社団法人 情報処理学会 研究報告
IPSJ SIG Technical Reports

2006－SLP－64（7）
2006／12／21

# コミュニケーション効率に基づく課題遂行型音声対話の評価

竹澤 寿幸†‡ 水島 昌英§ 清水 徹†‡ 菊井 玄一郎§

† 独立行政法人 情報通信研究機構 知識創成コミュニケーション研究センター
〒619-0288 京都府「けいはんな学研都市」光台二丁目2番地2

‡ ATR 音声言語コミュニケーション研究所 〒619-0288 「けいはんな学研都市」光台二丁目2番地2

§ 日本電信電話株式会社 NTT サイバースペース研究所

E-mail: † ‡ {toshiyuki.takezawa, tohru.shimizu}@{nict.go.jp, atr.jp}

§ {mizushima.masahide, kikui.genichiro}@lab.ntt.co.jp

**あらまし** 課題遂行型対話において課題達成に必須となる項目を伝達する観点からコミュニケーション効率を数量化する手法を提案する．音声翻訳システムを用いた対話実験結果と，システムの代わりに隠れた通訳者が翻訳を行ういわゆる Wizard of Oz (WOZ)による対話実験結果を示し，隠れた通訳者に対する音声翻訳システムの性能の相対的な数量化を行う．さらに，話し手の言語的な振舞い方とシステム性能の関係を議論する．システムの利用者は課題達成に必須となる項目の数は減らさずにそれ以外の言語表現を減らす傾向があることや，相手とのやり取りの仕方を制御することで見かけ上の効率を改善しようとしていることが知見として得られた．

**キーワード** 音声コミュニケーション，自然言語インタフェース，音声認識，インタラクティブシステム，ヒューマンファクタ．

# Method for Evaluating Task-Oriented Spoken Dialogs Based on Communication Efficiency

Toshiyuki TAKEZAWA†‡ Masahide MIZUSHIMA§ Tohru SHIMIZU†‡ and Genichiro KIKUI§

† Knowledge Creating Communication Research Center, National Institute of Information and Communication Technology 2-2-2 Hikaridai, Keihanna Science City, Kyoto, 619-0288 Japan

‡ ATR Spoken Language Communication Research Labs. 2-2-2 Hikaridai, Keihanna Science City, 619-0288 Japan

§ NTT Cyberspace Laboratories, Japan

E-mail: † ‡ {toshiyuki.takezawa, tohru.shimizu}@{nict.go.jp, atr.jp}

§ {mizushima.masahide, kikui.genichiro}@lab.ntt.co.jp

**Abstract** We propose a method for measuring communication efficiency from the viewpoint of conveying essential information in a task-oriented dialog. We show the results of a dialog experiment using speech-to-speech translation systems and one using the Wizard of Oz method, which was carried out using hidden interpreters instead of a speech-to-speech translation system. We also present a relative score for the performance of the speech-to-speech translation system which was obtained by measuring the performance of the machine against that of human, i.e., hidden interpreters. Finally, we discuss the relationship between users' linguistic behavior and system performance. We found that users of the system tended to make shorter utterances without decreasing the number of essential items needed to achieve a task and also to improve the transmission efficiency by controlling the strategy of dialogs.

**Keyword** Speech communication, natural language interfaces, speech recognition, interactive systems, human factors.

## 1. まえがき

音声対話翻訳を始めとする音声言語によるコミュニケーション支援に関する研究を進めている．コミュニケーション支援はユーザビリティの観点から最終的には評価されるべきものであるが，扱う課題の難しさや利用者の個性等の多くの要因の影響を受けるため，

§ ATR 音声言語コミュニケーション研究所滞在中になされたものである。

普遍的に議論することは難しい．そこで，ユーザビリティに関するISO[1]やJIS[2]の規格に則り，効果(effectiveness)，効率(efficiency)，利用者満足度(user satisfaction)の三つに分けて考えることにする．

効果は課題達成率あるいは対話成功率で数量化でき，課題の難しさと音声言語システムのカバー率の二つに分けて考えればよい．旅行に関する協調的な対話に限定すれば，セールストークやビジネストークのように文化によるストラテジの差はなく，確認等の際にやり取りする項目の順序に好みの差があるだけなので，課題の難しさは課題達成に必須となる項目の数で数量化できる．課題達成に必須となる項目数が増えれば難しくなることは直感的に明らかであり，実験的にも確認されている[3]．カバー率は場面や状況に依存する固有名詞等の取り扱いに関する課題として切り分けることができる．カバー率の高い音声言語システムが役に立つのは間違いないが，カバーされている領域内でどれだけ効率が良いかは別である．

効率は課題達成に要する時間あるいは発話数で数量化できる．利用者満足度はシステムの操作性や応答時間，さらに利用者の個性等の多くの要因の影響を受け，アンケート調査で数量化する以外の手段は知られていない．そこで，音声言語システムのコミュニケーション支援能力を数量化する第一歩として，効果，効率，利用者満足度のうち，本稿では効率に焦点をあてる．

人間と機械の音声対話システムについてはGlass等が必要な項目をシステムに伝達する効率の観点から評価尺度を提案している[4]．現在の音声言語処理技術が人間同士のコミュニケーション支援にどの程度役に立つのか数量化するために，Glass等の人間から機械への一方向の手法を双方向の情報伝達に拡張することにより，コミュニケーション効率を数量化する手法を考案した．翻訳システムや隠れた通訳者を介した日英・日中対話実験結果を示し，隠れた通訳者に対する音声翻訳システムの性能の相対的な数量化を行う．さらに，話し手の言語的な振舞い方とシステム性能の関係を考察する．

## 2. 評価尺度

課題達成に必須となる項目を伝達するための発話というものがある．例えば，客が「アイスコーヒー」「エルサイズ」「ホットドック」「マスタード抜き」等の希望する内容を伝達し，店員がメニューから品物の値段を伝達するようなものである．そのような課題達成に必須となる項目を用いて，コミュニケーション効率を数量化する尺度として，提示項目密度(PD: Provided Density)，伝達項目密度(TD: Transmission Density)，伝達効率(TE: Transmission Efficiency)を次のように定義する．

$$\text{提示項目密度}(PD)=\frac{\text{発話に含まれていた伝達項目総数}}{\text{伝達項目提示発話総数}}$$

$$\text{伝達項目密度}(TD)=\frac{\text{相手に伝わった伝達項目総数}}{\text{伝達項目提示発話総数}}$$

$$\text{伝達効率}(TE)=\frac{\text{相手に伝わった伝達項 目総数}}{\text{発話に含まれていた伝 達項目総数}}$$

提示項目密度は話し手の言語的な振舞いに着目して新たに定義した指標である．課題達成に必須となる項目を伝達するための発話を伝達項目提示発話と名付ける．提示項目密度は伝達項目提示発話に平均的に含まれる必須項目の数を表す．伝達項目密度は聞き手の立場からの指標であり，伝達効率は音声翻訳システムや隠れた通訳者のようなコミュニケーション支援の性能に関わる指標である．

文献[4]でGlass等は音声対話システムの効率に関わる二つの指標，質問密度(QD: Query Density)と概念効率(CE: Concept Efficiency)を提案している．「質問」は本稿の伝達項目提示発話に対応し，「概念」は本稿の課題達成に必須となる項目に対応する．質問密度は利用者の質問に対し音声対話システムが理解した平均的な概念の数を表す．概念密度は利用者が発話した概念の数に対し音声対話システムが理解した概念の数の比である．本稿では，対話に参加するすべての話者を考慮に入れる．したがって，*TD*は*QD*を一方向から双方向に拡張したもの．*TE*は*CE*を一方向から双方向に拡張したものに対応する．

## 3. 実験システム

携帯型多言語音声コミュニケーション支援プラットフォーム[5]を用いて，旅行会話全般を対象とした大規模コーパスベース音声対話翻訳技術の研究成果[6]をもとに，日英および日中の実験用音声翻訳システムを構築した．分散処理型の構成となっており，音声認識，機械翻訳，音声合成の各処理はサーバで行う．

### 3.1. 音声認識

日本語，英語，中国語の音声認識はともにATRASR[7]を使用した．音響モデルはMDL-SSSアルゴリズム[8]により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，旅行対話に関する大規模コーパス[9]を用いて作成したマルチクラス複合*N*-gram[10]の言語モデルである．音響モデル及び言語モデルの訓練データサイズを表1に示す．

表 1　音響及び言語モデルの訓練データサイズ

| | | 日本語 | 英語 | 中国語 |
|---|---|---|---|---|
| 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 | 852k | 710k | 510k |
| | 総単語数 | 8.7M | 6.1M | 3.5M |
| | 語彙サイズ | 66k | 44k | 38k |

### 3.2. 機械翻訳

音声認識の言語モデルを構築したものと同じ大規模コーパス[9]から自動構築した複数の翻訳エンジンを利用するマルチ・エンジン翻訳技術[11]の研究成果を採用した．具体的には統計翻訳エンジン SAT[12]と用例翻訳エンジン HPATR2[13]を利用した．その複数の翻訳エンジンの結果から SELECTOR[14]と呼ぶ選択器が良いものを選んで出力する．さらに，挨拶表現などの定型表現は対訳をそのまま出力するようにした．

### 3.3. 音声合成

日本語及び中国語の音声合成は，大規模コーパスベース音声対話翻訳技術の研究成果[6]からコーパスベース音声合成 XIMERA[15]を採用した．英語の音声合成は AT&T Labs' Natural Voices ™ を利用した．

## 4. 対話実験

### 4.1. 条件

実験結果の数値化を容易にするために，実験室における課題遂行型対話とした．システムの訓練用コーパスに多く含まれている場面から「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」を選び，被験者にわかりやすい課題を設定した．固有名詞の使用は必要最小限度にとどめ，それらを含め課題達成に必須となる語句はすべてコーパス中にあるものとした．

### 4.2. 音声翻訳システムを用いた対話実験

1 日 1 組で，日英，日中対話を 6 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，話者に重なりはない．話者は 1 台ずつ携帯情報端末(PDA)を持ち，PDA のディスプレイには，自身の発話の認識結果と相手話者の翻訳結果のみを表示させた．

英語，中国語話者は日本語が堪能な人も多かった．また日本人も英語が分かる人は多い．直接聞こえた相手の音声から発話内容を理解してしまうと，それが対話ないし課題達成に影響を与えることが別の実験により確認されている[16]．そこで，その影響を排除するために，相手の発話中にヘッドホンからマスキングノイズを再生し，互いに相手話者の音声が直接聞こえないようにした．また，誤認識されたテキストをそのまま翻訳すると誤訳する可能性が高いため，認識結果を発話者自身が確認した後に翻訳する方式を採用した．

被験者には，実験の前に，明瞭な声で，短く簡潔に話すよう教示した．さらに，指示書に書かれた目的を達するために必要な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繰り返し実験する．本稿で扱う結果はすべて午後の結果，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．なお，日英と日中で同等の課題を実施し，各課題実施の制限時間は 8 分とした．

図 1 に音声翻訳システムを用いた対話実験の様子を示す．

図 1　音声翻訳システムを用いた対話実験風景

### 4.3. 隠れた通訳者を介した対話実験

日中対話を実施した．通訳者を 2 名手配し，各々日本語から中国語，中国語から日本語の翻訳を担当させた．通訳者は，話者とは隔離された別室において，話者の音声を聞き，その翻訳テキストをキーボードから入力する．すると，翻訳テキストとその合成音声がパソコンを介して話者に提示される．話者には，人間が翻訳していることは知らせない．

実験は合計 6 組実施した．翻訳システムを介した対話実験を含めて話者に重なりはない．1 台のタッチパネルディスプレイを二人の話者が共有する形態とした．

翻訳システムを介した対話実験と同等な教示に加えて，発話時間を 1 発話 8 秒以内に制限した．翻訳システムを介した対話実験の発話と大きく乖離することを避けるためである．話者が発話ボタンを押してから最大 8 秒間のみ収録される．通訳者には，音声が途中で切れた場合には，話者に対して再発話する指示メッ

セージを送出するよう指示した．

なお，比較を容易にするために，翻訳システムを介した対話実験と同じ課題を実施した．

図2に隠れた通訳者を介した対話実験の様子を示す．図3に隠れた通訳者の操作風景を示す．

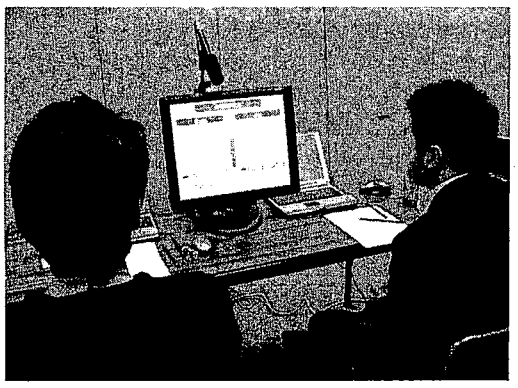

図2　隠れた通訳者を介した対話実験風景

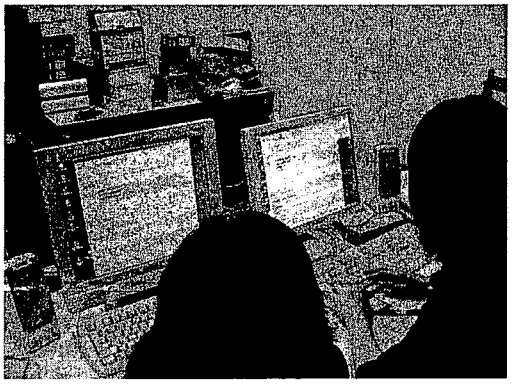

図3　隠れた通訳者の操作風景

## 5. 実験結果と考察

### 5.1. 基本特性とシステム性能

表2に翻訳システムを介した対話実験(S2ST実験)と隠れた通訳者を介した対話実験(WOZ 実験)のデータ量，及びS2ST実験における音声認識率，そして主観評価に基づく発話の正訳率を示す．正訳率とは，発話の内容をほぼ過不足なく相手に伝えることができるとみなせる発話の頻度である．文法の軽微な間違いなど聞き手が修復可能であると評価者が判断すれば正訳としている．なお，この音声認識率と正訳率は翻訳が実行された発話で計算した値である．翻訳が実行された発話を有効発話と呼ぶことにする．

表2　対話の基本特性及びシステム性能

| | S2ST実験 | | | | WOZ実験 | |
|---|---|---|---|---|---|---|
| 言語方向 | 日英 | 日中 | 英日 | 中日 | 日中 | 中日 |
| 課題数 | 40 | 46 | 40 | 46 | 67 | 67 |
| 発話数 | 326 | 419 | 392 | 572 | 363 | 366 |
| 平均発話長 | 6.9 | 6.7 | 6.0 | 5.4 | 10.5 | 9.2 |
| パープレキシティ | 32 | 28 | 38 | 92 | 25 | 139 |
| 単語正解率 | 96% | 97% | 88% | 84% | — | — |
| 発話正解率 | 82% | 87% | 64% | 55% | — | — |
| 翻訳中止率 | 23% | 22% | 29% | 45% | — | — |
| 正訳率 | 81% | 80% | 76% | 64% | — | — |

S2ST実験について，表2によれば，日英方向と日中方向は被験者が自分の認識結果を確認することにより，20%強の発話が翻訳されずにキャンセルされているものの，翻訳が実行された有効発話に対する日本語音声認識の単語正解率はいずれも95%を超え，正訳率もそれぞれ約80%となっていることがわかる．しかしながら，英日方向は被験者が自分の認識結果を確認することにより，約30%の発話が翻訳されずにキャンセルされているにも関わらず，有効発話に対する英語音声認識の単語正解率は90%弱であり，正訳率は75%強である．中日方向にいたっては，半分近くの発話が翻訳されずにキャンセルされているにも関わらず，有効発話に対する中国語音声認識の単語正解率は85%弱であり，正訳率は65%弱である．

### 5.2. コミュニケーション効率

表3にコミュニケーション効率に関わる三つの指標，提示項目密度(PD)，伝達項目密度(TD)，伝達効率(TE)の計算結果を示す．これらはすべて有効発話中の伝達項目提示発話に対して計算した値である．

表3　実験結果：コミュニケーション効率

| | S2ST実験 | | WOZ実験 |
|---|---|---|---|
| | 日英・英日 | 日中・中日 | 日中・中日 |
| 提示項目密度(PD) | 1.45 | 1.43 | 1.59 |
| 伝達項目密度(TD) | 0.97 | 0.88 | 1.46 |
| 伝達効率(TE) | 0.67 | 0.62 | 0.91 |

表3によれば，提示項目密度はいずれの条件でも大きな差はない．伝達効率はWOZ実験に比べてS2ST実験では値が小さくなっている．

コミュニケーションのための音声対話では，「いらっしゃいませ」のような挨拶発話や確認発話等がしばしば行われる．表3の実験結果を求める際の伝達項目提示発話としては，対話の相手に対して新たに課題達成に必須となる項目を発話したもののみを考慮するよ

うにした．しかしながら，翻訳が実行された有効発話には，伝達項目提示発話以外にも挨拶発話や確認発話が含まれる．しかも，確認発話の数は音声翻訳システムの性能により変化し，たいていシステム性能が劣化すると確認発話の数は増える傾向がある．そこで，すべての有効発話に対して相手に伝わった平均的な伝達項目の数とその相対比を計算してみた．その計算結果を表4に示す．

表4　有効発話に対する計算結果

| | S2ST 実験 | | WOZ 実験 |
|---|---|---|---|
| | 日英・英日 | 日中・中日 | 日中・中日 |
| 有効発話に対し相手に伝わった平均的な伝達項目の数 | 0.45 | 0.34 | 0.74 |
| 相対比 | 60% | 46% | 100% |

表3の伝達効率は伝達項目提示発話のみで求めた値であるので，見かけ上のコミュニケーション効率に近く，その条件下では日英・英日システムと日中・中日システムの差は小さい．しかしながら，挨拶発話や確認発話を含めたすべての有効発話で計算した表4の相対比は日英・英日システムと日中・中日システムの差が大きい．

表3と表4はいずれもそれぞれの条件に対して平均的な値を求めたものである．話者数は各条件に対していずれも6ペアずつであり，必ずしも多くはないが，それでも話者によって認識率の差や話し方の差が見受けられる．そこで，特に話者の言語的な振舞いに相当する提示項目密度について，話者の認識率との関係を図4に示す．JEが日英方向，JCが日中方向，EJが英日方向，CJが中日方向をそれぞれ表す．話者毎に個別に有効発話に対する単語認識率及び提示項目密度を求め，対応する話者ペアを破線あるいは点線でつないで示す．

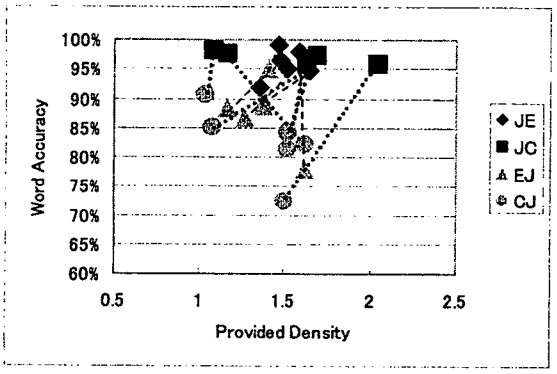


図4　音声認識率と提示項目密度

図4によれば，濃い色でプロットした日英方向と日中方向の提示項目密度は1から2まで広がっていることがわかる．その間で音声認識率の違いはあまりない．一方，薄い色でプロットした英日方向と中日方向の提示項目密度は1から1.6までの範囲であり，提示項目密度が小さくなるにつれて音声認識率が良い傾向が見られる．英語話者と中国語話者はシステムの制約にあわせて話し，日本語話者は比較的システムの制約が厳しくないために話し方の好みの差が出やすいと考えられる．

そこで，さらに，システム性能に関係の深い伝達効率について，話者の認識率との関係を図5に示す．図4と同様に，JEが日英方向，JCが日中方向，EJが英日方向，CJが中日方向をそれぞれ表す．話者毎に個別に有効発話に対する単語認識率及び伝達効率を求め，対応する話者ペアを破線あるいは点線でつないで示す．

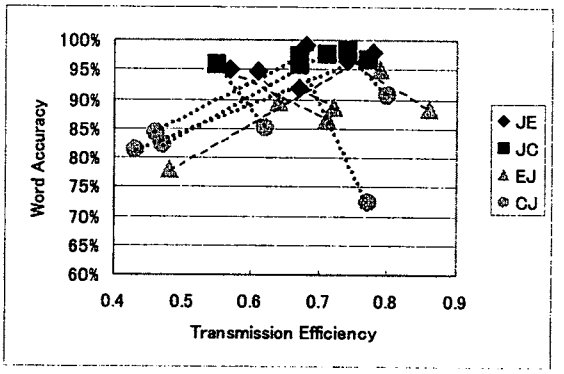


図5　音声認識率と伝達効率

図5によれば，濃い色でプロットした日英方向と日中方向の伝達効率は0.55から0.8までの範囲である．一方，薄い色でプロットした英日方向と中日方向の伝達効率は0.4近くから0.9近くまでの間に広がっている．旅行に関する協調的な対話では，システムの制約に応じて対話参加者は対話の戦略を変えることがある．例えば客がCDプレーヤーを購入したいことがわかり，色が黒，白，シルバーの3種類あった場合に，「どのような色がお好みですか」と尋ねて相手がなかなか希望の色が認識されないように見えると「黒はいかがですか」というように順にYes/No形式で答えるように誘導することで対話が課題解決に向かって円滑に進むことが多い．午前中に10対話前後繰り返し対話することで，話し方のみならずシステム性能にも慣れた被験者のうち接客業等の経験があるような人は対話の戦略を変えることでシステム性能が十分でない場合であっても伝達効率を上げることができるようであった．

## 6. 議論

### 6.1. 話し手の言語的な振舞い

表2によれば，S2ST実験はWOZ実験に比べて平均発話長は短くなっている．しかしながら，表3によれば，提示項目密度はいずれの条件でも大きな差はない．話し手は，課題達成に必須となる項目の数は減らさずに，それ以外の言語表現を減らしているといえる．

### 6.2. システム性能とコミュニケーション効率

表2によれば，日英方向と日中方向の認識・翻訳性能はほぼ同じであるが，英日方向は中日方向より認識・翻訳性能ともに良い．そのような性能差のために，表4に示したように，翻訳を実行した有効発話を基準とした場合に，WOZ実験と比べたシステムの相対性能は，日英・英日システムが60%，日中・中日システムが46%となった．ただし，旅行に関する協調的な対話では，相手とのやり取りの仕方を制御することで，見かけ上の効率に相当する伝達効率を改善しようとしている．結果的に，表3に示すように，日英・英日システムと日中・中日システムで平均的には伝達効率に大きな差はない．

なお，上記の議論は翻訳を実行した有効発話に基づくものであり，表2に示すように中国語の翻訳中止率が日本語や英語に比べて大きい点に注意する必要がある．つまり，翻訳がなされなかった発話を含めた場合には，日英・英日システムと日中・中日システムの性能差は有効発話を基準とした場合よりも大きくなる．

## 7. むすび

課題遂行型対話において課題達成に必須となる項目を伝達する観点からコミュニケーション効率を数量化する手法を考案した．翻訳システムや隠れた通訳者を介した日英・日中対話実験結果を示し，隠れた通訳者に対する音声対話翻訳システムの性能の相対的な数量化を行った．システムの利用者は課題達成に必須となる項目の数は減らさずにそれ以外の言語表現を減らす傾向があることや，相手とのやり取りの仕方を制御することで見かけ上の効率を改善しようとしていることが知見として得られた．今後は音声対話翻訳に限定することなく，コミュニケーション支援のための音声言語処理技術の研究開発をさらに進める．

## 文　献


[1] ISO (International Standardization Organization), "ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on usability," 1998, http://www.iso.org.

[2] JIS Z8521, "人間工学―視覚表示装置を用いるオフィス作業―使用性についての手引き," 1999.

[3] 水島昌英，竹澤寿幸，清水徹，菊井玄一郎，"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，" 情報処理学会研究報告，2006-SLP-60(10), pp. 49-54, 2006.

[4] J. Glass, J. Polifroni, S. Seneff, and V. Zue, "Data collection and performance evaluation of dialogue system: The MIT experience," Proc. ICSLP, vol. IV, pp. 1-4, 2000.

[5] 葦苅豊，木村法幸，清水徹，"携帯型多言語音声コミュニケーションプラットフォーム，" 日本音響学会秋季研究発表会講演論文集，1-2-22, pp. 43-44, 2006.

[6] 中村哲，菊井玄一郎，佐々木裕，清水徹，"大規模コーパスベース音声翻訳技術と全体性能の評価，" 日本音響学会春季研究発表会講演論文集，2-1-9, pp. 87-88, 2006.

[7] 伊藤玄，葦苅豊，實廣貴敏，中村哲，"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講演論文集，1-P-30, pp. 221-222, 2004.

[8] T. Jitsuhiro, T. Matsui, and S. Nakamura, "Automatic generation of non-uniform context-dependent HMM topologies based on the MDL criterion," Proc. EUROSPEECH, pp. 2721-2724, 2003.

[9] G. Kikui, E. Sumita, T. Takezawa, and S. Yamamoto, "Creating corpora for speech-to-speech translation," Proc. EUROSPEECH, pp. 381-384, 2003.

[10] H. Yamamoto, S. Isogai, and Y. Sagisaka, "Multi-class composite *N*-gram language model," Speech Communication, vol. 41, pp. 369-379, 2003.

[11] 隅田英一郎，パウル・ミヒャエル，今村賢治，大熊英男，"多言語音声翻訳のためのマルチ・エンジン翻訳技術，" 言語処理学会第12回年次大会発表論文集，E4-2, pp. 853-856, 2006.

[12] T. Watanabe and E. Sumita, "Example-based decoding for statistical machine translation," Proc. MT Summit IX, pp. 410-417, 2002.

[13] K. Imamura, T. Watanabe, and E. Sumita, "Practical approach to syntax-based statistical machine translation," Proc. MT Summit X, pp. 267-274, 2005.

[14] 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.

[15] H. Kawai, T. Toda, J. Ni, and M. Tsuzaki, "XIMERA: A new TTS from ATR based on corpus-based technologies," Proc. 5th ISCA Speech Synthesis Workshop, pp. 179-184, 2004.

[16] 水島昌英，竹澤寿幸，菊井玄一郎，"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，" 情報処理学会研究報告，2004-HI-09(19)/2004-SLP-52(19), pp. 99-106, 2004.