

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

水島 昌英[†] 竹澤 寿幸[†] 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

音声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調べた。情報の過不足を伴う誤訳が生じた際、翻訳テキストや相手話者音声がマスクされ合成音声だけしか聞こえないと、相手話者に再発話を要求する頻度が大幅に増えた。このことから、翻訳テキストや相手話者音声は、誤訳で生じた情報の欠落を補う効果があることが明らかになった。

Effects of Audibility of Partner's Voice and Visibility of Translated Text in Machine-Translation-Aided Bilingual Spoken Dialogues

Masahide MIZUSHIMA[†] Toshiyuki TAKEZAWA[†] and Genichiro KIKUI[†]

[†]ATR Spoken Language Translation Research Laboratories, Kyoto 619-0288, Japan

This paper reports on the effects of audibility of partner's voice and visibility of translated text in Machine-Translation-Aided Bilingual Spoken Dialogues. When MT errors lacked important information, subjects who could hear only the synthesized voice from the translation required re-speaks ("Pardon?" or "Again, please." etc.) more frequently than subjects who were able to see the translated text and hear the partners' original voices. This fact shows that people can compensate for the lack of information from MT error more easily by using displays of translated text and by listening to the original voices of their partners.

1. まえがき

ATR では、実世界で利用可能な多言語の音声対話翻訳(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互いにPDA(携帯情報端末)のような小型の通信機器を持ち、相手話者の音声を翻訳した結果をPDA画面に表示すると共に合成音声で出力するといった実現形態を想定している。

我々は、このような翻訳システムの研究開発の一環として、実際に翻訳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対話コーパスMAD(Machine-translation-Aided-bilingual spoken Dialogue corpus)の構築を進めている[2]。これまでに、当該コーパスを分析することで、翻訳システムを介した対話音声与会話文を読み上げた音声に近い発話スタイルになることや、実験前の教示が言語表現に大きな影響を与えることなどを明らかにしてきた[3],[4]。

本研究では、音声翻訳システムを介した対話に影響を与えると思われる二つの要因について検討を

試みた。一つは相手話者音声の影響である。MADでは、日本語話者と英語話者で実験を行っている。被験者の選定において「英語の分からない日本語母語者」、「日本語の分からない英語母語者」という条件を出しているが、日本語被験者は学校教育による英語教育を受けていない人はまれであろう。また英語被験者は、日本で生活していることから、多少の日本語は理解できると考えられる。相手話者の話していることが翻訳システムを介さずに直接理解できるとすれば、仮に翻訳(あるいは認識)に誤りがあったとしても、対話がそのまま進んでしまう可能性があり、システム性能の評価等に影響を与えるであろう。また、ATRでは、日英だけではなく、日中等の多言語での翻訳システムの開発も進めている。日英で得られた知見等を他の言語対に適用する際、相手話者の言語の理解度がどの程度対話に影響するかは無視できない可能性がある。

もう一つは、ディスプレイに表示される翻訳テキストの影響である。先に述べたように想定しているシステムでは、PDAや携帯電話等の画面に翻訳されたテキストを表示すると共に合成音声でそれを読み上

げる。テキスト表示だけでも情報の伝達は可能であるが、画面が小さく見づらい場合や一つの端末でやり取りをする場合などでは、合成音声の有効性が高まると考えられる。さらに通常は合成音声で内容を理解し、数字や名前やわかりにくかった表現等についてのみ表示テキストで確認するように出来れば、S2ST を使ったコミュニケーションがよりスムーズになるともいえよう。

本研究では、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を調べるために、それらをマスクした対話実験を実施し、分析を試みた。

以下、2節において、対話実験の概要を説明する。3節において実験の結果を示し、4節でそれらを考察する。5節でまとめを述べる。

2. 対話実験の概要

2.1 実験システム構成

今回の実験においても、従来の MAD 実験同様に、機械翻訳部に重点を置き、音声認識部の代わりにタイピストが発話内容を書き起こし、翻訳部に入力する方法で実施した。実験システム構成の概要を図1に示す。英語話者音声を英語タイピストが聞き取り、コンピュータに入力する。入力されたテキストを英日機械翻訳部で翻訳し、翻訳された日本語文を日本語話者の目前に置かれた小型PCのディスプレイに表示すると共に音声合成部で合成した日本語音声をヘッドホンから再生する。日本語話者音声も同様に、翻訳結果をテキストと合成音声で英語話者に提示する。これらを繰り返して、翻訳システムを介した対話が行われる。

小型PCのディスプレイには、自身の発話を書き起こされたものと相手話者の発話の翻訳結果が図2のように対話の進行に従って表示される。相手話者発話の書き起こしは表示されない。一発話の長さにもよるが、おおよそ3ターン程度の発話が同時に表示される。

日英、英日翻訳には、句構造等のパターンに基づく翻訳 HPAT[5] と文を単位とする用例に基づく D3[6] に最良翻訳選択器[7] を組み合わせたシステムを使用した。日本語音声合成は XIMERA[8]、英語音声合成には、AT&T Labs' Natural Voices™ を使用した。

相手話者の肉声を聞こえなくする場合は、一方の話者が発話中に、もう一方の話者のヘッドホンからマスキングノイズを再生した。翻訳テキストを提示し

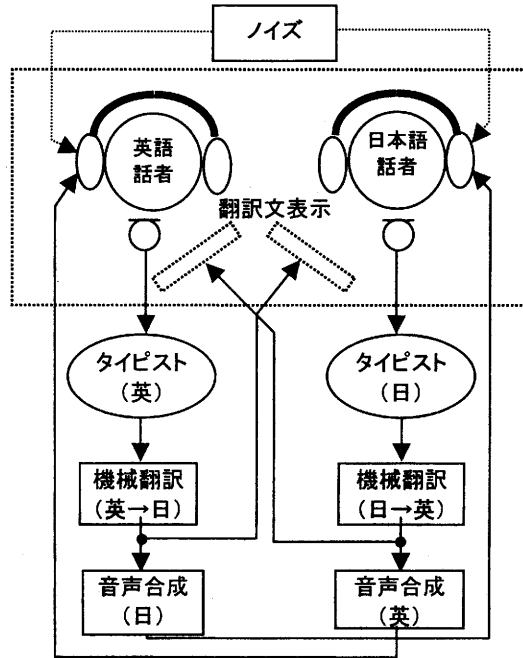


図1 実験システムの構成

英翻>観光局ございます。
日認>すみません、ホテルを探しているのですが。
英翻>何名様でしょうか。
日認>三人です。
英翻>宿泊の予定は市でしょうか____…。
日認>駅に近いホテルを希望しているのですが。

図2 対話実験時のテキスト表示の例
(日本語話者用)

ない場合は、ディスプレイを紙で覆った。

2.2 対話実験の実施

実験は、新たに追加した実験条件、即ち相手話者の音声と翻訳テキストをマスクする以外は、従来の MAD 実験と同等な条件で実施した。一連の対話実験の 5 回目の実験であるので、MAD5 と呼ぶことにする。以下にその概要を述べる。

【被験者】日本語話者8名、英語話者8名で、いずれも本実験は初めての人である。原則的に相手の言語は分からないことを条件としているが、日本語話者は英語教育を受けており、また英語話者は、日本在住の人であることから、ある程度は、理解できると思われる。ただし、個々の被験者がどの程度の理解度を持つかは調べていない。

【教示】MAD4 と同様な教示をした[4]。

1. 大き目の声で明瞭に話す。
2. 1回の発話は10秒以内とする。
3. 時々誤りが発生するが、確認や再発話をする
ことにより、対話を続ける。
4. 伝えたい情報を分割して、短く簡潔に話す。

【実験条件】相手話者の音声聞こえる(条件 v)又は聞こえない(条件 n)、翻訳テキストが見られる(条件 d)又は見られない(条件 h)の組み合わせになる。全部で4つの条件になるが、一人の被験者につき、テキスト表示条件(d 又は h)は固定し、相手話者音声の提示(v 又は n)のみ、条件を振った。また対話中の英日話者の条件は等しくした。

【課題設定】MAD4と同じ16パターンの課題を設定した。空港でのトラブルやレストランでの注文、観光地での道案内等、海外旅行時に体験しそうな様々な場面を設定している。各実験条件別で実施したのべ課題数を表1にまとめた。

【発話数】発話総数は、日本語が1438、英語が1466、総数2904発話になった。各実験条件別の発話総数を表1にまとめた。

3. 実験結果の分析

3.1 発話の長さへの影響

実験条件毎に、一発話あたりの平均単語数、同平均文数に加え、日本語側のみ、文の単文率[9]を調べた。結果を表2に示す。平均単語数は、翻訳テキスト表示がない場合(条件 h)で、日本語、英語共に有意に少なくなっている(危険率1%)。相手話者音声の条件(v 又は n)では、若干聞こえないほうが少なくなっている傾向があるが、有意差はなかった。日本語における単文率も同様に、翻訳テキスト表示に関してのみ有意に表示がないほうが単文率が増加した(危険率1%)。一発話あたりの平均文数では、条件vより条件nのほうがやや短く、英語側では条件dよりhのほうが短い。英語側のvdとnh以外に統計的な有意差はなかった。

3.2 誤訳への対応

相手話者の発話が正しく翻訳され、正しく合成音声で再生されていれば、相手話者の音声や翻訳結果テキスト表示の有無の影響は比較的小さいと考えられる。即ち、翻訳や合成の間違いがあつた場合、これら追加情報の影響が顕在化すると予想される。そこで、影響が特に大きいと思われた機械翻訳における誤訳が次の話者の発話にどのような影響を与えるかを調べてみた。

3.2.1 誤訳の抽出

表1 各実験条件における実施課題数と発話総数

	課題	日本語	英語	合計
vd	30	337	354	691
nd	27	364	385	749
vh	25	406	404	810
nh	26	331	323	654
合計	108	1438	1466	2904

表2 各実験条件における発話の長さ

		平均単語数	平均文数	単文率
日本語	vd	9.06	1.26	79.3%
	nd	8.84	1.23	77.6%
	vh	8.14	1.32	86.0%
	nh	7.85	1.26	87.8%
英語	vd	10.13	1.98	
	nd	9.84	1.83	
	vh	8.37	1.72	
	nh	9.03	1.70	

誤訳の影響を調べるためには、まず翻訳結果から誤訳を抽出する必要がある。ここでの実験のようなタスク遂行型の対話の場合、翻訳文の自然さより、情報が過不足なく伝えられるかで評価することが重要と思われる。従来、機械翻訳結果を評価する方法はいくつか提案されている。一つは評価者が翻訳文について数段階で主観的に評価する方法である。例えば Sumita ら[10]は、翻訳結果を以下の4つのカテゴリで評価した。

- A: 訳文は情報的にも文法的にも問題がない[完全訳]
- B: 訳文は重要度の低い情報が欠けていたり文法的に間違いがあつたりするが、容易に理解できる[部分訳]
- C: 訳文は broken だがなんとか理解できる[可能訳]
- D: 訳文からは重要な情報を正しく理解できない[不可訳]

また、両語に精通した人が作った正解訳文と翻訳結果文との類似性で評価する方法もある。例えば音声認識の単語正解精度のように、正解訳文と翻訳結果文とをDPマッチングにより比較し、単語単位の一致率を計算する方法[11]や、正解訳文と翻訳結果文との一致率を n-gram 単位で計算する方法[12]などが提案されている。

前者の方法では、情報の過不足と文法的間違いが総合的に判断される。後者は、情報の過不足を客観的に評価出来るが、正解訳文を作成する手間がかかる。そこで今回は簡便に情報の過不足の有無を評価する方法として、発話単位で、以下の条件を一つでも満たすものを誤訳として抽出した。評価は、

2名の英語に精通した評価者(日本人)が行った。

挿入: 不要な単語の挿入がある。

削除: 必要な単語が削除されている。

置換: 必要な単語が不必要な単語に入れ替わっている。

疑問肯定否定の間違い: 疑問文を肯定文に又はその逆に訳している等。

文法等の間違い: 係り受け, 語順, 文法の大きな間違いや複数の意味を持つ語の過度に不適切な訳。

一発話あたりの出現数は考慮せず, 一つでも誤訳と判断された箇所があれば, 誤訳が生じた発話としてカウントした。挿入, 削除, 置換, 疑問肯定否定の間違いの四つを情報の過不足の誤訳と考え, それらと文法等の間違いを分けて以下分析を試みた。

表3に抽出した誤訳が生じた発話の頻度を示す。小計は挿入, 削除, 置換, 疑問肯定否定いずれかを含む発話の頻度であり, これらを重複して含む発話もあることから, 四つの合計よりも小さい値になっている。また文法等のみは, 文法等に間違いのあるもののうち, これら四つの誤訳を含まない発話の頻度である。条件によって, 誤訳を生じる発話の頻度にほとんど違いがないことが分かる。

3.2.2 誤訳後の発話の分析

誤訳が生じると, 文章自体として意味を取りにくかったり, それまでの対話の流れと食い違う内容だったりすることが多い。そのため, 次に発話する被験者の発話に大きな影響が出ると考えられる。誤訳後の発話を, 以下の四つに分類して分析を試みた。

確認発話: 相手の意図を汲み取れなかったり, 不確実であった場合に, 相手に対して, 再発話を要求したり, 内容を確認する発話。

例1: 日> すいませんが, もう一度お願いします。

例2: 日> 扇子を一本プレゼント用ですか?

適切発話: 相手の元の発話に対して, 適切と思われる発話。

不適切発話: 相手の元の発話に対して, 不適切と思われる発話。

例: 日本国外の空港カウンター。日本人話者の荷物が出てこない場面。

英> How many bags are missing.

英訳> …… バッグが見当たらないのです。

日> どこで待ってればいいですか。

表2 各実験条件における発話単位の誤訳の頻度

	挿入	削除	置換	疑・肯 否	小計	文法等 のみ	総計
vd	5.1%	6.7%	9.8%	4.2%	23.4%	15.5%	38.9%
nd	2.3%	6.4%	11.3%	3.3%	21.8%	18.8%	40.6%
vh	3.8%	6.5%	12.6%	4.7%	25.1%	14.4%	39.5%
nh	4.9%	7.0%	9.5%	5.0%	22.2%	16.7%	38.8%

この例では, 英語話者が“行方不明のバッグは何か?”と質問しているが, 正しく訳されていないため, 日本語話者はそれに答えられず, 別の質問を開始している。

不明発話: 発話自体はおかしくないが, 話題が切り替わるなどして, 正しく伝わったかどうか第三者には判断できない発話。

例: 京都市内で日本語話者が英語話者に道案内をしている場面。

日> バスは京都駅から発車します。

日訳> The bus stops at Kyoto Station.

英> Where is Kyoto station.

英訳> 京都はどこですか。

日> この道をまっすぐ行きます。三つ目の交差点を右に曲がると京都駅です。

日訳> I'll go down this street _ _ _ turn right at the 00th Kyoto Station.

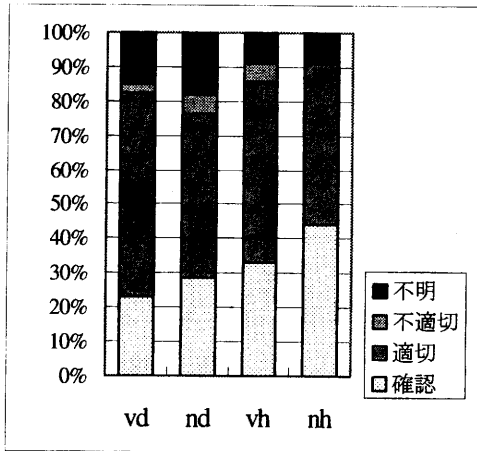
英> I understand . How much is the bus fare to Kiyomizu temple.

英訳> 分かりました。清水寺までバス代はいくらですか。

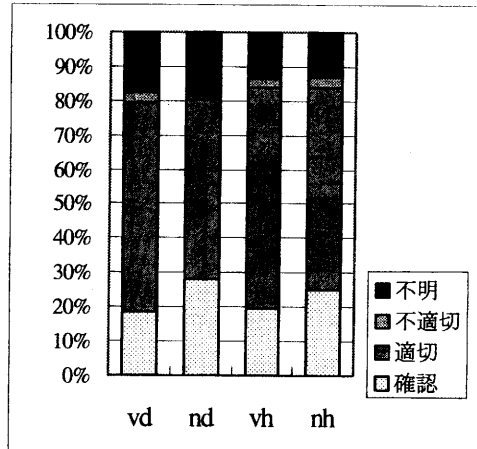
日本語話者の京都駅までの行き方の説明において, “三つ目の交差点で(右に曲がる)”が訳されていないが, 英語話者は納得して別の質問をしている。正しく伝わっていない可能性が高いが, 例えば相手話者の元の音声から理解した可能性も否定できないので, このような場合は不明発話に分類した。

評価は, 誤訳が生じた発話直後の発話に対して, 対話ログから判断した。必要に応じて, 対話の前後関係も判断材料にした。

図3(a), (b)に分析結果を示す。(a)は情報の過不足がある場合と考えられる挿入, 削除, 置換, 疑問肯定否定4つの誤訳が生じた直後の発話を分類したもので, (b)は, それ以外の文法間違い等のみが生じた場合のものである。図3(a)から明らかなように, 条件 vd, 即ち翻訳テキストが表示され, 相手話者の音声も聞こえる場合には情報の過不足があっても確



(a) 挿入・削除・置換・疑問等の誤訳後



(b) 文法等のみの誤訳後

図3 誤訳直後の発話の分類

認発話は少なく、条件 nd や vh、即ち相手話者の声か翻訳テキストどちらか一方がマスクされると、確認発話が増える。さらに、条件 nh、即ち相手話者の声と翻訳テキストいずれもマスクされると、最も確認発話が多くなる。また、条件 h のほうが条件 d よりも不明な発話が少ないことも分かる。一方文法等のみの場合には、比較的條件による差が小さい。

3.2.3 確認発話の詳細

前節で相手話者の声や翻訳テキストがマスクされると、誤訳後に確認発話が増えることがわかった。そこで、確認発話について、さらに細かく見てみた。前節で定義したように、確認発話には、相手に再発話を要求するものと、自ら相手の発話内容を繰り返して、相手にその正否を確認するもの(復唱確認発話と呼ぶことにする)がある。図4(a), (b)は図3における誤訳後の確認発話の頻度を、新たにこの二つに分類しなおして示したものである。さらに(c), (d)は、誤訳の有無と確認発話の頻度を比較するために、誤訳がない場合の確認発話の頻度を示したものである。(c)は挿入・削除・置換・疑問等の情報の過不足がない場合、(d)はそれに加えて文法等の誤訳もない、即ち正訳後の確認発話の頻度である。図4(a)を見ると、翻訳テキストが見えない条件 h においては、再発話要求の割合が多く、逆にテキストが見えると復唱確認が多い。この傾向は文法間違いでも同様だが、割合を見ると、復唱確認が全体的に増加し、特に条件 vd ではほとんどが復唱確認で、再発話要求はまれになっている。さらに誤訳がないと実験条件に関わらず、確認発話そのものが大幅に減り、特

に再発話要求はほとんどないことが分かる。

4. 考察

4.1 発話の長さ

3.1 の結果から翻訳テキストの情報がないと発話が短くなることが分かった。発話が短いほど翻訳文も短くなることが期待される。テキスト表示がない場合に発話が短くなることは、誤訳直後の再発話要求が増えることから分かるように、合成音声だけで長い翻訳文を理解することの困難さを被験者が感じていることを示していると考えられる。

4.2 誤訳の抽出

表3から分かるように、情報の過不足が生じる誤訳は 20%強、文法等を含めると約 40%の誤訳率であった。翻訳テキストが見えない条件(vh, nh)では、一発話における単語数が少なくなることから誤訳が減ることも期待されたが、実験条件間に大きな誤訳率の違いは生じなかった。実験前に短く簡潔に話すことを教示されているため、既に被験者は誤訳が生じにくい比較的簡易な表現を使っている可能性がある。

4.3 誤訳後の反応

図3(a), (b)のように誤訳、特に情報の過不足の誤訳(挿入、削除、置換、疑問肯定否定の間違い)が生じると、翻訳テキストの表示や相手話者音声の有無によって、その直後の発話行動に差が生じることが分かった。どちらも提示されている場合(vd)は、誤訳があってもそのまま対話を進める割合が約 78%であるのに対し、どちらも提示されていない場合(nh)で

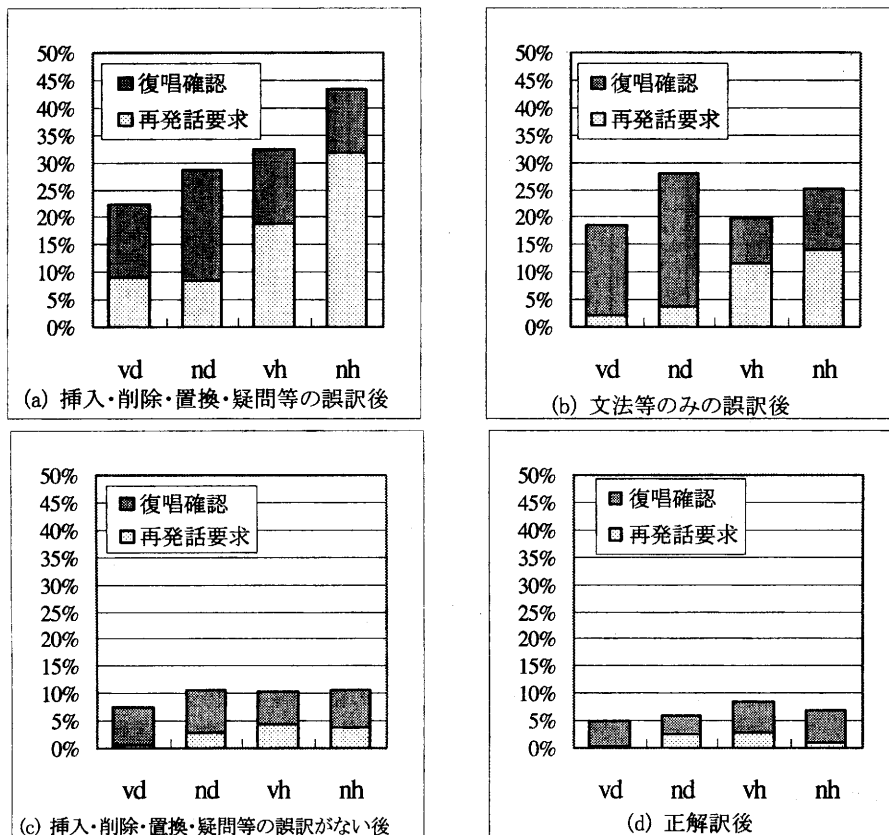


図4 誤訳の有無と直後の確認発話の頻度

は、約 57%に減る。どちらか一方が提示されない場合は、その中間的頻度になった。このことから誤訳により情報の過不足が生じた場合、翻訳テキストと相手話者音声の情報が共に相手話者の意図の理解に重要な役割を果たしていることが分かる。以下、誤訳が生じた直後の発話が適切発話になった例をあげる。

例1:ホテルの予約の場面(条件 vd)

日> ダブルルームとツインルームがありますが、どちらがよろしいですか。

日訳> There is a twin room and a twin room which would you prefer.

英> I'd prefer a double room.

英訳> ダブルルームが良いのですが。

“ダブルルーム”が訳されていないが、日本語話者の音声から直接理解したものと思われる。

例2:日本食レストランでの注文の場面(条件 nd)

日> 生のお魚は大丈夫ですか。

日訳> Is fish of all right.

英> Fish is ok. What's the difference between the four thousand and five thousand course.

英訳> 魚で構いません。四千人から五千コースの違いは何時間ですか。

日> 一品五千元の方が多くなっています。

コース料理の違いを聞いている英語話者の発話が正確に訳されていないが、日本語話者はきちんと対応している。翻訳テキストを見るほうが合成音声で聞くよりも意図を推理する、即ち過不足の部分を補う(あるいは無視する)ことが容易になる要因として、合成音声は一回しか聞けないこと(実験条件でそのようにした)、テキストは過去の2, 3個の対話ログも見られるため、全体から推理しやすいことその他、誤訳を含む不自然な文章は、時系列の流れ(いわゆる Left

to right)で理解することが困難になることも一因と考えられるが、詳細はさらに検討が必要である。

条件 nh では、確認発話が多いだけでなく、不適切な発話もほとんどなく、不明発話も他の条件より少なくなっている。小さな誤訳もあいまいにせず、対話を慎重に進めているともいえよう。実際、対話中に起きる誤解、例えばホテルの宿泊日数の確認ミス等は、画面表示のない条件 vh, nh では一回も起きなかったのに対し、vd では30課題で7回、nd では27課題で2回、合計9回確認出来た(なお対話ログからは分からない誤解はチェックしていない)。誤訳を含む不自然な翻訳テキストから過度に推測して対話を進行してしまうと、誤解が生じる恐れがあることを示している。

文法等のみの誤訳後の発話では、特に条件 vh, nh の確認発話が大幅に減り、実験条件の間で大きな差が生じていない。このことから、本実験のような比較的簡易なタスク遂行型の対話においては、情報の過不足があることが、より大きく対話の進行に影響を与えているといえよう。

4.4 確認発話

図4(c), (d)から、誤訳が生じなければ確認発話の頻度は大幅に少なく、また実験条件間の差も小さいことが分かる。誤訳が生じなければ、合成音声だけでも、対話の進行に大きな支障がないことを示している。なお確認発話のうち、相手の発話を繰り返す「復唱確認」が多くの割合を占めているが、これは実験に際して、自分の意図が相手に正確に伝わっているか、相手の意図を自分が正確に受け取ったかどうかを、お互いに確認しあいながら対話を進めるように指示しているため、ある程度確証がある内容であっても、復唱して相手に確認を求める発話も多いと思われる。

誤訳がない時の再発話要求の発生事例は以下のようであった。

例 1: 外国の観光案内所で、日本語話者がホテルの予約をしている。ホテルの予約完了後、接客係の英語話者の発話より(条件 nd)。

英> Is there anything else today.

英訳> 今日何か他の物はありますか。

日> もう一度お願いします。

英語話者は、“(本日は)他になにか御用はございますか?”と聞いたと思われる。英訳は、直訳として間違いではないが、厳密には“用事”と訳すべきところを“物”と訳したために日本語話者には通じなかつ

たようである(なお、この程度の訳は全て正訳としてしている)。同様の状況が vd でも数回起きているが、その場合は、全て適切な対応がされていた。

例 2: 日本のみやげ物屋で、英語話者が日本のお土産をいろいろ探している場面(条件 nd)。

英> I would also like a tanuki mask.

英訳> たぬきマスクにします。

日> もう一度お願いします。

“tanuki mask”は、“たぬきのお面”と訳されれば通じたと思われる。あるいはこれは翻訳テキスト表示のない条件での対話だったが、テキストがあれば、「たぬきマスク→たぬきのお面」と解釈出来たかもしれない。音声で聞くと聞きなれない表現であったため、理解できなかったのであろう。これらの例のように誤訳とまではいえなくても、直訳や普段あまり使われない表現に訳された場合に、特に合成音声だけでは理解されないことがあるものの、図4(d)から明らかのようにその頻度は非常に少なく、他の誤訳に比較すれば、影響は小さいと見てよいであろう。

次に誤訳が生じた後の確認発話について考察する。先に述べたように画面や相手話者音声がないと特に情報の過不足があったときに確認発話が増えることが分かっているが、図4(a)でその増え方を細かく見ると、復唱確認の頻度はあまり変わらず、再発話要求の頻度が大きく違うことが分かる。特に翻訳テキスト表示のない vh, nh の再発話要求が多い。これは、翻訳テキスト表示があると、確認発話が減るだけでなく、確認する際にも、相手の発話した内容を出来る限り推理して、“あなたのおっしゃったことは〇〇ですね?”のように相手に確認を求められるが、表示がないとそれが難しく、単に“もう一度お願いします”のように再発話を要求せざるを得ないものと思われる。以下に、誤訳から復唱確認をしている例をあげる。

例: 日本で日本語話者が英語話者に対して道案内をしている場面(vd)

日> ツーブロック目の角に本屋さんがあります。

日訳> Two sees the bookstore on a block eye corner.

英> So the bookstore is two blocks away.

英訳> 書店は二ブロック先です。

日> そうです、二本通りを通過します。

最初の日本語話者の説明は正確に訳されていないが、日本語話者音声の“ツーブロック”と訳の

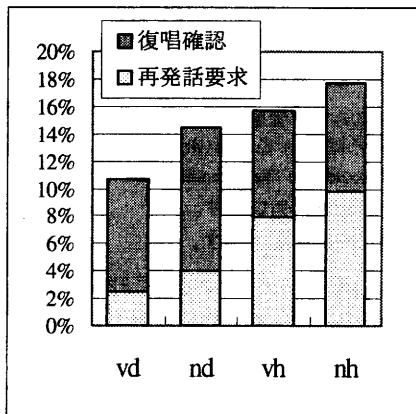


図5 発話全体の中での確認発話の頻度

“bookstore”を組み合わせたのか、英語話者はほぼ正確に推理して、復唱確認をしている。確認を求められた日本語話者は“そうです。”と答えればよく、対話が進行しやすい。誤訳が生じた際の翻訳テキスト表示、あるいは相手話者音声の有効性をここでも示している。

文法等のみの誤訳後の確認発話も、翻訳テキスト表示がある場合の再発話要求が少ない傾向は同じだが、全体的に復唱確認発話の割合が多くなっており、この結果からも、情報の過不足が生じていなければ、文法等に間違いがあっても相手の発話内容を推理することが比較的容易に出来る事が分かる。

図5は発話全体の中での確認発話の頻度である。復唱確認発話の頻度は実験条件でほとんど違いがなく8から10%程度であった。一方再発話要求は、翻訳テキスト表示があることで5,6%, 相手話者の音声がかえることで2%ほど減っている。これが、今回用いた翻訳システムの誤訳発生率(情報過不足が20%強、文法等を含めると約40%)における確認発話(再発話要求)の頻度という尺度で見たときの、翻訳テキスト表示と相手話者音声の対話の進行に与える総合的な効果を示す数字であるといえる。

5. むすび

音声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、それらをマスクした状態での対話実験の結果から分析を試みた。その結果、特に情報の過不足を伴う誤訳が生じた際、過不足を補い相手話者の意図を推理して対話を進めるために翻訳テキストや相手話者音声の情報を有効に使っている事が分かった。一方、

誤訳が生じない場合には、相手の意図を確認する発話の頻度はそれらの提示の有無で大きな差はなく、誤訳が生じていない時は、合成音声だけでも比較的スムーズに対話は進行することが分かった。

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, 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] Takezawa, T. and Kikui, G., "A Comparative Study on Human Communication Behaviors and Linguistic Characteristics for Speech-to-Speech Translation," Proc. of LREC2004, pp.1589-1592, 2004.
- [4] 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.
- [5] Imamura, K., "Application of Translation Knowledge Acquired by Hierarchical Phrase Alignment for Pattern-based MT", Proc. of TMI-2002, pp74-84, 2002.
- [6] Sumita, E. "Example-based machine translation using DP-matching between word sequences, Proc. of ACL-2001, pp.1-8, 2001.
- [7] Akiba, Y., Watanabe, T., and Sumita, E., "Using Language and Translation Models to Select the Best among Outputs from Multiple MT systems", Proc of COLING-2002, pp.8-14, 2002.
- [8] Kawai, H., Toda, T., Ni, J., Tsuzaki, M. and Tokuda, K., "XIMERA: A New TTS from ATR Based on Corpus-Based Technologies", 5th ISCA Speech Synthesis Workshop, 14th-16th June 2004, Carnegie Mellon University (to appear), 2004.
- [9] 丸山, 柏岡, 熊野, 田中, "節境界自動検出ルールの作成と評価", 言語処理学会第九回年次大会発表論文集, pp.517-520, 2003.
- [10] Sumita, E., Yamamoto, S., Paul, M., Kashioka, H., Ishikawa, K. and Shirai, S., "Solutions to Problems Inherent in Spoken-language Translation: ATR-MATRIX Approach", Proc. of MT Summit '99, pp.229-235, Sep. 1999.
- [11] Thompson, H., "Automatic Evaluation of Translation Quality: Outline of Methodology and Report on Pilot Experiment", Proc. of Evaluators' Forum, pp.215-223, 1991.
- [12] Papineni, Kishore and Roukos, Salim and Ward, Todd and Zhu, Wei-Jing, "Bleu: a Method for Automatic Evaluation of Machine Translation", Proc. of ACL-2002, pp.311-318, 2002.