

多言語対応ロボットの対話応答実証実験 (ロボットが東京駅でお客さまを案内)



菅澤 学^{*1}



坂入 整^{*2}

Talking Humanoid Robot Verification in Tokyo Station (Technical Verification of a Multilingual Talking Humanoid Robot)

Manabu SUGASAWA^{*1}, Sei SAKAIRI^{*2}

^{*1} Researcher, Frontier Service Development Laboratory, Research and Development Center of JR EAST Group

^{*2} Chief Researcher, Frontier Service Development Laboratory, Research and Development Center of JR EAST Group

Abstract

We conducted a technical verification of Hitachi-made humanoid robot that recognizes English, Chinese and Japanese in JR East Travel Service Center of Tokyo station. We verified the performance of language switching by the voice of conversation partner, the performance of noise reduction. Also, we confirmed some technical problems to realize smooth conversation, especially processing time. As a result, the success ratio of language switching is about 60%, the dialog failure due to the noise reduction failure is about 1%, the interval time required between the conversation partner's utterance and the robot's utterance is about 3 seconds.

●**Keywords:** Voice interaction technology, Artificial intelligence, Humanoid robot, Multilingual dialogue, Question answering

1. 目的

少子高齢化に伴う経営環境の変化や訪日外国人の急増への対応策の1つとして、ロボットおよび人工知能(AI)による接客案内の実現に向けた研究開発を行なっている。すでにAIを用いたコールセンター業務支援システムの実導入が決定しているが、駅はコールセンターと比べて雑音が多いため、音声認識技術に着目し、メーカーが開発中のロボットを用いて実験を行なった。

本実験の目的は、多言語案内を視野に入れた言語切り替え性能および雑音が多い駅環境で重要となる雑音除去性能について既存の技術水準を確認すること、そして人とロボットのスムーズな対話を成立させるために解決しなければならない課題を把握することである。

1・1 音声に基づく言語切り替え性能の確認

対話言語の切り替え方式次第でユーザビリティは大きく変わってくる。今回は英語、中国語、日本語の3言語について、タッチパネル上で言語を選択するのではなく、ワンアクション少なくして利用者の音声に応じて言語が切り替わる仕組みの実用性を確認した。

1・2 雑音除去性能の確認

駅構内は列車の走行音や案内放送の音声、雑踏が発する音など対話を阻害する音が多い。マイクで集音した音の中から利用者の声だけを抽出する雑音除去性能は欠かせない技術であるため、その性能を確認した。

1・3 スムーズな対話実現への課題確認

利用者が質問をしてからロボットが回答するまでの処理時間は、実用化を考えると極めて重要である。駅環境でどこまでスムーズなやり取りが可能か、対話ログから分析を行う。

2. 実施概要

2・1 利用機材および実施場所

本実験では1・1実施に向けた作り込みが比較的容易であり、1・2の性能が高いとされる、株式会社日立製作所が公共スペースなどにおいて接客・案内サービスを行うことを目的に開発中のヒューマノイドロボット「EMIEW3」(図1)を利用することとし、対話が成立しないケースの発生に備えてロボット近傍に補助要員を置いた。実施場所には訪日外国人向けの旅行案内窓口である東京訪日旅行センター(訪旅セ)を選定した。

2・2 質問応答リスト

外国人旅客からよく質問される事柄について、訪旅セ係員とグランスタの案内係員にヒアリングを実施し、216種類の質問とその答えのリストを英中日の3言語分作成して、EMIEW3登録した。そのうち大部分が一問一答形式の順路に関する質問応答であるが、「空港への行き方」「新幹線乗り場への行き方」などは更問の必要があるため、羽田空港なのか成田空港なのか、新幹線の下車駅はどこかなど、ロボットが質問相手に対して逆に質問するものも含まれている。

2・3 画面表示

実験では回答内容の伝達性を高めるためにロボットの背後に、画像や動画を表示するためのディスプレイを併設した。順路に関する質問応答は、当初東京駅構内や駅周辺の詳細な地図を用意し、目的地までの順路を表示することを試みたが、かえって分かりづらくなるおそれがあるため、シンプルなデザインを採用した(図2)。その他、ロボットが回答可能な範囲を示すイントロ画面(図3)も表示することとした。

2・4 言語切り替えとフローの改変

本実験の言語切り替えは、特定のフレーズを利用者が発することで、そのフレーズに対応する言語に切り替える設計にした。フレーズの決定に際しては少しでもスムーズな対話になるよう配慮し、ロボットであっても人間の係員相手に話しかけるようなフレーズとした。例えば、英語を希望する場合は利用者がロボットに“Can I ask you something?”(質問してもいいですか)と伝えると英語に切り替わる(中国語の場合は“我可以问你几个问题吗”、日本語の場合は“質問があります”)。この言語切り替えの流れについては、利用者が分かりやすいよう併設の画面上で案内をしたが、実験開始から2週間の実態として、多くの利用者が画面を見ることなくロボットに視線を向けてしまい、言語切り替えフェーズを理解しないまますぐに質問を始めてしまうケースが多くみられた。また、外国人利用者の8割程度が英語での対話を選択したことから英語を対話言語のデフォルトに設定し、中国語と日本語に変更を希望する場合にのみ言語切り替えを実行するよう対話フロー切り替えた。その他、言語切り替えの流れと同様に利用者が分かりやすいようイントロ画面上にロボットが対応している質問の具体例を記載したが、具体例をそのまま質問する利用者が多くみられた。我々としては利用者がロボットにどのような質問をするかについても関心があったため、回答可能な質問の種類のみを表示することとし、質問の具体例を削除した。

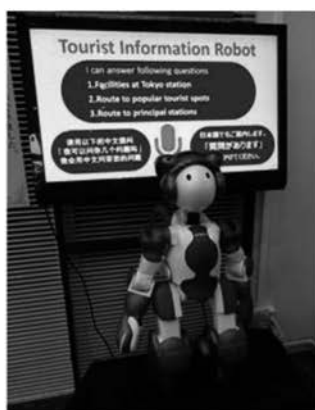


図1 EMIEW3



図2 順路案内

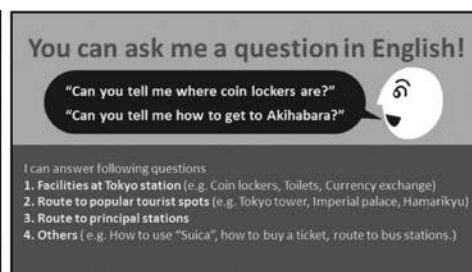


図3 イントロ画面

3. 実験結果

対話ログを集計した結果、実験参加者はのべ634名で、そのうち英語を選択した参加者が43%、中国語を選択した参加者が8%、日本語を選択した参加者が49%であった。

3・1 音声に基づく言語切り替え性能の確認

言語切り替えと質問応答の成功率を表1に示す。表中の試行数は、参加者がロボットに話しかけた回数を示し、成功数は言語切り替えまたは質問応答が成立した回数を示している成否は人が対話ログから判定)。日本人が発してしまいがちな「えー」や「あー」などの意味のない音声を除外するよう後半でプログラムを改変した結果、日本語への切り替えは23ポイント改善した。

表1 対話成功率

		言語切替			質問応答		
		試行数	成功数	成功率	試行数	成功数	成功率
前半	英語	282	176	62%	636	174	27%
	中国語	50	40	80%	68	22	32%
	日本語	407	205	50%	689	223	32%
	小計	739	421	57%	1393	419	30%
後半	英語	-	-	-	290	88	30%
	中国語	12	11	92%	27	6	22%
	日本語	153	111	73%	336	164	49%
	小計	165	122	74%	653	258	40%

3・2 雑音除去性能の確認

表1で成功数にカウントされなかったについて、その原因を特定して整理したものが表2である。このうち雑音除去が上手くいかなかったことが原因で失敗と判定した対話をVに区分した。発話部分の正確な検出(音声データから利用者の声だけを正確に切り出す処理)は、雑音除去精度に依存するところが大きい。V列は他の区分と比較して値が非常に小さく、雑音除去性能が高いことが見てとれる。

表2 対話失敗原因整理表

		I	II	III	IV	V	VI	VII	
前半	英語	9.9%	41.2%	6.2%	14.1%	0.2%	16.3%	12.1%	I. 質問時言語識別不良 II. 質問内容認識不良 III. 回答不正確 IV. 質問時の顔認識不良 V. 発話区間検出不良 VI. 回答情報なし VII. その他
	中国語	16.0%	42.0%	10.0%	6.0%	2.0%	2.0%	22.0%	
	日本語	10.8%	30.7%	12.3%	14.1%	0.6%	16.8%	14.7%	
	小計	10.6%	36.2%	9.4%	13.7%	0.5%	15.9%	13.8%	
後半	英語	1.0%	45.5%	4.0%	6.9%	4.0%	36.1%	25.0%	I. 質問時言語識別不良 II. 質問内容認識不良 III. 回答不正確 IV. 質問時の顔認識不良 V. 発話区間検出不良 VI. 回答情報なし VII. その他
	中国語	12.5%	45.8%	16.7%	8.3%	0.0%	4.2%	12.5%	
	日本語	4.6%	20.7%	17.8%	5.7%	2.9%	43.1%	5.2%	
	小計	5.6%	23.7%	17.7%	6.1%	2.5%	38.4%	6.1%	
合計		8.6%	35.8%	9.7%	11.7%	1.2%	21.7%	11.2%	

3・3 スムーズな対話実現への課題確認

参加者の発話が終了してからロボットの発話が始まるまでの時間の平均値を表3に示す。これは雑音除去とその他の言語処理時間を合わせたもので、言語の切り替えと質問応答では処理に違いがあるため個別に整理した。質問と回答の間が3秒程度開いたことで、ロボットに自分の質問が認識されていないと思って質問を繰り返した参加者が多く、体験後のアンケートでは日本人外国人ともに「応答スピードが遅い」が不満理由の2番目に挙げられた。

表3 処理時間 (秒)

		言語切替	質問応答
前半	英語	3.0	2.6
	中国語	3.5	3.0
	日本語	3.2	2.5
	小計	3.1	2.6
後半	英語	-	2.6
	中国語	3.1	1.9
	日本語	3.1	2.2
	小計	3.1	2.4

4. 考察

4・1 音声に基づく言語切り替え

理想としては、利用者が問いかけるだけで言語切り替えと質問内容の識別が同時に処理されることであるが、実際のところ両方同時に処理することはできない。実験では利用者の視線がディスプレイに向かないため、何をすべきか戸惑う参加者が多く、手順に沿って進めた場合でも切り替え成功率は前半で57%、後半で74%にとどまった。そのため分かりやすさの点からしても、音声認識技術の成熟度合いからしても、ロボットとの対話を音声だけで進めていくのは困難であるため、タッチ画面入力と併用することが現時点で妥当なアプローチであると考ええる。

4・2 雑音除去性能

表3で示したとおり雑音除去に起因する対話失敗率は非常に小さいことが分かった。しかし、人の声とそれ以外の音の分離は実現できても複数の人の声を聞き分けるまでには至っていないため、現状ではロボットの正面方向に立つ1人としか対話することができない。

4・3 スムーズな対話への課題

処理時間は対話環境の騒音レベルに影響されるため静粛性の高い空間では雑音除去に要する時間を短縮できるとされている。しかし事前の調査で訪旅セ内の騒音値は最大で65dB、平均55dBであり、改札やコンコースの騒音は更に大きいことを考えると、さらに処理時間が増大する可能性がある。そうした点からも言語切り替え等の選択式の対話の場合はタッチ画面による入力方式を採用するなど、音声認識に拘らない姿勢が求められる。また、音声認識処理実行中の旨を音声や描画で利用者に示すことも不満抑制に重要と考えられる。

5. おわりに

人間の案内係は、たとえ知らない質問がきたとしても背景知識や常識によってうまく対処することができるが、ロボットにはそうしたものがいないため、未登録の話題や言い回しが質問に含まれていると対話が成立しない。今回は、頻出とされる216の質問とその回答をあらかじめ学習させて実験に臨んだが、ロボットに学習させていない質問対話が成立しなくなるケースや、216の質問に含まれている話題ではあるが、言い回しの違いで質問意図を正確に認識することができずに対話が成立しないケースが数多く発生した(表3のⅡとⅣ)。我々ロボットを設置する側は、利用者が質問する話題や言い回しをあらかじめ全て把握できないため、多様な話題や言い回しに対応させるには実際の対話を通じてロボットに追加学習させていくべきと考えている。ただし、今回のようにロボット近傍に補助要員を配すると、ロボットではなく補助要員にばかり質問者が集まってしまうため、何らかの工夫が必要である。現在この課題に対応するための要素技術の開発がすすめられており、我々としては今後も引き続きロボットや人工知能(AI)による接客案内の実現するために必要な研究開発にメーカーやベンダーと共同で取り組んでいく予定である。