

課題遂行型対話実験による日英及び日中音声翻訳システムの評価

水島 昌英[†] 竹澤 寿幸[†] 清水 徹[†] 菊井 玄一郎[†]

[†] ATR 音声言語コミュニケーション研究所 〒619-0288 京都府相楽郡精華町光台 2-2-2

ATR で開発中の日英及び日中音声翻訳システムのコミュニケーションツールとしての性能評価を、課題遂行型対話実験を通じて試みた。日英と比較すると、中国語の音声認識性能が低く、それが日中対話の課題達成率を低下させたが、誤認識の影響を除いて考えれば、日英、日中ともに、概ね一つの情報を、互いの約2回の発話で伝達できることが分かった。さらに一発話を短くし、誤認識の頻度を減らせた話者が、効率よく確実に課題を進められていた。

Task-Oriented Dialogue Experiments for an Evaluation of the ATR Speech-to-Speech Translation System in Japanese-English and in Japanese-Chinese

Masahide MIZUSHIMA[†] Toshiyuki TAKEZAWA[†] Tohru SHIMIZU[†] and Genichiro KIKUI[†]
[†]ATR Spoken Language Communication Research Laboratories, Kyoto 619-0288, Japan

We conducted task-oriented dialogue experiments in order to evaluate the ATR speech-to-speech translation (S2ST) system, which supported not only Japanese-English but also Japanese-Chinese, as a communication tool. On average, one piece of information could be transmitted from the information source speaker to his/her counterpart after speaking a total of two times. The lower speech recognition performance of Chinese, as opposed to other languages, decreased the task achievement rates of the experiments between Japanese and Chinese speakers. Speakers who could divide information and use short sentences in each utterance, thereby decreasing recognition errors, could perform the tasks more steadily and more effectively.

1. まえがき

ATR では、実世界で利用可能な多言語の音声翻訳(S2ST)システムの実現を目指して研究を進めている[1]。異なる言語話者同士が PDA (携帯情報端末) のような小型の通信機器を持ち、相手話者の発話した音声、自分の PDA 画面に表示される翻訳テキストと合成音声によって理解しながらコミュニケーションする実現形態を想定している。

著者らは、このようなシステムの研究開発の一環として、同システムを介した日本語話者と英語話者による課題遂行型の模擬対話実験を実施してきた[2]。これまでに、S2ST システムを介した対話音声は、会話例文を読み上げた音声と近い発話スタイルになること[3,4]、また、このような発話スタイルになるためには、話し方の事前教示やある程度の“習熟”も必要であることが分かっている[5]。

本研究の目的の一つは、我々の S2ST システムが“異言語話者間のコミュニケーションにどの程度使えるのか”，という問いに対する一つの目安を示すことである。以前より、TOEIC (Test of English for International Communication) を指標とし、人の能力との対比によるシステム評価を実施してきており、同テスト 600 点の人と同程度の音声翻訳能力があるこ

とが分かっている[6]。しかし、このようなシステムを(英語が全く分からない人が)使ってどの程度のコミュニケーションが可能かは直接的には評価されていなかった。例えば、翻訳結果を理解できない時に、相手に再発話要求すれば、コミュニケーションの“効率”は低下するが、逆に不完全な翻訳結果でも、うまく推測して対話が進む場合もある。

本研究のもう一つの目的は、日本語-中国語の S2ST システムの評価である。ATR では、多言語での音声翻訳システムを大規模な一つの対訳コーパスから自動構築することを目指している。その一つとして、中国語音声認識モデル、日中及び中日の翻訳モデルの開発を進めてきた。従来から開発してきた日英システムと比較しながら日中システムの現時点での到達点を示す。

以下、2節において対話実験に使用したシステムの概要を説明する。3節において対話実験方法を、そして4節で同実験の結果を示す。5節でそれらを考察する。6節でまとめと今後の課題を述べる。

2. S2ST システムの概要

2.1 システムの全体構成

我々の S2ST システムは音声翻訳全体を制御するゲートウエー(以下 GW)を中心に2台の携帯情報端

末(PDA), 日英及び日中の双方向音声翻訳に必要な音声認識(ASR), 機械翻訳(MT), 音声合成(SS)のエンジン群をネットワークで結んだ構成である。PDAとGWとは無線インフラを介して接続する。

日英翻訳を例に全体の動作を簡単に説明する。日本語話者が携帯する PDA に日本語の音声を入力すると, その音声はGWに送られる。GWは送られてきた信号を入力としてまず ASR(J)を呼び出して日本語テキストに変換する。次に得られた日本語テキストを入力として MT(JtoE)を呼び出して, 英語テキストを求め, 最後に, この英語テキストを SS(E)に送ることにより英語の合成音を得て, 英語話者が携帯する PDA に送る。この過程で, 音声認識及び翻訳テキストは, それぞれ日本語話者, 英語話者の PDA に送られ表示される。

英語から日本語への翻訳, 日中翻訳に関してもエンジンがそれぞれの言語(対)用に置き換わるだけで動作は同じである。つまり, 将来的に他の言語(対)エンジンを用意すれば, 多言語間の音声翻訳が同一システム上に実現できる構成になっている。

2.2 構成エンジンの概要

音声認識, 機械翻訳そして音声合成エンジンの概要について述べる。

[音声認識]ATR で開発した ATRASR[7]を使用した。音響モデルは, 読み上げ音声(日本語のみ模擬対話音声)が2割程度含む)から構築した性別依存不特定話者モデルである。言語モデルは, 旅行会話例文集や模擬対話実験データから構成される大規模コーパスから作成した N-gram ベースの言語モデルを用いた[8]。学習データサイズを表1に示す。

表1 音響及び言語モデルの学習データ

		日	英	中
音響 モデル	総話者数[人]	400	384	540
	総発声時間[時間]	38	150	257
言語 モデル	総文数[x10 ³]	852	710	510
	総単語数[x10 ⁶]	8.7	6.1	3.5
	語彙サイズ[x10 ³]	66	44	38

[機械翻訳] ATR の機械翻訳は, 音声認識の言語モデルを構築したのと同じ大規模旅行会話コーパスから自動構築した複数の翻訳エンジンを, SELECTOR [9]と呼んでいる一種の選択器の下に配置した構成をとっている。エンジンには, 統計翻訳エンジン SAT [10]と用例翻訳エンジン HPATR2 [11], 及び挨拶言葉などの定型文では対訳文をそのまま出力する EXACTMATCH の3種類の方式を採用し

ている。

[音声合成]日本語及び中国語は, ATR が開発したコーパスベースの音声合成方式 XIMERA[12]を, 英語に関しては, AT&T が開発した AT&T Labs' Natural Voices™を用いた。

3. 対話実験の設計

3.1 基本方針

我々のシステムは旅行会話全般をカバーすることを目指しているが, 一口に旅行会話といっても幅広く, また地名等の固有名詞は, 語彙になれば翻訳できない。本研究においては, 音声翻訳システムを介して, “十分達成が見込める課題”を, “どこまで正確に且つ効率的に遂行できるか”に焦点を当て, 以下のような方針で課題を設定した。

- ・ 場面はコーパスの中に多く含まれるものから選ぶ。具体的には“買い物”, “ホテル予約”, “ホテル及びレストランでの簡単なトラブル対応”とした。
- ・ 固有名詞の使用は必要最小限にとどめ, 課題達成に必要なキーワードは全てコーパスの中にあるものとする。

3.2 課題設定

本研究の目的は, (模擬)対話実験の結果に基づくシステムの評価である。ここでいう“結果”とは, 対話を通じて, お互いに自分の持っている情報(要望含む)をいかに的確にそして速やかに相手に伝達できたかで測ることができよう。そこで, 伝えるべき情報の数で課題の難易度を, 伝えられた情報の数で達成度を, そして, それに費やした発話数で効率を評価することを考えた。

上記の目的のための課題設定を, 具体例をあげて説明する。例えば買い物の場面において, 難易度が低い課題では, 客役に対して, “エスサイズのアレンジジュースを二杯注文すること。氷無しにすることを願うこと。一杯の値段を確認すること。”のような課題を与える。これは“1:オレンジジュース, 2:エスサイズ, 3:2杯, 4:氷無し”と, 4つの情報を店員役の相手に伝えることになる。店員に正しく伝われば, あらかじめ店員にのみ渡してあるメニューからその値段を客役に伝えて, その課題は終了する。難易度を上げるためには買う品目を多くする。例えば, “エルサイズのアイスコーヒーを二杯とピクルスを抜いたハンバーガーを一つとミックスサンドイッチを一つ注文すること。各々の値段を店員から確認すること”などである。これで客役から店員役へは8つ, 店員から客へは3つの情報を伝えることになる。

3.3 実験方法

対話実験は、前節で述べた S2ST システムを使用して ATR 内の実験室で実施した。なお、音声収録にはヘッドセットマイクを使用した。

[被験者] 一日一組で、日英、日中対話を六日間ずつ実施した。即ち英語、中国語話者は6名ずつ、日本語話者は計 12 名である。英語、中国語話者は日本に比較的長期に滞在中の人たちであるため、日本語が堪能の人も多かった。また日本人も英語が出来る人は多い。過去の実験で、相手言語が分かると、それが対話に影響することが明らかだった [13] ので、相手が発話する時にはヘッドホンからマスキングノイズを流し、互いに相手話者の音声は直接聞こえないようにした。

[教示・練習] 被験者には、実験の前に、音声翻訳されやすい話し方として“明りょうな声で”、“短く簡潔に”話すように教示した。さらに、設定した課題の達成率等を正しく評価するために、あらかじめ指示された情報を過不足なく相手に伝え、その目的から外れる発話はなるべくしないように指示した。

被験者は朝から夕方まで同一ペアで 20 課題以上繰り返し実験する。本論文は全て午後の結果、つまり、既に午前中に 10 対話前後繰り返し対話をして、システムに十分に慣れた後のものである。

[課題] 課題は、先に述べた三つの場面について、難易度の高いもの、低いもの、そして想定場所を日本と海外に設定して実施した。即ち互いが店役と旅行客役を交代で担当する。課題は、日英、日中で基本的に同一のものを実施した(想定場所が海外の課題における通貨単位を除く)。各課題の制限時間は8分とした。一日の実験予定課題数と全体の実験時間との関係で最大に取れる時間に設定したが、途中で対話が大きく停滞しない限り、どの課題も十分に余裕を持って終了できる時間である。

[翻訳の中止] 誤認識されたテキストをそのまま翻訳すれば、誤訳する可能性は高い。そこで、認識結果

を発話者自身が確認した後に翻訳する方式も試みた。誤認識があれば、話者は翻訳を中止して、発話しなおすことが可能となる。以後、この方法を手動モード、認識結果の良し悪しに関わらず必ず翻訳する方法を自動モードと呼ぶことにする。

4. 実験結果

4.1 発話特性と認識及び翻訳結果

表2に対話実験で収集した各言語の発話の基本特性、(実験時の)音声認識率、そして翻訳主観評価に基づく発話の正訳率を示す。ここで正訳率とは、発話の元の内容をほぼ過不足なく相手に伝えることが出来ると思われる発話の頻度である。文法の間違いなどは修復可能であると評価者に判断されれば許容する。

手動モードは、話者によって翻訳が中止された発話を除いた結果である。つまり日本語から英語(JtoE)及び中国語(JtoC)では、2割強の発話が捨てられることで、単語認識率が95%を超え、約8割の発話が、文法の良し悪しを除けば正しく翻訳されたことになる。同様に英語から日本語(EtoJ)は3割弱の発話が無駄になるが、正訳率は約15%向上する。中国語から日本語(CtoJ)は半分近くも捨てなければならないが、その結果約5割だった正訳率が6割強となる。

比較のために、従来のテストセットによるオフライン音声翻訳実験(認識出力を翻訳)の結果も示した。“タイピスト”は、音声認識率100%を模擬するために音声認識の代わりにタイピストを導入した対話実験データから選んだもの、“会話文読み上げ”は、前述の ATR コーパスの会話例文から選んだテストセット文を読み上げた音声データの結果である。今回の実験の発話長は、タイピスト実験時よりも大幅に短くなり、会話例文のそれとほとんど同じになった。

4.2 発話エントロピーと認識及び翻訳誤りとの関係

発話が長くなれば、誤認識そして誤訳する可能性は高まる。さらに発話の長さが同じでも、キーワードと

表2 対話実験音声の発話特性と認識及び翻訳評価結果

言語方向	今回の実験								従来のテストセット				
	自動モード				手動モード				タイピスト		会話文読み上げ		
	JtoE	JtoC	EtoJ	CtoJ	JtoE	JtoC	EtoJ	CtoJ	JtoE	EtoJ	JtoE	EtoJ	CtoJ
平均発話長	7.3	7.1	6.3	5.4	6.8	7.0	5.9	5.4	9.9	9.3	7.0	6.0	5.5
パープレキシティ	25	20	39	91	25	24	31	63	23	29	19	25	34
単語正解精度[%]	89	93	78	74	96	97	88	84	91	86	95	88	84
発話正解率[%]	66	75	51	44	82	87	64	55	60	48	82	69	62
正訳率 翻訳実行率 [%]	69 -	78 -	60 -	51 -	81 77	80 78	76 71	64 55	54 -	58 -	75 -	72 -	65 -

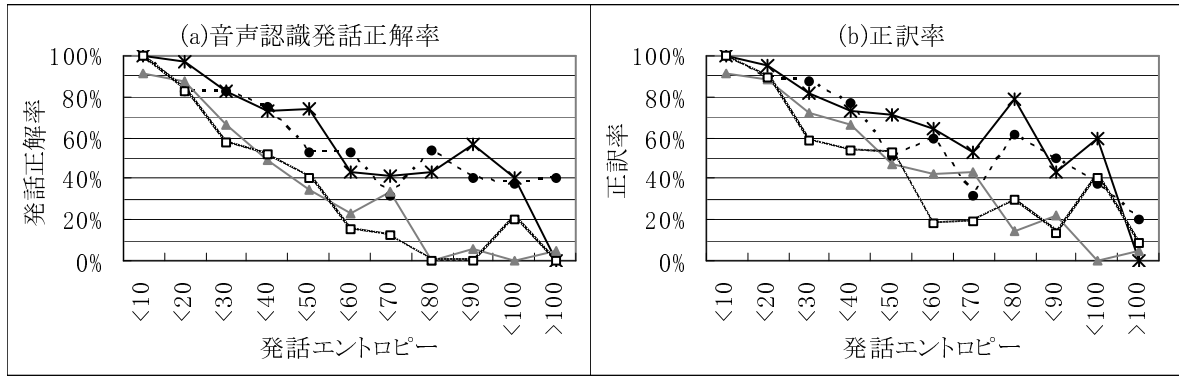


図1 発話エントロピーと認識及び翻訳成否との関係

なる名詞など、コーパス中の出現頻度が低い単語をより多く含むほうが認識や翻訳は難しくなる。そこで一つの発話全体のエントロピーと認識及び翻訳の成否の関係を図1(a), (b)に示した。(a)が音声認識の発話正解率、(b)が翻訳の正訳率で、自動モードの結果のみ示した。発話エントロピーが高いほど、認識そして翻訳誤りを生じ易いことは明らかである。言語別に見ると、JtoE及びJtoCは音声認識性能が比較的高く、発話エントロピーが30以下(発話長では6単語前後)で話せば8割強、40以下(7単語前後)でも8割弱は正しく翻訳出来る。EtoJ, CtoJでは発話エントロピー20以下ではJtoE, JtoCとほとんど違いはないが、30以下で7割と6割、40以下では、6割前後の正訳率となり、言語間で差が出てくる。

4.3 課題達成率と発話効率

表2及び図1より、現状システムの音声翻訳の性能は、JtoEとJtoCが同程度、次にEtoJ、最も低いのがCtoJということが分かった。この差は実際のコミュニケーション(情報伝達の良し悪し)に影響を与える。表3は全体としてどの程度の課題達成(情報の伝達)ができたかをまとめたものである。

ここで**課題達成率(TAR)**、**全発話効率(UE_A)**及び**有効発話効率(UE_V)**を、対話システムの評価法[14,15]を参考に、以下のように定義した。

$$TAR = \frac{TI_T + TI_C}{AI_T + AI_C}$$

$$UE_A = \frac{AU_T + AU_C}{TI_T + TI_C}, \quad UE_V = \frac{VU_T + VU_C}{TI_T + TI_C}$$

表3 課題達成率と発話効率

		課題難易度(伝達必要情報数)					
			<4	4-5	6-8	>8	計
日英	実施課題数	自動 手動	87	1110	811	1412	4140
	課題達成率(TAR) [%]	自動 手動	87 100	94 89	91 81	81 83	86 85
	発話効率(UE)	自動 UE _A	3.2	2.3	2.7	2.5	2.5
		手動 UE _A UE _V	2.6 1.8	3.4 2.4	3.5 2.4	2.6 1.7	3.0 2.0
日中	実施課題数	自動 手動	97	1112	711	1116	3846
	課題達成率(TAR) [%]	自動 手動	80 94	88 96	82 77	74 65	79 75
	発話効率(UE)	自動 UE _A	3.8	2.7	3.6	3.0	3.2
		手動 UE _A UE _V	4.3 2.0	3.0 2.3	3.6 2.3	4.1 2.3	3.8 2.3

ここで、AIは実施した課題の中で相手に伝える必要があった情報の総数、TIは実際に伝えられた情報の総数で、添え字は旅行者役Tと店員Cを意味する(以下同じ)。AUは発話総数、VUは翻訳が実行された発話(有効発話と呼ぶ)の総数である。いずれも対話の冒頭や終了時にあるような定型的な挨拶発話や、明らかな発話の失敗は除いた。

課題達成率TARは互いに伝えるべき情報の中でどれだけの情報が伝達出来たかを表す。伝達が必要な情報数が多くなれば、制限時間内に全ての情報を伝えられなかったり、伝達をミスしたりする可能性が高まるので、TARは低下する。

発話効率UEは、一つの情報を伝えるのに必要な互いの発話数の平均を表す。そして、手動モードにおいて、有効発話だけを考慮するのが有効発話効率UE_V、翻訳が中止された発話を含めた全発話で計算するのが全発話効率UE_Aである。

日英の手動モードにおいて、全体平均のUE_Aが3.0、UE_Vが2.0ということは、話者が二人で3発話する間に、1発話の翻訳が中止されて、一つの情報が伝わるということの意味する。日中の場合には、一つの情報が伝わる間に4発話弱費やされ、1.5発話の翻訳が中止される。日英と日中では、中国語の中止される発話数の多さが、特に手動モードで伝達が必要

な情報数が多い場合の課題達成率に大きな影響を与えている。

4.4 発話エントロピーと発話効率の関係

対話実験を観察していると、話者(ペア)によって、課題を比較的スムーズにこなしていくペアと、そうでないペアが見られた。その違いは音声翻訳の成否に大きく依存するが、4.2節(図1)で示したように、発話エントロピーが大きい(発話長が長い)ほど誤認識や誤訳は生じやすい。一方、一つの発話により多くの情報を含めたほうが情報伝達の効率は良くなる。

図2は手動モード時の各話者ペアの正訳発話エントロピーと発話効率の関係を示したものである。

ここで**正訳発話エントロピー(CUE)**とは、“正しく伝わったと推定される”一発話あたりのエントロピーで、次式で定義する。

$$CUE_A = \frac{\sum UE_{Tcorrect} + UE_{Ccorrect}}{AU_T + AU_C}$$

$$CUE_V = \frac{\sum UE_{Tcorrect} + UE_{Ccorrect}}{VU_T + VU_C}$$

AU, VU 及び添え字 T, C は 4.3 節の定義と同じである。UE_{T,C|correct} は正訳と評価された発話のエントロピーである。その総和を全発話数で割った CUE_A を全発話の正訳発話エントロピー、有効発話で割った CUE_V を有効発話の正訳発話エントロピーと呼ぶ。図2は、CUE_A に全発話効率 UE_A を、CUE_V に有効発話効率 UE_V を対応させ、CUE_A-UE_A を●(日英)、▲(日中)で、同様に CUE_V-UE_V を○、△でプロットし、対応する話者ペアのそれらを線で結んである。

全体的に右上がりの相関があり、正しく翻訳された発話エントロピーが大きいほど発話効率も良いという妥当な結果となった。線の長さは翻訳の中止頻度に対応する。全体的に日英に比べて、日中のほうが長く、日中の翻訳中止頻度の多さ、そしてそれが結果的に全発話効率を低下させていることがこの図からも分かる。特に日中においてが CUE_V が 30 を超える二組は、発話長が長いいため、認識が成功して翻訳に渡せれば、発話に含まれる情報数が多いために UE_V は良くなる(それでも日英と同等)が、翻訳中止(誤認識)の頻度は高くなるため、CUE_A 及び UE_A は大幅に悪くなっている。

5. 考察

5.1 日英と日中の差

各エンジンの性能差の原因などについては本論

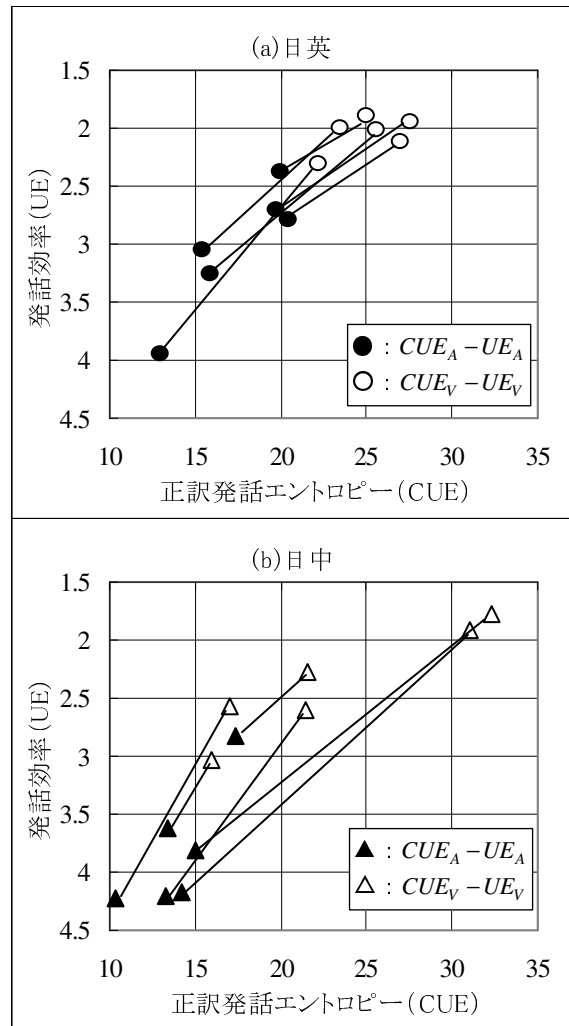


図2 正訳発話エントロピーと発話効率の関係

文では議論しないが、結果として、中国語の音声認識率が低いことが日中対話の課題達成率や発話効率が低くなった最大の要因と考えられる。手動モードにおける話者による翻訳中止率の高さに加えて、翻訳が実行された発話の認識率も他言語より低く、それが正訳率の低下を招いている(表2参照)。

5.2 自動モードと手動モード

表3を見ると、自動モードと手動モードでは、全体の課題達成率はほとんど同じだった。しかし、自動と手動には、課題達成率の分布に違いが見られる。伝達が必要な情報数が少ない場合には、手動モードの“確実性”が勝り、達成率が高い。一方、伝えるべき情報数が多くなると、無駄になる発話数が多くなる手動モードは時間切れになる可能性が高まり、達成率は逆転する。確かに誤認識すれば誤訳する確率は高いが、仮に誤訳してもそこにキーワードさえ含まれていれば、それから相手の意図などを推測できる

場合も多く、“何らかの情報が送られる”自動モードのほうが伝わる情報の数という面では有利に働いたようである。

5.3 発話効率

図2(a), (b)の CUE_V-UE_V (白色プロット)に着目すると、有効発話の正訳発話エントロピー(CUE_V)が25あたりから、その発話効率(UE_V)が、2発話前後で頭打ちになっていることが分かる。つまり我々の音声翻訳システムの発話効率は、誤認識発話を無視すれば、“2発話で1情報程度”といえる。一つの情報要求発話に一つの情報提示発話というパターンだけではなく、一発話に二つ以上の情報を含めれば、その分、誤訳の確率が高まり、確認発話や再発話要求等が増え、結果的にこの値になっていると考えられるが、詳細な発話分析は今後の課題である。

5.4 話者の話し方の課題進行への影響

全発話を考慮した発話効率(UE_A)の差は誤認識による翻訳中止頻度の大小で生じてくる。つまり、誤認識発話の頻度を減らすように、常に“短く簡潔な”発話を心がけた話者ペアの効率が結果的に高くなっている。中国語の音声認識率が低い日中対話においてそれが顕著であり、唯一 UE_A が3発話以下のペアは、お互いの発話長がかなり短い(日 4.9, 中 4.8)。日本人側も協力して情報を分割して提示、また要求していた。つられて中国語話者も短く話せるようになり、互いの誤認識そして誤訳が減り、発話効率が日中対話の中では顕著に高くなったと思われる。一発話の情報量が減ることで、有効発話効率はそれほど高くないが、無駄になる発話を減らし、少しずつ着実に対話を進めていることで、結果的に課題達成率も高かった(94%)。

6. まとめと今後

ATR で開発中の日英及び日中音声翻訳システムを使った課題遂行型対話実験を実施し、同システムのコミュニケーションツールとしての性能評価を試みた。日英と比較すると、日中は中国語の音声認識性能が低く、それが全体の課題達成率を低下させたが、翻訳が実行された有効発話だけでみると、日英、日中ともに“2発話で1情報”という発話効率であることがわかった。また、一発話を短くし、情報を分割することで、誤認識を減らせた話者が効率よく対話を進められたことも分かった。

今後は今回の知見などを元に、“適切な発話”へ話者を誘導するための自動教示手法を検討するとともに、より自然なコミュニケーションのための音声翻

訳性能の向上も必要であると考えている。

謝辞

本研究は情報通信研究機構の研究委託「大規模コーパスベース音声対話翻訳技術の研究開発」により実施したものである。

参考文献

- [1] Yamamoto, S. “Toward Speech Communications beyond Language Barrier - Research of Spoken Language Translation Technologies at ATR -,” Proc. of ICSLP2000, Vol. IV, pp.406-411, 2000.
- [2] Takezawa, T. and Kikui, G., “Collecting Machine-Translation-Aided Bilingual Dialogues for Corpus-Based Speech Translation,” Proc. of EUROSPEECH2003, pp. 2757-2760, 2003.
- [3] 水島, 竹澤, 菊井, “翻訳システムを介した対話音声の発話スタイルについて-自然発話, 朗読発話との関係-”, 第3回話し言葉の科学と工学ワークショップ講演予稿集, pp.135-142, 2004.
- [4] 水島, 竹澤, 菊井, “音声翻訳システムを介した対話の評価-誤認識及び誤訳が対話に及ぼす影響-”, 言語処理学会第11回年次大会, 229-332, 2005.
- [5] 水島, 竹澤, 菊井, “実環境における音声翻訳システムを介した対話実験-実験室環境との発話スタイルの比較-”, 日本音響学会秋季研究発表会, 1-P-17, 2005.
- [6] 中村, 佐々木, 菊井, 清水, “音声翻訳技術の展開”, 情報の科学と技術, Vol.55, No.8, pp.345-350, 2005.
- [7] 伊藤, 葦莉, 實廣, 中村, “音声認識統合環境 ATRASR の概要と評価報告”, 日本音響学会秋季研究発表会, 1-P-30, 2004.
- [8] H. Yamamoto, S. Isogai, and Y. Sagisaka, “Multi-class composite N-gram language model,” Speech Communication, vol. 41, pp. 369-379, 2003.
- [9] Y. Akiba, T. Watanabe, and E. Sumita, “Using language and translation models to select the best among outputs from multiple MT systems,” Proc. COLING, pp. 8-14, 2002.
- [10] T. Watanabe and E. Sumita, “Example-based decoding for statistical machine translation,” Proc. MT Summit IX, 2002.
- [11] K. Imamura, H. Okuma and E. Sumita, “Practical approach to syntax-based statistical machine translation,” Proc. MT Summit X, pp. 267-274, 2005.
- [12] H. Kawai, T. Toda, J. Ni, M. Tsuzaki, and K. Tokuda, “XIMERA: A new TTS from ATR based on corpus-based technologies,” Proc. 5th ISCA Speech Synthesis Workshop, 2004.
- [13] 水島, 竹澤, 菊井, “翻訳システムを介した音声対話における相手話者音声と翻訳テキスト表示の影響について”, SLP-52, pp.99-106, 2004.
- [14] J. Glass, J. Polifroni, S. Seneff, and V. Zue, “Data collection and performance evaluation of dialogue system: The MIT Experience”, Proc ICSLP2000, Vol. IV, pp.1-4, 2000.
- [15] 橋田ほか, “DiaLeague-自然言語処理システムの総合評価-”, 人工知能学会誌, Vol.12, No.3, pp.390-134, 1997.