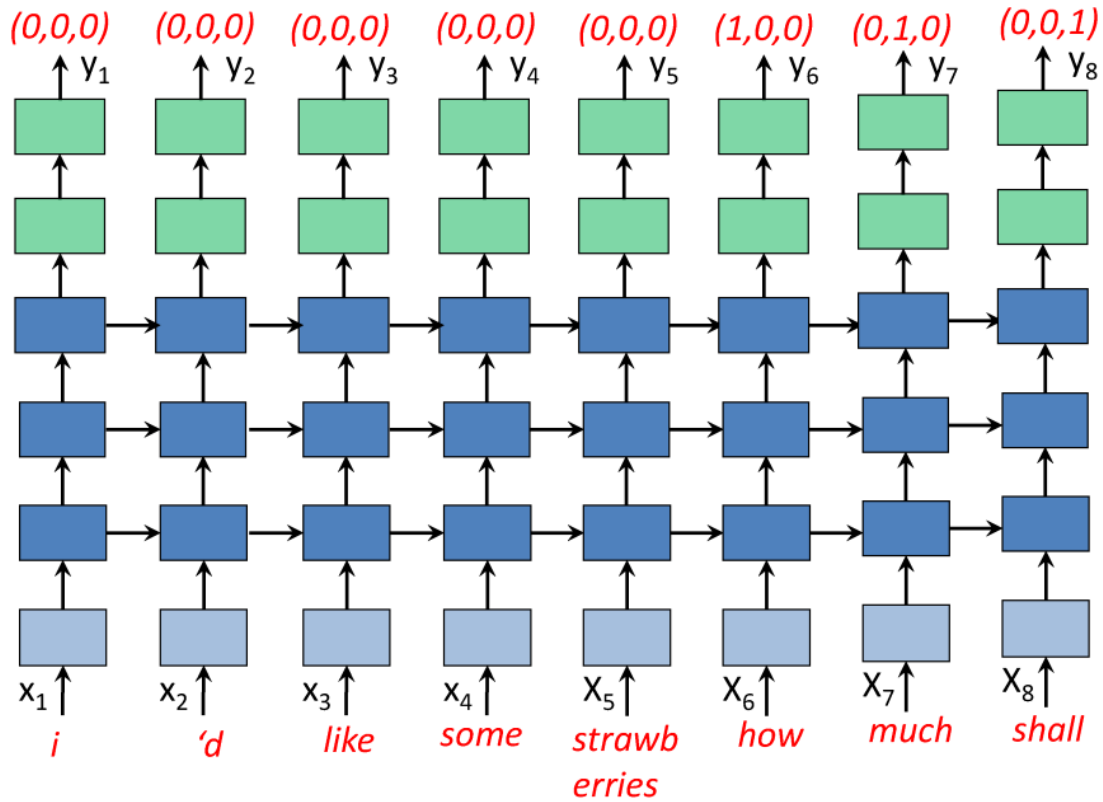


Figure 2: Unrolling a Sample on Multi-shifted RNN Sentence Segmentor. The ASR transcript is “i ’d like some strawberries how much does it cost” where the first sentence ends after “strawberries”. Note that y_t ’s last dimension $y_t^{(4)} = 1 - y_t^{(1)} - y_t^{(2)} - y_t^{(3)}$ is omitted for simplicity.

デコーディング例

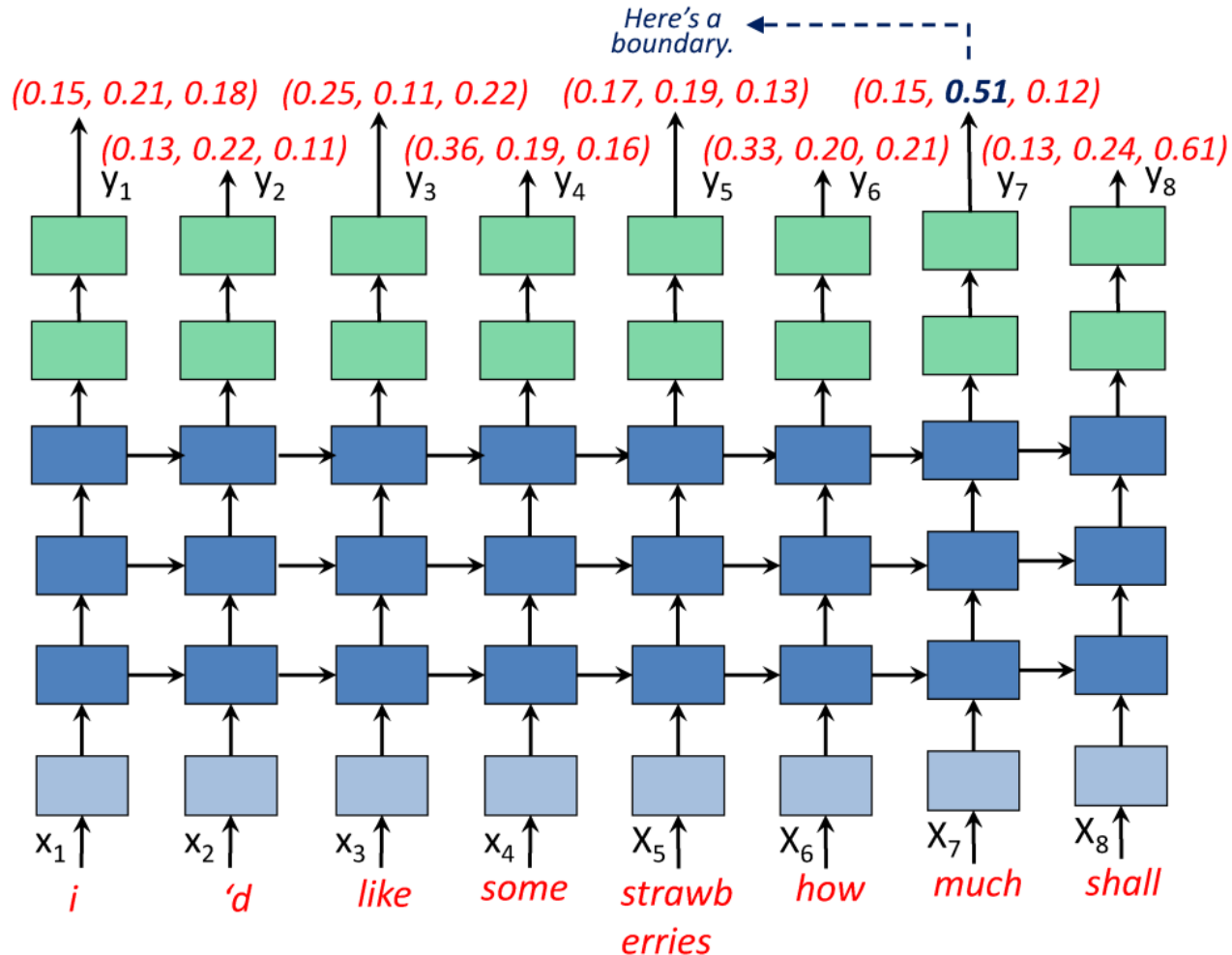


Figure 3: Decoding a Sample on Multi-shifted RNN Sentence Segmentor. Suppose the threshold vector is $(0.40, 0.50, 0.60)$. $y_7^{(2)}$ is the first value that exceeds the corresponding threshold $\theta^{(2)}$. This correctly predicts a sentence boundary after the time stamp 5. Note that y_t 's last dimension $y_t^{(4)} = 1 - y_t^{(1)} - y_t^{(2)} - y_t^{(3)}$ is omitted for simplicity.

実験条件

Corpus	Sentences	Src. Tokens	Trg. Tokens
IWSLT-Train	209,491	4,270,869	6,050,169
News Commentary	223,153	5,689,117	5,660,789
OpenSubtitle(subset) [†]	1,000,000	8,682,476	1,047,208
Dev (test2010 test2011)	2,815	55,426	83,317
Test (test2012 test2013)	2,658	52,766	74,822

Table 3: Experimental Corpora. [†] The subset consists of the first one million sentence pairs.

評価尺度

$$F_1 = 2 \frac{\textit{precision} \cdot \textit{recall}}{\textit{precision} + \textit{recall}} \quad \textit{score} = F_1 - \alpha \cdot \textit{latency},$$

$\alpha = 0.01$

latency = 単語ごとの出力にかかった時間の平均(文頭は長い。文末は短い。平均文長に近い)
12.4 - 8.1 = 4.3 は、文末から4単語で、文全体が出力

実験結果

Methods	Parameters		Performance				
	n_f	thresh.	Precision	Recall	F ₁	Latency	Score
Oracle	先読み		1.000	1.000	1.000	8.126	0.9187
n -gram LM	1	-0.6	0.1402	0.2432	0.1779	8.3410	0.0945
	2	-0.6	0.1862	0.3087	0.2323	9.6480	0.1358
	3	-0.6	0.1928	0.3005	0.2349	11.2520	0.1224
	4	-0.6	0.1944	0.2993	0.2357	12.2930	0.1128
	5	-0.6	0.1935	0.2959	0.2340	13.2410	0.1016
	6	-0.6	0.1927	0.2937	0.2327	14.1570	0.0912
RNN LM	1	-0.8	0.2686	0.3213	0.2926	10.3503	0.1891
	2	-0.6	0.3289	0.3683	0.3475	11.9733	0.2277
	3	-0.8	0.3255	0.3743	0.3482	12.7531	0.2207
	4	-0.8	0.3372	0.3845	0.3593	13.8317	0.2210
	5	-0.8	0.3342	0.3822	0.3566	14.8643	0.2080
	6	-0.8	0.3256	0.3740	0.3481	15.7449	0.1907
Proposed Improve [†]	1 – 6 (...) [‡]		0.3583	0.4387	0.3945	12.3863	0.2706
			8.9%	19.1%	13.5%	-3.4%	18.8%

Score
最大

Table 4: Performance after Training on Standard Set. [†] Improvement versus the stronger baseline of RNN LM. [‡] The optimal threshold vector is (1.0, 0.8, 0.8, 0.5, 0.5, 0.3).

必要な訓練データ

- 文分割
 - ➡ 文分割された生コーパスから学習可能
ASRの出力を模擬するために、句読点除去
- チャンク分割
 - ➡ チャンク分割されたコーパスが必要
 - ➡ 本プロジェクトにおいて同時翻訳コーパスとして作成中

文分割モデル

- ➡ ファインチューニング
- ➡ チャンク分割モデル

チャンク分割例（日英）

原文	MT(文脈なし)
おはようございます	Good morning.
F誌記者のカーティスタッカーと申します	I'm Curtis Tucker, a reporter for F Magazine.
プレゼンの中で	In the presentation.
この製品の開発に七年かかったことに言及されていました	It was mentioned that it took 7 years to develop this product.
特にご苦労なされた点は何だったと思われませんか	What do you think was the most difficult point?
設置手順を簡単にするため	In order to simplify the installation procedure.
持ち運びできる大きさにすることが	You should make it portable.
開発当初から主な目標のひとつでした	It has been one of the main goals since the beginning of development.

| は正解の文分割

チャンク分割例（英日）

原文	MT(文脈なし)
yes	はい、そうです。
it's great for field reps that way when they're traveling	出張中の現場担当者にとっては、そういう使い方はありがたいですね。
they can still have access to all of the information back at the office tracy	オフィスにあるすべての情報にアクセスできます。トレイシー。
tell them how you start a typical day	いつもの1日の始まり方を教えてあげてください。
the first thing I do is login to product a	最初にすることは製品aにログインすることです。
and check my home page at the top of the page	トップページにある私のホームページを見てください。
are my tabs	私の担当です。
which quickly help me find all of the information I need.	すぐに必要な情報を見つけることができますね。

|| は正解の文分割（英語の方が難しい）

ASR出力

- ➡ 文・チャンク分割 (CytonNSS)
- ➡ 句読点付与・要約・省略補完 (実装中)
- ➡ NMT (分野タグ等)

W1 w2 w3 -> GPMT -> t1 t2 t3

W1 w2 w3 w4 w5 w6 w7 -> GPMT -> t1 t2 t3 t4 t5 t6 t7

W1 w2 w3 w4 w5 w6 w7 w8 w9 -> GPMT -> t1 t2 t3 t4 t5 t6 t7 t8 t9

W10 w11 w12 -> GPMT -> t10 t11 t12 (新しい文)

W10 w11 w12 w13 w14 w15 -> GPMT -> t10 t11 t12 t13 t14 t15

GPMT: General Purpose Machine Translation

- チャンク毎に全体を訳し直し。チャンク単位N-to-N翻訳
- 新規原文チャンクに対応する部分を新規訳文チャンクとする
- 要約・省略補完→単言語コーパス必要
- 各種タグ付き対訳コーパス必要(ファインチューニングで訓練)

文末で訳し直し

- 長所: チャンク単位翻訳の遅れ < 文単位翻訳の遅れ
- 短所: チャンク単位翻訳の精度 < 文単位翻訳の精度
- **文末でチャンクと同期**して、文単位翻訳を出力可能

原文チャンク	チャンク単位翻訳	文単位翻訳
前回	Previously	
ちょっと太り気味という数値が出ていて	the results showed I was a little too fat,	
心配だったんだけど...。(1文終了)	and I was concerned.	Last time, I was a little bit overweight, and I was worried.
要注意がなくなって	I didn't receive a warning,	
ほっとしたよ。(1文終了)	and I'm relieved.	I'm relieved that I don't have to pay attention.

③ 同時翻訳を実施するために開発するコーパス

本プロジェクトにおける
NICTの同時翻訳の研究開発で開発中の
同時翻訳コーパスのご紹介

前提：同時通訳とは

◆ 定義（「応用言語学辞典」より一部抜粋）

- 起点言語による話し手の発話を、その進行とほとんど同時に目標言語に通訳していく通訳方式。3～8秒のずれがある。通訳者が使う技術として予測、待機、区切りなどがある。

◆ 英日同時通訳の例（水野的「同時通訳の理論」より）

(1) The relief workers (2) say (3) they don't have (4) enough food, water, shelter, and medical supplies (5) to deal with (6) the gigantic wave of refugees (7) who are ransacking the countryside (8) in search of the basics (9) to stay arrive.

(1) 救援担当者たちの (2) 話では、(4) 食料、水、宿泊施設、医薬品が (3) 足りず、(6) 大量の難民たちの (5) 世話ができないということです。
(7) 難民たちは今村々を荒らしまわって、(9) 生きるための
(8) 食料を求めているのです。

- 大幅な構造の変更（順送りの訳）が行われる
- 情報の省略（欠落）・抽象化なども起こる

同時翻訳コーパスの概要

◆ 同時翻訳モデルに求められる要件（書き言葉の翻訳と比較）

- 原文のスタイルに対応: 話し言葉（非流暢、冗長）
- 訳文のスタイルに対応: 順送り形式、（時間的制約があれば省略も必要）
- 同時性: 原文の発話が終了する前に訳出を開始しなければならない

◆ 構築した同時翻訳コーパスの特徴

1. speech-to-speech 形式ではなく **text-to-text 形式** の対訳（効率重視）
2. 原文を話し言葉調のテキスト（ビジネス会話、講演）とした
3. 通訳者により、原文に **チャンク**（一度に翻訳する単位）の情報を付与
4. 通訳者により、先頭の原文チャンクから順に訳文チャンクを作成し、**順送りの対訳**とした
5. 通訳者が省略する・しない情報を区別できた方が良くと仮定し
原文の内容の **60-80%程度、100%を訳出した2段階の訳文**を作成

入力\出力	テキスト	音声	同時性
テキスト	本コーパス	サイトラ	有
	通常の翻訳		無
音声	字幕	音声同時通訳	有
		音声逐次通訳	無

同時性あり:
原文のある部分と訳文の対応部分が
時間的・空間的に近い位置に配置
されることが望まれる場合

サイトラ: sight translation
テキストを前から順送りでチャンク単位で訳す

同時翻訳コーパス構築作業の指示内容

◆ 一次訳の作成

- 原文が読み上げられる状況を想定し、先頭から意味的なまとまり（チャンク）に分割し、順送り式で訳した訳文を作成する
- 原文の内容の60--80%程度を目安として訳出
- 原文・訳文の分割した箇所に記号「|」を挿入する⇒チャンクに相当
 - Though it seems really simple.₍₁₎ |
it's actually pretty hard.₍₂₎ |
because it forces people to collaborate very quickly.₍₃₎
 - 簡単そうですが、₍₁₎ |
実はかなり難しいんです₍₂₎ |
他の人と素早く協力しなければいけませんから。₍₃₎

◆ 二次訳の作成

- 一次訳に対し、原文の内容の100%を含むように情報を補った訳文を作成する
(例文中の ■ の部分なども明示的に訳出する)

※2021年度作業では、100%訳を作成→60--80%訳作成の作業順とした

同時翻訳コーパスの内容 (2020年版)

◆ データソース

- BSD (中澤 2021): シナリオ作成された日・英の原文とその対訳からなるビジネス対話コーパス。<https://www.aclweb.org/anthology/D19-5204.pdf>
- JNPC (松下 2020): 日本記者クラブでの通訳付き記者会見の書き起こし。<https://www.gsk.or.jp/catalog/gsk2020-a/>
- SIDB (遠山 2004): 名大にて収録・書き起こしされた同時通訳コーパス。<http://sidb.jp/index.html>
- TED (Cettolo 2012): TEDトークの原発話・翻訳からなる対訳コーパス。IWSLT2017のコーパス

データ	重複付与なし			重複付与あり		
	文	文字	文字/文	文	文字	文字/文
日英 BSD-ja	35,673	789,014	22.1	1,279	42,154	33.0
JNPC	1,298	76,144	58.7	789	26,266	33.3
SIDB	2,327	125,533	53.9	344	21,581	62.7
合計	39,298	990,691		2,412	90,001	

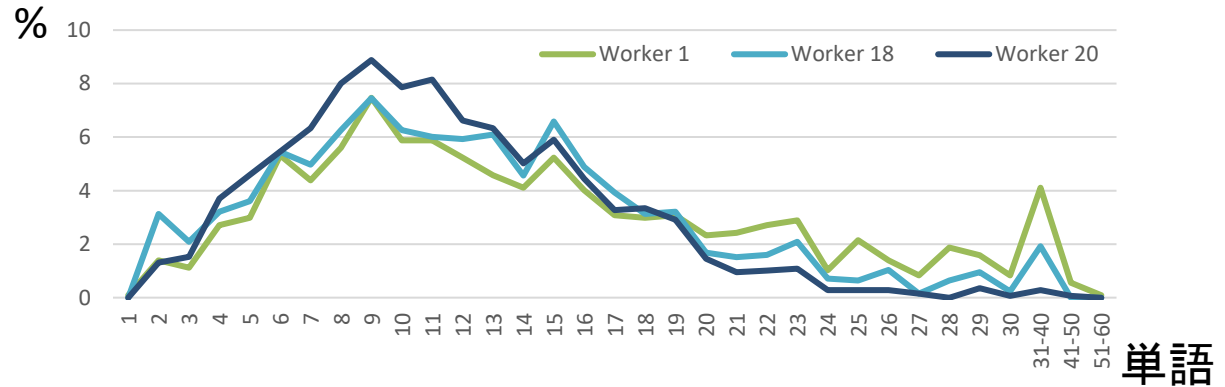
データ	重複付与なし			重複付与あり		
	文	単語	単語/文	文	単語	単語/文
英日 BSD-en	25,978	257,281	9.9	1,040	15,636	15.0
TED	8,146	138,727	17.0	1,194	20,367	17.1
合計	34,124	396,008		2,234	36,003	

※重複付与あり: 異なる3名によりチャンクと訳文を付与したデータ

評価：作業者別チャンク長の分布（日本語）

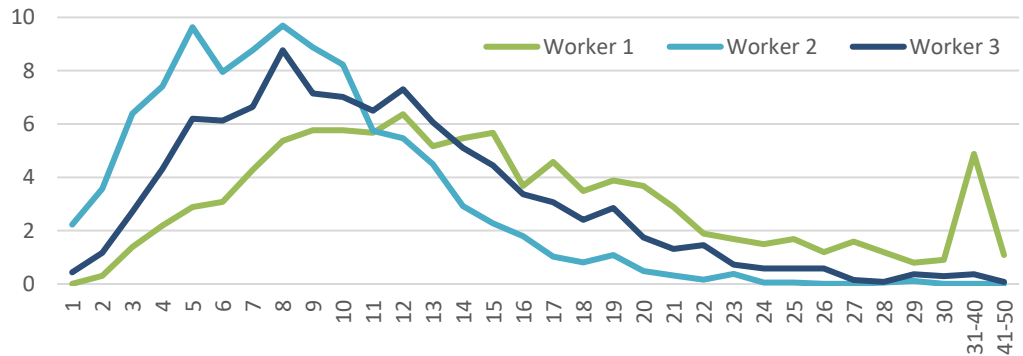
BSD-ja ビジネス会話

文数	単語/文	文字/文
764	20.59	33.08



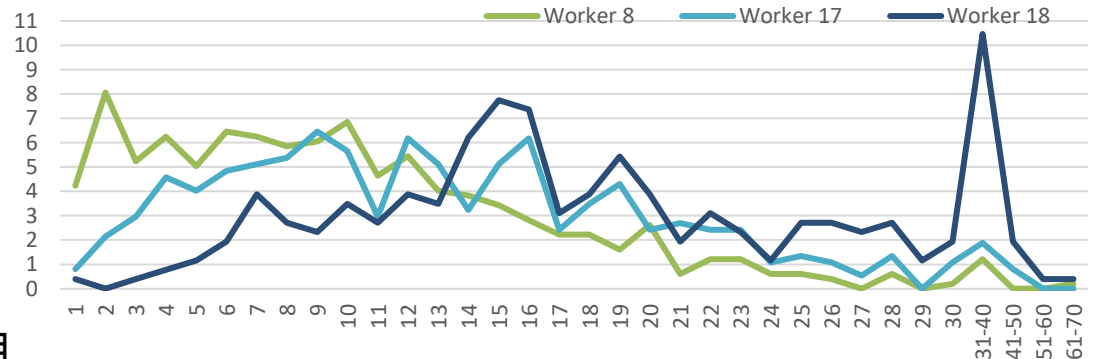
JNPC 講演

文数	単語/文	文字/文
789	19.54	33.29



SIDB 講演

文数	単語/文	文字/文
104	47.8	79.02

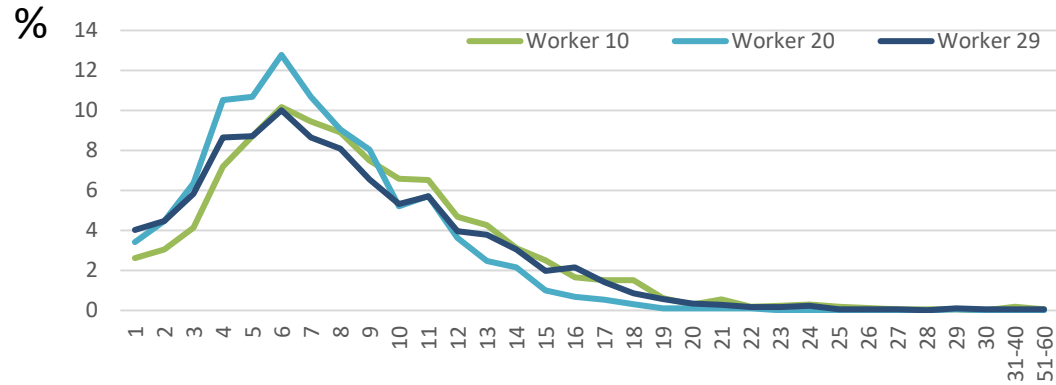


※単語分割は mecab-unidic を使用

評価：作業者別チャンク長の分布（英語）

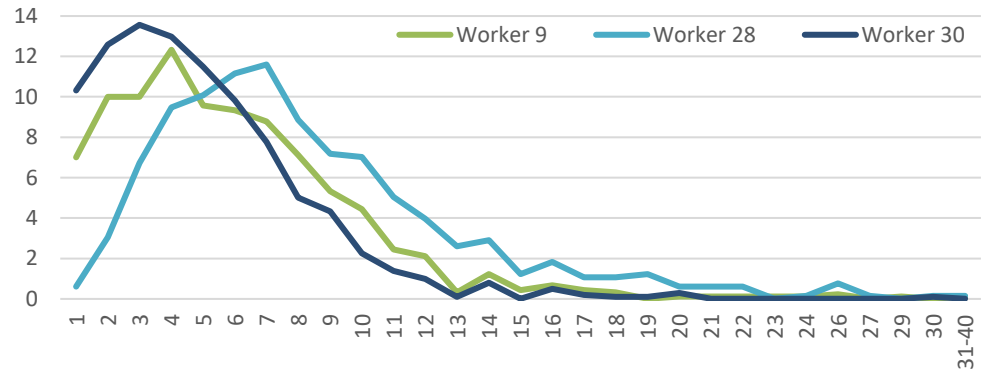
BSD-en ビジネス会話

文数	単語/文
988	13.64



TED 講演

文数	単語/文
392	12.82



単語

評価：一次(60-80%)・二次(100%)訳の差分の内容例

◆ Insert

- (“it’ s out of question”, “its’ out of **the** question”)
- (“I will have an hour”, “I will have an hour **of your time**”)
- (“I mean”, “, **to be more accurate**, I mean”)

◆ Delete

- (“**a** part of”, “part of”)
- (“we **will** advertise”, “we advertise”)
- (“we requested all employees **to** participate in”,
“we requested all employees participate in”)

◆ Replace

- (“we **hold**”, “we **are holding**”)
- (“you can also **see** them”, “you can also **read about** them”)
- (“please **receive** them **at** the reception desk”,
“please **get** them **from** the reception desk”)
- (“**potatochips**”, “**potato chips**”)

④ 同時翻訳のこれまでの成果の概略

本プロジェクトにおける
NICTの同時翻訳の研究開発で
2022年度までの成果

1. 入力 X は、複数文(発話)を1行につなげたもの
2. 分割エンジンが X を翻訳単位 S_i に分ける
3. 翻訳エンジンが S_i を翻訳 $\rightarrow Y_i$
4. Y_i をつなげて Y とする。

評価指標1. S_i の平均長。これが短い方が遅れが短いので良い。

評価指標2. BLEU(Y , 参照訳(X))。複数文を1文に見立てる

入力例：日本語文（発話）を適宜連結して1入力とする

- 議長 どうもありがとうございました。
- 只今ご紹介いただきました木村でございます。
- 本日はここ名古屋市にお招きいただきまして 環境問題の国際的動向と地域からの取り組みについて報告する機会を与えられ 大変光栄に思っております。
- まず最初に 本会議のホストであります名古屋市長に 心から御礼を申し上げたいと思います。
- 昨日 来市長が先頭に立たれまして 私どもいろいろと案内していただき そして名古屋市についてご説明いただきました。
- 非常にバイタリティ溢れ 若さと行動力とそれからリーダーシップ大なりということでお見受けするところ 平成の徳川家康のような方ではないかと密かに思っておる次第でございます。
- それから中村さんを始めとする 本会議主催者の皆様にも 御礼 申し上げたいと思います。
- 私はこの度 名古屋市に初めて参りました。
- 従いましてあまり予備知識はございませんでした。
- 徳川家康のことそれから 小さい頃読みました童話集ですね そういった童話を読んだ記憶がかすかにございます。

入力例：英語文（発話）を適宜連結して1入力とする

- You know, one of the intense pleasures of travel and one of the delights of ethnographic research is the opportunity to live amongst those who have not forgotten the old ways, who still feel their past in the wind, touch it in stones polished by rain, taste it in the bitter leaves of plants.
- Just to know that Jaguar shamans still journey beyond the Milky Way, or the myths of the Inuit elders still resonate with meaning, or that in the Himalaya, the Buddhists still pursue the breath of the Dharma, is to really remember the central revelation of anthropology, and that is the idea that the world in which we live does not exist in some absolute sense, but is just one model of reality, the consequence of one particular set of adaptive choices that our lineage made, albeit successfully, many generations ago.
- And of course, we all share the same adaptive imperatives.
- We're all born. We all bring our children into the world.
- We go through initiation rites.
- We have to deal with the inexorable separation of death, so it shouldn't surprise us that we all sing, we all dance, we all have art.
- But what's interesting is the unique cadence of the song, the rhythm of the dance in every culture.

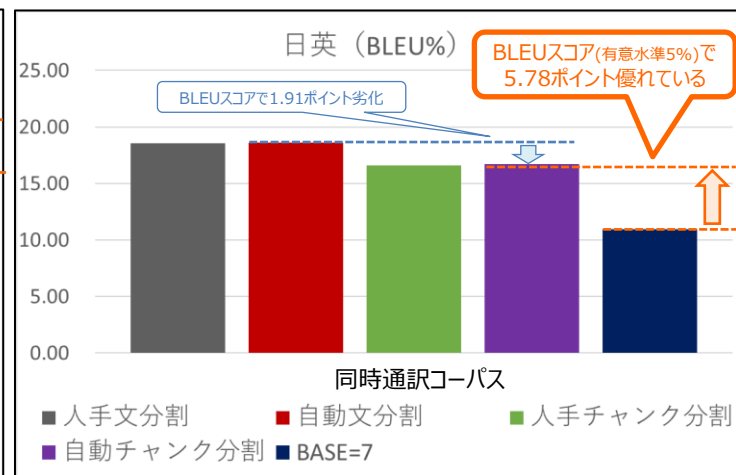
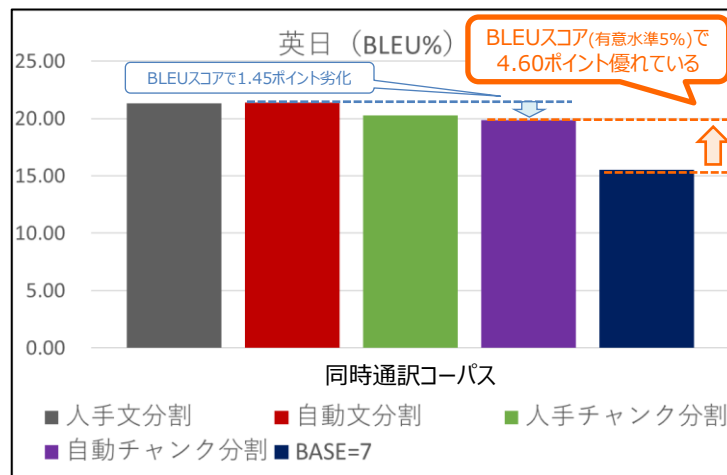
同時翻訳コーパスでのBLEU比較

人手文分割 = ~ 自動文分割

> 人手チャンク分割 = ~ 自動チャンク分割

>> 固定長分割(7単語)

※文(発話)連結数1,3,5,7,10の平均値で評価

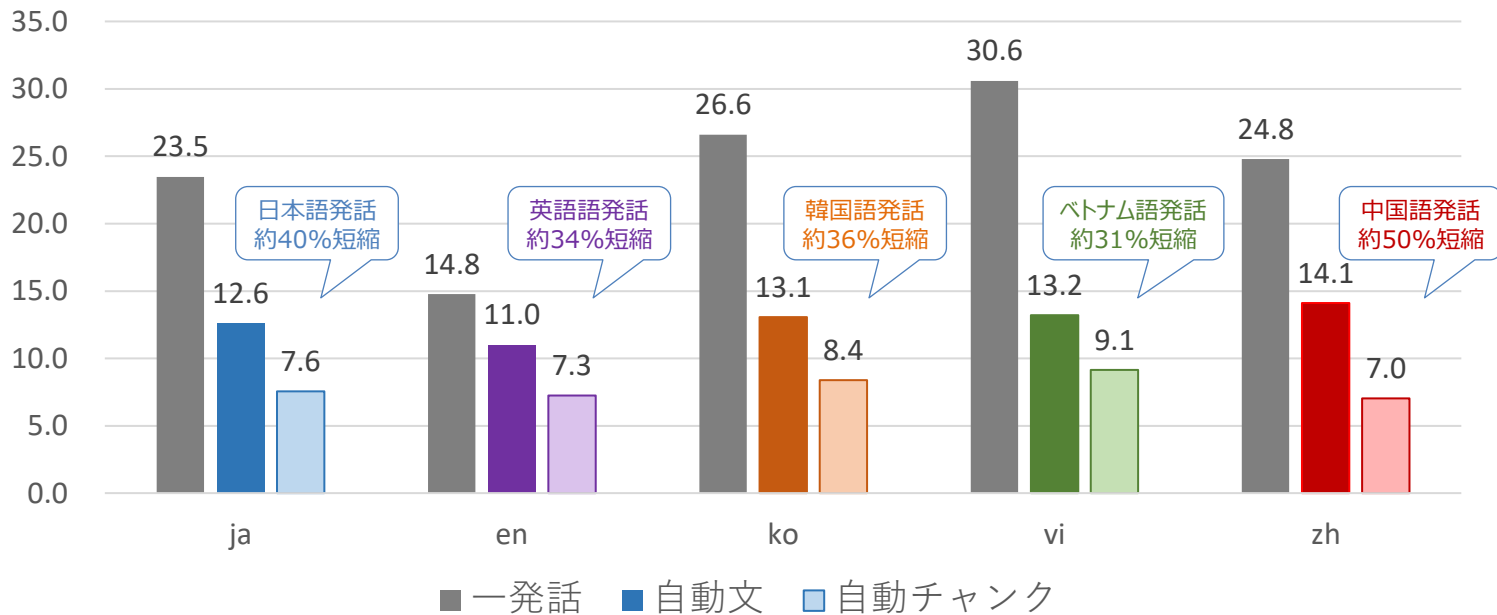


みんなの自動翻訳@TexTraで公開している汎用翻訳エンジンを
ファインチューニングした翻訳エンジンを利用

翻訳単位の平均単語数の比較（同時翻訳コーパス）

翻訳単位長が大幅に短縮 → 遅延の改善
一発話 > 自動文分割 > 自動チャンク分割
※文(発話)連結数1,3,5,7,10の平均値で評価

翻訳単位の平均単語数



文分割エンジンを人手チャンク分割コーパスでファインチューニング

短文化エンジンの公開@みんなの自動翻訳

日本語 → 日本語



短文化 【日本語... 1

翻訳

訳文コピー



オプション

ヘルプ

議長どうもありがとうございました。 ×

只今ご紹介いただきました木村でございます。

本日はここ名古屋市にお招きいただきまして環境問題の国際的動向と地域からの取り組みについて報告する機会を与えられ大変光栄に思っております。

まず最初に本会議のホストであります名古屋市長に心から御礼を申し上げますと思います。

昨日来市長が先頭に立たれまして私どもいろいろと案内していただきそして名古屋市についてご説明いただきました。

非常にバイタリティ溢れ若さと行動力とそれからリーダーシップ大なりというところでお見受けするところ平成の徳川家康のような方ではないかと密かに思っておる次第でございます。

それから中村さんを始めとする本会議主催者の皆様にも御礼申し上げたいと思います。

私はこの度名古屋市に初めて参りました。

従いましてあまり予備知識はございませんでした。

徳川家康のことそれから小さい頃読みました童話集ですねそういった童話を読んだ記憶がかすかにございます。

どうもありがとう。

今ご紹介いただいた木村です。

本日は、名古屋市を訪問し、環境問題の国際的な動向や地域の取り組みについて報告する機会を得たことを光栄に思います。

まず、本会議を主催した名古屋市長に心から感謝したいと思います。

昨日、市長が先頭に立って、名古屋についてご説明いただきました。

非常に活力があり、行動力とリーダーシップがあることから、密かに平成の徳川家康のような人ではないかと思っている。

また、中村さんをはじめとする主催者の皆様にも感謝したいと思います。

初めて名古屋に来ました。

予備知識がありませんでした。

徳川家康と、小さい頃に読みました童話集を読んだ記憶が少しあります。

左はSIDB (遠山 2004)から引用➡短文化

プロジェクト目標再掲：2025年における社会実装のイメージ

- ・ 実証結果に基づくアジャイルな研究開発と普及活動の両軸を推進し、社会実装を目指す。
- ・ 2025年の大阪・関西万博をショーケースとして、日本の自動同時通訳技術を世界に発信。

令和2年 ～ 令和7年（2025年）

※出典：グローバルコミュニケーション計画2025（2020年3月 総務省）

本委託研究開発

同時通訳技術の社会実装

美術館・博物館

イベント

教育

自治体

応対

防災

講演

国際会議

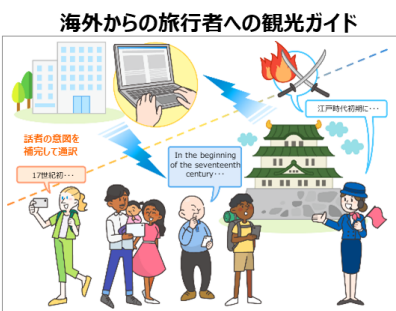
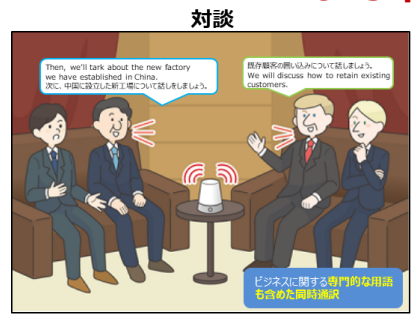
遠隔作業 等々

GCP協議会

連携

第1期研究成果を活用した製品・サービス

2025年における社会実装のイメージ（一例）



ショーケースとしての2025年日本国際博覧会（大阪・関西万博）における利活用例

パビリオン内



同時通訳システムによる来場者とのインタラクティブなプレゼンテーションの実現

会場周辺



会場への移動空間を利用した周辺観光地等への誘客や、会場内のイベント・スケジュールの紹介や告知

会場周辺の宿泊施設受付

