

同時通訳における訳出遅延時間の定量的分析

小野 貴博[†] 遠山 仁美^{††} 松原 茂樹^{††}

[†] 名古屋大学大学院情報科学研究科 〒464-8601 名古屋市千種区不老町
^{††} 名古屋大学情報連携基盤センター 〒464-8601 名古屋市千種区不老町
E-mail: [†]{ono, hitomi}@el.itc.nagoya-u.ac.jp, ^{††}matubara@nagoya-u.jp

あらまし 本論文では、同時通訳における単語レベルでの訳出遅延時間の定量的分析について述べる。同時通訳は人間の高度な言語処理活動である。しかし、訳出プロセスの複雑さに加え、大規模な実データが存在しなかったこともあり、そのメカニズムについてはほとんど解明されていない。本研究では、名古屋大学同時通訳データベースに対して、単語単位の時間情報を音声認識ツールを用いて付与し、分析を行った。その結果、日英同時通訳及び英日同時通訳における訳出遅延現象の違い、ならびに、遅延の要因が明らかになった。

キーワード 英日同時通訳, 日英同時通訳, 対訳コーパス, 訳出遅延時間

Quantitative Analysis of Time Delay in Simultaneous Interpretation

Takahiro ONO[†], Hitomi TOHYAMA^{††}, and Shigeki MATSUBARA^{††}

[†] Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, 464-8601 Japan
^{††} Information Technology Center, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, 464-8601 Japan

E-mail: [†]{ono, hitomi}@el.itc.nagoya-u.ac.jp, ^{††}matubara@nagoya-u.jp

Abstract In this paper, we discuss word-level time delay in simultaneous interpreting by conducting quantitative analysis. The mechanism of simultaneous interpreting has been hardly understood because of the complexity in its translation process and lack of large-scale databases. We have provided word-level temporal information for the simultaneous interpretation database at Nagoya University by utilizing speech recognition techniques, and investigated time delay between words. Based on the results of the analysis, we observed different characteristics of time delay phenomena between Japanese-to-English and English-to-Japanese simultaneous interpreting, and also confirmed factors of time delay occurrences.

Key words English-to-Japanese simultaneous interpretation, Japanese-to-English simultaneous interpretation, bilingual corpus, time delay

1. はじめに

高度な言語処理活動である同時通訳は、訳出プロセスの複雑さに加え、大規模な実データが存在しなかったこともあり、そのメカニズムはほとんど解明されていない。同時通訳における訳出は最大限に瞬時であることが期待されるが[1]、日本語と英語のように文法構造上の差異が大きい言語間の通訳では、文の構造が決定するタイミングや語順が言語ごとに異なるため、訳出に遅延が生じる。

近年、同時通訳に対して、認知言語学的見地から訳出遅延を分析し、訳出時間と概念化の関係について考察した研究[1]、日本語の構文構造から日英同時通訳における予測の可能性を論じ、

日英同時通訳理論の構築を試みた研究[2]、遅延時間の巻き返しストラテジーや訳出タイミングを通訳者間で比較分析した研究[3]などが行われている。しかし、小規模のデータを用いた定性的な分析や理論的考察が中心であり、大規模な同時通訳事例を用いて定量的に分析した試みは存在しない。

本論文では、日英、及び、英日同時通訳における訳出遅延時間の定量的分析について述べる。分析には、名古屋大学同時通訳データベースに収録されている、独話の同時通訳データを用いた。分析にあたり、書き起こしテキストの各単語を、書き起こし元の音声信号と対応付けることにより、発話単位よりさらに詳細な単語レベルでの発声開始・終了時刻を付与した。その結果、講演者発話と通訳者発話に含まれる対訳語の訳出遅延時

間を大規模に観察することが可能になった。これらの時間情報を用いて、日英、および、英日同時通訳における対訳語の訳出遅延時間を算出し、訳出遅延の現象、および、その要因を比較分析した。

2. 名古屋大学同時通訳データベース

名古屋大学同時通訳データベース [4] は、182 時間におよぶ日英・英日双方向の同時通訳音声（独話・対話）を収録している。本分析には、本データベースに収録されている独話データを用いた。

音声収録は教室レベルの環境で行われ、講演者と通訳者の音声を 16bit, 16kHz でステレオ録音している。通訳者は、話者をガラス越しに観察できる通訳専用のブースに入り、ヘッドホンから流れる講演者の音声を同時通訳している。講演者には通訳者の音声が聞こえないようになっており、実際の同時通訳付き講演と同様の環境を再現している。通訳者は、第一線で活躍するプロの通訳者を起用し、訳出品質を保證している。

音声の書き起こしは、日本語話し言葉コーパス (CSJ) [5] の書き起こし基準に準拠しており、また、話し言葉に特徴的に現れるフィラーや言い直し、言い誤りといった発話現象に対して談話タグを付与している。講演者、及び、通訳者の音声を 200msec 以上のポーズ（無音声区間）で分割し、無音区間で挟まれた発話区間を発話単位とした。全ての発話単位には発声の開始・終了時刻を付与している。

同時通訳の訳出プロセスを観察し、様々な知見を収集するためには、話者発話とそれに対応する通訳者発話の対訳データを大規模に分析する必要がある。そのため、発話単位レベルならびに文レベルの対訳対応データを大量に作成している [6]。

本分析では、名古屋大学同時通訳データベースにおける単語に時間情報を付与し、データベースの高度化をはかることにより、単語レベルでの訳出遅延の観察を可能にした。

3. 訳出遅延時間の測定

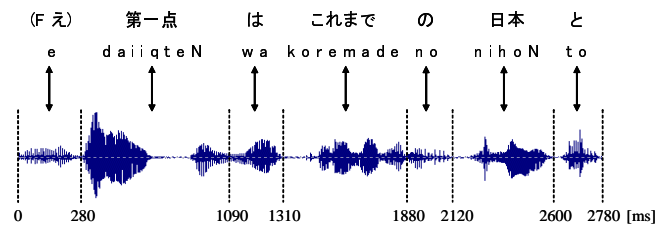
講演者発話と通訳者発話の中から対訳表現を抽出し、それらの間の訳出遅延時間を測定する。訳出遅延時間の測定手法は以下の通りである。

- (1) 単語の発声開始・終了時刻の付与
- (2) 対訳表現の抽出
- (3) 訳出遅延時間の算出

3.1 単語の発声開始・終了時刻の推定

講演者発話、および、通訳者発話に含まれる各単語に、発声開始・終了時刻を付与する。発声開始・終了時刻は、単語を書き起こし元の音声信号と対応付けることにより推定する (図 1)。まず各単語の読みを音素で表記し、音響モデルを用いて単語と音声信号の区間を対応付ける。単語の発声開始・終了時刻は、対応する音声信号の開始・終了時刻とする。

日本語単語の読みは、形態素解析器茶筌 [7] が出力するカタカナで表記された読みを Julius 付属の読み付与ツールで音素



単語	(F 発)	第一点	は	これまで	の	日本	と
開始時刻 (ms)	0	280	1090	1310	1880	2120	2600
終了時刻 (ms)	280	1090	1310	1880	2120	2600	2780

図 1 単語の発声開始・終了時刻の推定

表 1 単語発声開始・終了時刻推定の精度

	日本語	英語
平均誤差 (ms)	31.1	27.9
最大誤差 (ms)	98	144

列に変換した。英語単語の読みは、TIMIT Acoustic-Phonetic Continuous Speech Corpus 付属の辞書 (6,229 エントリ) と The CMU Pronouncing Dictionary version 0.6 (129,481 エントリ) を用いて与えた。

単語と音声信号の対応付けには、音声認識エンジン Julius [8] を用いた。日本語の音響モデルには、Julius 付属の不特定話者 PTM トライフォンモデルを用いた。英語の音響モデルは、HTK [9] を用いて TIMIT Acoustic-Phonetic Continuous Speech Corpus [10] から構築した 2 混合モノフォンモデルを用いた。

上記の手法によって、名古屋大学同時通訳データベースの日英および英日同時通訳に対して単語の発声開始・終了時刻の付与を行ったところ、日本語では 98.8% (182,464/184,750)、英語では 93.9% (128,769/137,100) の単語に時刻を付与できた。時刻付与に失敗した主な原因として、読み付与の失敗や Julius の解析中断が挙げられる。表 1 に、人手による作業結果に対する自動推定した時刻の誤差を示す。日本語は無作為に選んだ 10 発話単位、英語は 7 発話単位に対する結果である。日本語の評価に用いた 10 発話単位での人手による作業の揺れは平均 14.4 ミリ秒、最大 75 ミリ秒であり、それを鑑みると十分な精度であると言える。

3.2 対訳表現の抽出

講演者発話と通訳者発話の中から対訳表現を抽出する。講演者発話と通訳者発話の文レベルの対訳対応に含まれる全ての単語、および、フレーズの組に対して、対訳語辞書を用いて対訳関係にあるか否かを判定し、一対一で対応する単語またはフレーズの組を対訳表現として抽出する。ただし、以下の条件のいずれかに該当する組は、抽出の対象から除いた。

- (1) 講演者側または通訳者側の単語が機能語である。
- (2) 通訳者の発声開始が講演者の発声開始より早い。

誤った対訳表現の抽出を避けるため、本分析では、一対一の対

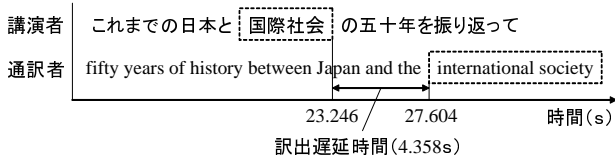


図 3 対訳表現の訳出遅延時間

表 2 日本語講演及び日英同時通訳の基礎統計

項目	日本語講演	日英同時通訳
講演数	12	18
収録時間 (s)	20,464	31,032
文数	1,532	2,848
発話単位数	7,196	11,691
単語数	45,981	47,385
異なり単語数	4,425	4,266

訳対応にある組のみを抽出の対象とした。フレーズは単語の列とし、10 語以下の単語からなるフレーズを対訳表現抽出の対象とした。対訳辞書には、英辞郎 [11] から作成した 203,508 エントリを用いた。

名古屋大学同時通訳データベースの日英同時通訳から 4,731 組、英日同時通訳から 3,833 組の対訳表現を抽出した。図 2 に日英同時通訳から抽出した対訳表現の例を示す。異なる 2 つの同時通訳事例から抽出した対訳表現が、実際に正しく対訳関係にあるかを人手によって検証したところ、正解率は 94.2% (181/192) であり、十分なスコアであった。

3.3 訳出遅延時間の計算

対訳表現の訳出遅延時間は、講演者の発声終了時刻と通訳者の発声開始時刻の差とする。図 3 に、“国際社会”と“international society”の間の訳出遅延時間の計算例を示す。講演者が“国際社会”の発声を終えてから、通訳者が“international society”の発声を開始するまでにかかった 8.397 秒が、この場合の訳出遅延時間になる。訳出遅延時間が計算できた対訳表現は、日英同時通訳で 4,444 組、英日同時通訳で 3,384 組であった。

4. 訳出遅延時間の分析

4.1 分析に用いたデータ

日英同時通訳には、名古屋大学同時通訳データベースに収録されている 12 個の日本語講演と、それらの日英同時通訳 18 個を用いた。日本語講演及び日英同時通訳の基礎統計を表 2 に示す。18 個の日英同時通訳は、全て異なる通訳者が行っている。これらの 18 組の日本語講演・日英同時通訳から抽出した 4,444 組の対訳表現を分析に用いた。

一方、英日同時通訳には、名古屋大学同時通訳データベースに収録されている 14 個の英語講演と、それらの英語同時通訳 14 個を用いた。分析に用いた英語講演及び英日同時通訳の基礎統計を表 3 に示す。14 個の英日同時通訳は、全て異なる通訳者が行っている。これらの 14 組の日本語講演・日英同時通訳から抽出した 3,384 組の対訳表現を分析に用いた。

表 3 英語講演及び英日同時通訳の基礎統計

項目	英語講演	英日同時通訳
講演数	14	14
収録時間 (s)	14,014	14,047
文数	1,528	1,879
発話単位数	5,496	5,436
単語数	27,926	36,611
異なり単語数	3,360	2,885

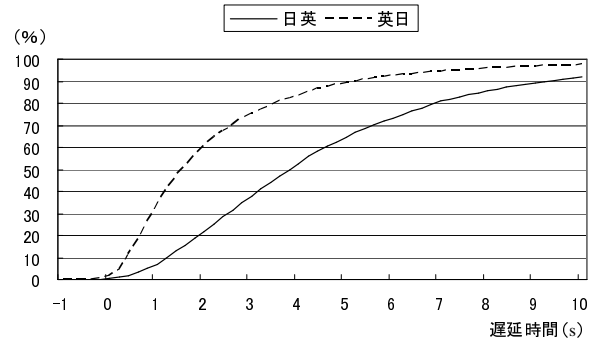


図 4 日英及び英日同時通訳の訳出遅延時間

表 4 日英同時通訳における品詞別訳出遅延時間

	名詞	動詞	形容詞	副詞
対訳表現数	2,836	319	90	143
平均遅延時間 (s)	4.8	4.3	4.9	4.4

4.2 訳出遅延時間の分布

日英・英日同時通訳における対訳表現の訳出遅延時間の累積割合を図 4 に、それぞれ実線と破線で示す。日英同時通訳における訳出遅延時間の平均は 4.8 秒、英日同時通訳での平均は 2.4 秒である。英日同時通訳で遅延時間の 80% (2,712/3,384) が 3.5 秒以下であるのに対し、日英同時通訳では 6.9 秒以下の遅延時間で全体の 80% (3,568/4,444) を占める。英日同時通訳では入力された表現を即座に訳出する傾向があり、日英同時通訳では表現をより長い時間、記憶しながら訳出していることがわかる。

4.3 品詞と訳出遅延時間の関係

SOV 言語である日本語と SVO 言語である英語では、構文構造の違いから語順が大きく異なる。このため、入力された単語の構文的役割を即座に決定できない場合もあり、訳出遅延を生じる原因の一つになっていると考えられる。文の構造を決定する名詞や動詞を中心に、品詞ごとに訳出遅延の特徴を観察した。

日英同時通訳における品詞ごとの対訳表現の数と遅延時間の平均を表 4 に、名詞及び動詞の遅延時間の累積割合を図 5 に示す。名詞の訳出遅延時間は平均 4.8 秒、動詞の訳出遅延時間は平均 4.3 秒であり、名詞に比べて動詞の訳出遅延時間が小さい。日本語で文末に位置する動詞の出現を待ってから、動詞とその目的語を訳出しているためであると考えられる。

英日同時通訳における品詞ごとの対訳表現の数と遅延時間の平均を表 5 に、名詞及び動詞の遅延時間の累積分布を図 6 に示す。名詞の訳出遅延時間は平均 2.2 秒、動詞の訳出遅延時間は平均 4.7 秒であり、名詞に比べて動詞の訳出遅延時間が大きい。

51	この時期 <u>165:(A 千九百七十五;1975)</u> 年には <u>164第一</u> 回の <u>167.七</u> か国 <u>166.サミット</u> が (R ヴァンブイ) ヴァンブイエで開催されました <SB>	In <u>165:(A nineteen seventy five, 1975)</u> the <u>164.first</u> <u>166.summit</u> was held in (? Vamburia) (F ah) (?R par)) participated by the <u>167.seven</u> industrialized countries .
52	<u>168.それ</u> 以来この <u>170.七</u> か国 サミットのメンバーとしてわが国としては <u>169.外交</u> を展開しているわけでございます <SB>	And <u>168.since then</u> our country has been promoting <u>169.diplomacy</u> as a member of the <u>170.seven</u> member countries .
53	<u>171.サミット</u> は (A 七十九;79) 年には <u>172.東京</u> で開催されております <SB>	The <u>171.summit</u> (F ah) was held in <u>172.Tokyo</u> in (A nineteen seventy nine; 1979) .
54	(R こないだ;この間)この時期におきましては (A 千九百七十六;1976) 年には (F え<H>) <u>173.日本</u> は <u>174:(A エスピーティー;NPT)</u> の (R ひすい;批准)を行っております <SB>	During this period in (A seventy-six, 76) <u>173.Japan</u> (F ah) ratified the <u>174.NPT</u> .
55	また <u>176.冷戦</u> のさなかでございましたが (F え<H>) <u>178:(A 千九百八十三;1983)</u> 年には <u>179:(A ケーエーエル;KAL)</u> 大韓航空の (F え<H>) (R しゅ) (F え<H>)突撃事件でございますとか 或はラングーンでの <u>177.韓国</u> <u>181.大統領</u> の暗殺 <u>180.未遂</u> 事件といった <u>175.非常に</u> <u>182.緊張</u> の <u>183.高い</u> 時期でもございました <SB>	(R And during the) Although it was during the <u>176.Cold War</u> in <u>178:(A nineteen eighty three; 1983)</u> the <u>179.KAL</u> (R accident) plane crash (?R occu)) occurred and (F ah) there was a assigation <u>180.attempt</u> of the <u>181.president</u> of <u>177.South Korea</u> and <u>182.tensions</u> were <u>175.very</u> <u>183.high</u> .

図 2 日英同時通訳から抽出した対訳表現

表 5 英日同時通訳における品詞別訳出遅延時間

	名詞	動詞	形容詞	副詞
対訳表現数	1,973	432	354	141
平均遅延時間 (s)	2.0	4.2	2.1	2.9

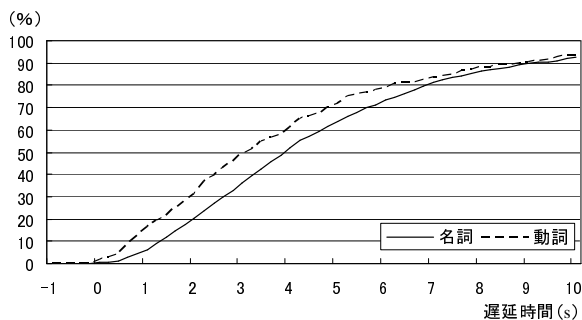


図 5 日英同時通訳における名詞及び動詞の訳出遅延時間

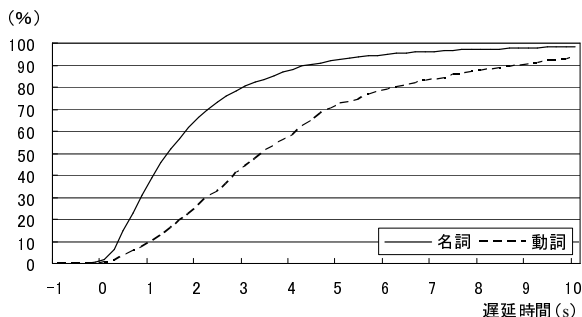


図 6 英日同時通訳における名詞及び動詞の訳出遅延時間

日本語の動詞を訳出するには文末まで待つ必要があり、動詞の訳出は遅れるが、その一方で、日本語では名詞句の語順が柔軟であり、名詞を即座に訳出することができるためであると考えられる。

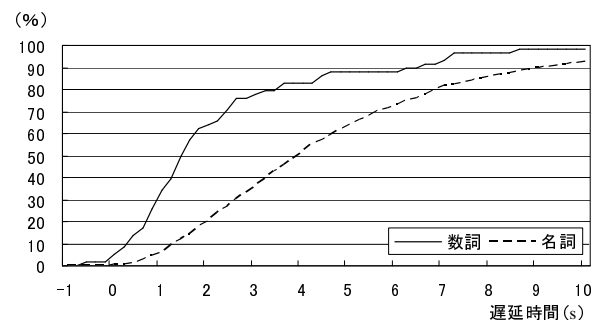


図 7 日英同時通訳における数詞の訳出遅延時間

4.4 数詞の遅延時間

通常の名詞が概念として捉えることができるのに対して、年号や面積、人数などのランダムな数字の列を記憶することは通訳者にとって負荷が大きいと予想される。図 7 と 8 に日英及び英日同時通訳における数詞の訳出遅延時間の分布を示す。日英同時通訳では、通常の名詞の遅延時間が平均 4.8 秒であるのに対して、数詞の平均遅延時間は 2.2 秒と非常に小さい値である。英日同時通訳でも、名詞の平均遅延時間が 2.0 秒であるのに対し数詞では 1.8 秒であり、その差はあまり大きくないが数詞の方が遅延時間が小さい。通訳者がランダムな数字の記憶を避けるため、数詞をできるだけ早く訳出しているためであると考えられる。

4.5 訳出遅延と通訳者の発話状態

訳出遅延には、直前の発話を訳出し終わらないため訳出のタイミングが遅れる場合と、講演者の発話内容が十分に決定されるまで必要な情報を待つために遅れる場合が考えられる。講演者の発声終了から通訳者の訳出開始までの間の通訳者の発話時間を図 9 に示す。横軸は、訳出遅延時間全体に対して通訳者が訳出していた時間の割合である。遅延している間に通訳者が全

表 6 日英同時通訳における予測を伴う訳出例

講演者 非常に強い訛りの英語を話しますし...

通訳者 they speak their own English with a strong accent

単語	非常	に	強い	訛り	の	英語	を	話し	ます	し
開始時刻 (s)	17.698	18.108	18.278	18.738	19.168	20.118	20.368	20.428	20.858	21.148
終了時刻 (s)	18.108	18.278	18.738	19.168	19.388	20.368	20.428	20.858	21.148	21.438

単語	they	speak	(R the)	their	own	English	with	a	strong	accent
開始時刻 (s)	20.556	20.816	22.666	22.726	23.056	23.766	24.456	24.846	24.926	25.316
終了時刻 (s)	20.816	21.496	22.726	23.056	23.456	24.456	24.846	24.926	25.316	25.916

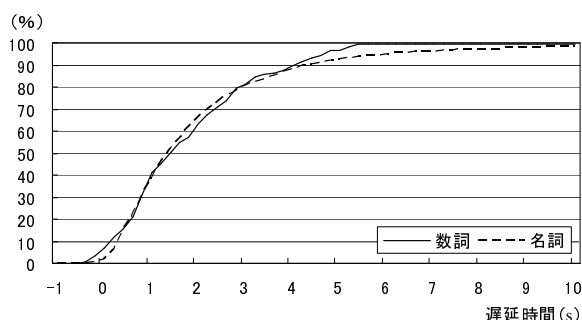


図 8 英日同時通訳における数詞の訳出遅延時間

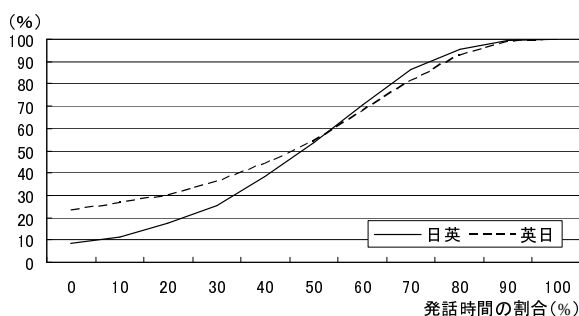


図 9 訳出遅延時間と通訳者の発話状態

く発話しない場合が、日英同時通訳で 9% (383/4,443)、英日同時通訳で 23% (836/3,592) 存在しており、講演者発話の構文の決定に必要な入力を待つために遅延している場合があることを確認した。

4.6 講演者の発声終了を待たない通訳者の発声

通訳者は講演者の表現を聞いてから訳出すると考えられるが、講演者が発声を終える前に通訳者が訳出を開始する対訳表現が存在した。これらの対訳表現では、講演者の表現と通訳者の表現がほぼ同時に発声されており、通訳者が講演者の発声を完全に確認しているとは考えにくく、通訳者が講演者の発話内容を予測しながら訳出していると考えられる。通訳者が講演者発話を待たずに訳出開始できる要因を調査したところ、日英同時通訳事例の分析では、予測を伴う訳出が 6 組観察された。6 組中 5 組が動詞の訳出であり、それらは、speak, think, learn, play であった。図 6 に、「話し」が「speak」と訳出される例を示す。これらの単語はほぼ同時に発話されており、通訳者は「強い訛りの英語」に対する動詞として「speak」を予測しながら訳出し

ていると考える。

5. まとめ

本論文では、大規模同時通訳データベースを用いて、日英及び英日同時通訳における対訳表現の訳出遅延時間の特徴を分析した。分析の結果、以下の特徴が明らかになった。

(1) 日英同時通訳に比べて英日同時通訳の訳出遅延時間は小さい。

(2) 日英同時通訳では名詞より動詞の訳出が早く、英日では逆に名詞の訳出が早い。

(3) 日英・英日同時通訳のいずれでも、ランダムな数字を含む表現の訳出が早い。

(4) 通訳者によっては、構文の決定に必要な入力を待つために遅延している場合がある。

(5) 日英・英日同時通訳で即時的な訳出のために予測が行われることがある。

今後は、各単語の文中での役割を考慮しながら訳出遅延の観察を行い、日英・英日同時通訳を特徴付ける要因、ならびに、訳出遅延が発生する原因を調査する予定である。

文献

- [1] 船山仲他, 笠原多恵子, 西村友美, “同時通訳における訳出遅延のメカニズム”, 平成 12-13 年度科学研究補助金 (基盤研究 (c)(2)) 研究成果報告書『同時通訳における対訳遅延の認知言語学的研究』, pp. 3-24, 2004.
- [2] 水野的, “日英同時通訳研究ノート”, 通訳理論研究, Vol. 5, No. 2, pp. 4-21, 1995.
- [3] 遠山仁美, 松原茂樹, “訳出遅延時間と訳出開始タイミングに着目した同時通訳者の原発話追従ストラテジーに関する分析”, 通訳研究, No. 6, pp. 113-128, 2006.
- [4] 遠山仁美, 松原茂樹, 河口信夫, 稲垣康善, “CIAIR 同時通訳データベースの構築と利用”, 信学技報, Vol. 103, No. 487, pp. 7-12, 電子情報通信学会, 2004.
- [5] 前川喜久男, 籠宮隆之, 小磯花絵, 小椋秀樹, 菊池英明, “日本語話し言葉コーパスの設計”, 音声研究, 第 4 巻, 第 2 号, pp. 51-61, 2000.
- [6] 高木亮, 松原茂樹, 稲垣康善, “同時通訳コーパスの対訳アライメント手法とその評価”, 情報処理学会第 64 回全国大会, 2000.
- [7] 松本裕治, 山内啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原正幸, “形態素解析システム『茶釜』 version 2.3.3 使用説明書”, 奈良先端科学技術大学院大学松本研究室, 2003.
- [8] 大語彙連続音声認識システム Julius, <http://julius.sourceforge.jp/>
- [9] Hidden Markov Model Toolkit (HTK), <http://htk.eng.cam.ac.uk/>
- [10] TIMIT Acoustic-Phonetic Continuous Speech Corpus, <http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC93S1>
- [11] 英辞郎, <http://www.eijiro.jp/>