

タスク指向対話における 発話意図の対話リズムへの影響

藤原 敬記† 伊藤 敏彦† 荒木 健治†

†北海道大学大学院 情報科学研究科
〒060-0814 札幌市北区北14条西9丁目
E-mail: {fujiwara, t-itoh, araki}@media.eng.hokudai.ac.jp

あらまし 従来、自然で滑らかな人対機械のコミュニケーション実現において、音声認識や言語理解の精度、合成音声の質などが重要であると考えられていたが、我々の先行研究ではたとえ認識精度などが高くとも、対話リズムやシステム発話の韻律情報が十分でないとは機械と自然な対話を行えないことを示した。しかしながら、タスク指向対話における対話リズムの研究についてはほとんど行われていない。そこで本稿では、自然で滑らかな対話を行えるタスク指向型の音声対話システム実現のために、タスク指向対話を収集し、「対話構造、発話意図（発話内容）、相槌、復唱、間投詞」と「対話リズム（発話タイミング、F0、発話速度）」の関係について分析した。その結果、対話リズムは対話構造と発話意図による影響が強いこと、相槌・復唱には、対話リズムを維持するための厳しい制約が存在することが明らかとなった。

キーワード 発話意図、対話リズム、発話タイミング、F0、発話速度

Influence of Dialogue Acts on Dialogue Rhythm in Task-oriented Dialogues

Noriki FUJIWARA† Toshihiko ITOH† Kenji ARAKI†

†Graduate School of Information Science and Technology, Hokkaido University,
N14, W9, Kita-ku, Sapporo, 060-0814, JAPAN
E-mail: {fujiwara, t-itoh, araki}@media.eng.hokudai.ac.jp

Abstract We consider that factors such as prosody of systems' utterances and dialogue rhythm are important to attain a natural human-machine dialogue. However, it has been not revealed the relations between dialogue rhythm and speaker's various states in task-oriented dialogue. In this study, we collected task-oriented dialogues and analyzed the relations between "dialogue structures, kinds of dialogu acts (contents of utterances), *Aizuchi* (*backchannel/acknowledgment*), *Repeat* and *interjection*" and "dialogue rhythm (response timing, F0, and speech rate)". From the results, we understood that dialogue rhythm is affected by dialogue structures and dialogue acts significantly, moreover, utterances of *Aizuchi* and *Repeat* conform to restrictions to keep dialogue rhythm.

keyword Dialogue Act, Dialogue Rhythm, Response Timing, F0, Speech Rate

1. はじめに

近年、音声対話システムの性能は飛躍的に向上してきており、実用化を目指し、多くの研究がなされている。しかしながら、現在のところ人対機械のコミュニケーションが人間同士のものと同等であるとは言えない。従来、人対機械におけるコミュニケーションにおいては、音声認識や言語理解の精度、合成音声の質などが重要であると考えられ、これらに関して多くの研究が行われている [1]。しかし、我々の先行研究 [2] では、たとえ認識精度などが高くとも、対話リズムやシステム発話の韻律情報が十分でないとは機械と自然な対話を行えないことを示した。

北岡らは、この対話リズムに着目し、非タスク指向型の音声対話システムを開発している [3]。このシステムは、音響的特徴とキーワードを素性とし、

人間同士の対話を機械学習することにより適切な発話タイミングで相槌や応答を行うものである。我々も人対機械の自然な音声対話実現のために対話リズムに着目し、対話リズムを考慮したタスク指向型の音声対話システムの開発を行っている [4]。我々のシステムは、基本的な発話タイミングの決定方法は北岡らと同じであるが、ユーザ発話の音響的特徴にシステム発話の音響的特徴を同調させる機能を有する点が異なっている。これは、人間同士の協調的な対話においては、発話者の音響的特徴（発話タイミング・F0・発話速度）が対話が進むにつれ同調する傾向がある [5] というに基づいている。

我々のシステムは、対話リズムや同調傾向を考慮しない場合に比べて、対話の自然性に関しては若干の改善が見られたものの人間同士の対話の滑らかさや自然さには達していない。この原因として、シス

テムの発話タイミングが対話相手（ユーザ）の音響的特徴やキーワードのみに依存し、システム側の状態を一切考慮していないことや、同調機能に関しても常に相手の発話の音響的特徴に無条件にシステム発話の音響的特徴を同調させているためであると考えている。発話タイミング、F0、発話速度などの対話リズムや音響的な同調の度合いなどは、発話者の状態（対話の盛り上がり、発話意図、感情など）により大きく変化すると考えられる。しかしながら、我々のシステムも北岡らのシステムもそれらを考慮しているとは言えない。また、我々は発話者の状態の中でも、特に発話意図が対話リズムに与える影響が強いと考えているが、発話意図と対話リズムの関係を分析した研究はほとんど存在せず、またタスク指向対話における同調傾向についても研究されていない。スコギンズらは、「ボケ」と「ツッコミ」の発話意図と発話タイミングに関して分析を行っている[6]が、対象とした対話が漫才であり、発話意図の種類が一般的な音声対話システムが用いられる状況とは大きく異なる。

そこで我々は、人間同士のコミュニケーションの自然さで音声対話が可能な音声対話システムを実現するために、実際人間同士のタスク指向型の音声対話を収集し、発話意図と対話リズム（発話タイミング・F0・発話速度）に関する分析を行った結果について述べる。

2. 対話コーパス

タスク指向型の音声対話における対話リズムを分析するために、人間同士の音声対話を収録し、発話意図のタグ付けと音響ラベリングを行った。以下にその手順について述べる。

2.1 対話データ

対話タスクは、ホテルの検索・予約タスクとした。対話は Customer 役（ホテルを予約する側）と Agent 役（ホテル検索・予約を行う側）の 2 人によって行われる。Agent 役は、Customer 役が提示する条件を満たすホテルをデータベースから検索し、該当するホテルがある場合にはそのホテルに関する情報を伝える。もし Customer 役が提示する条件を全て満たすホテルがない場合には、一部の条件を満たすホテルの情報を伝えたり、代替案を提示したりと言ったネゴシエーションを行うこともある。Customer 役は、予め用意した状況の条件をできるだけ満たすように予約を行う。この状況は全部で 7 種類用意し、以下にその 1 例を示す。

- 9 月 4 日に海外旅行に出発するため、前日の夜は東京で 1 泊する必要がある。
- 飛行機の出発時刻は午前 9 時 45 分である。

- 海外で豪遊するため、出国前にはできるだけ出費を抑えたい。
- 飛行機の出発空港は羽田空港である。

対話の収集は、非対面の環境で行っており、ジェスチャーや視線、表情などの音声以外の情報伝達は行えないようにした。Agent 役・Customer 役を演じた被験者は大学生・大学院生の 17 人、収録した対話の総数は 50 対話である。なお、被験者は常に Agent 役または Customer 役を行うのではなく、ある対話では Agent 役を、別の対話では Customer 役を行うこともある。そのため、Agent 役は本対話タスク遂行における専門的な知識（効率的な交渉、など）はまったく持っていない。表 1 に作成した対話コーパスについて示す。

表 1: 対話コーパス

対話数	50
被験者数	17
全発話数	3844
Agent	3510
Customer	2215
全発話意図数	2520
Agent	2520
Customer	2520
平均対話時間	4 分 57 秒

2.2 音響ラベリング

Agent 役と Customer 役の各発話開始位置と発話終了位置に発話位置ラベルを付与した。なお、300ms 以上のポーズ（無音区間）があった場合には別の発話とみなした。このため、1 つの発話意図（2.3 節にて詳細に述べる）が複数の発話から構成されることもある。ラベリングには音声分析ツール“Wavesurfer”を使用し、波形と音声の両方を利用して行った。

2.3 意図タグ付け

通常、対話は 1 つ以上の exchange（やりとり）から構成され、その exchange は基本的に“Initiate-Response-(Follow-up)”の構造から成る。Initiate（働き掛け）は、新しい exchange を開始するための働き掛けの機能を持ち、Response（応答）は Initiate に対する反応の機能を持つ。Follow-up（了解）は、現在の exchange を終了させる機能があり、しばしば省略される。それぞれの構成要素はさらに 1 つ以上の発話意図に分類することができる。タグ付けに使用した発話意図の種類や選択基準などに関しては、談話タグワーキンググループが提案した談話単位タグ標準化案 [7] に従った。なお、本稿では収集した対話において出現頻度の高い発話意図だけを分析対象とした。以下に分析対象とした発話意図の種類とその意味を示す。これらの発話意図が対話全体に占める割合は約 93% であった。また、括弧内の数値は、Agent 役と Customer 役ごとにその発話意図を発話した数を表示している。

Initiate (働き掛け) :

- 確認 (Agent: 547, Customer: 55):
話し手が文脈または何らかの知識から聞き手の応答に対して予測を持って発する質問。
- 未知情報要求 (Agent: 248, Customer: 184):
話し手が聞き手の応答に対する予測を持っていない質問で、何らかの値または表現を応答として要求するもの。
- 依頼 (Agent: 138, Customer: 131):
聞き手に対する行為の要求で、聞き手が諾否または何らかの応答を返す必要のあるもの。
- 情報伝達 (Agent: 50, Customer: 7):
話し手の知識や意見、または話し手が事実と思っていることを述べるもの。
- 真偽情報要求 (Agent: 47, Customer: 35):
話し手が聞き手の応答に対する予測を持っていない質問で、「はい」または「いいえ」で答えられるもの。

Response (応答) :

- 未知情報応答 (Agent: 270, Customer: 242):
未知情報要求に対して、その値を返す発話。
- 肯定・受諾 (Agent: 127, Customer: 563):
真偽情報要求に対してその命題内容を肯定する際の返答、および依頼や勧誘に対してその要求を受け入れることを示す際の返答。
- 否定・拒否 (Agent: 5, Customer: 28):
真偽情報要求に対してその命題内容を否定する際の返答、および依頼や勧誘に対してその要求を受け入れないことを示す際の返答。

Follow-up (了解) :

- 了解 (Agent: 165, Customer: 168):
Response の後に続き、やりとりの目的が達成されたことを伝えるもの。

相槌・復唱:

- 相槌 (Agent: 295, Customer: 870):
話を聞いていることを表す際や相手の次発話を促す際に発話されるもの。(明確な応答ではなくスムーズな対話を実現するためのもの)
- 復唱 (Agent: 187, Customer: 38):
先行する相手の発話に含まれる重要な単語(キーワード)を繰り返すもの。

相槌と復唱は基本的には上記の “Initiate-Response-(Follow-up)” の構造においては Follow-up に含まれるが、我々は対話リズムに与える影響は Follow-up とは異なると考え、今回は Follow-up には含めずに調査することとした。これは、円滑な対話を実現するためにロボットの相槌・復唱の機能に注目した藤江らの考え [8] と一致する。図 1 に対話例と付与し

た意図タグの例を示す。

3. 分析

本稿では、対話リズムを構成する重要な要素として発話タイミング、F0、発話速度を考え、これらについて以下のような方法で求めた。発話タイミングは、対象発話意図の開始時刻から先行発話意図の発話終了時刻の差とした。これは話者交替が起きた場合には相手発話との間隔になり、話者交替が起きなかった場合には、自身の先行の発話意図との間隔になる。また、オーバーラップが発生した際には、この発話タイミングは負の値となる。なお、5秒以上の間隔がある発話タイミングに関しては、ホテルのデータベース検索を行うなど、対話リズムとは別の理由であるとみなし、本稿では分析対象から除いた。F0 に関しては、平均 $\log(F0)$ を求めた。これは有声区間の $\log(F0)$ の和を有声区間の分析フレーム数で割ったものである。F0 の抽出の際には、“ESPS/waves+” [9] を使用した。また、F0 が計算できなかった発話意図は、分析対象から除いた。発話速度に関しては、1つの発話意図における発話内容のモーラ数を発話持続時間で割り求めた。また、発話速度の個人差を取り除くために、各発話の発話速度から発話者の平均発話速度(「相槌」・「復唱」も含む全発話の平均)を減ずることで相対的な発話速度も同時に求めた。

3.1 Initiate・Response・Follow-up 間の比較

まず Initiate, Response, Follow-up のグループごとの対話リズムについて分析を行った。表 2 に、「対話全体」・「話者交替時」・「話者継続時」の「発話タイミング」・「平均 $\log(F0)$ 」・「相対発話速度」のそれぞれの平均値と標準偏差を示す。なお参考のために「相槌」・「復唱」についても示す。括弧内の数字は、各発話意図の数を表している。これより、「対話全体」・「話者交替時」・「話者継続時」の全ての場合において、Initiate は Response よりも発話タイミングが遅いことが分かる。対話における Initiate の基本的な役割とは、新たな exchange や話題を開始することであり、タスク達成のための対話の流れを決定し、制御することと考えられる。それ故、Initiate は、Response に比べ発話内容を考える時間が長いと考えられ、発話者自身が思考にかけても良いと思っている時間も、対話相手が許容できる時間も長いと考えられる。逆に、Response は Initiate に対する反応であり、情報検索などの場合を除けば、Response の思考時間は Initiate よりも短い。また、Response の発話は exchange 構造における時間的制約を受け、許容できる発話開始までの時間は短くなると考えられる。一方、発話速度に関しては、「話者継続時」を

C: 東京でホテルを取りたいんですけど。	[依頼]
A: 東京のホテルでよろしいですか？	[確認]
C: はい。	[肯定・受諾]
A: お名前のほうよろしいですか？	[未知情報要求]
C: 鈴木です。	[未知情報応答]
A: お電話のほうもよろしいですか？	[未知情報要求]
C: 012, (はい, [相槌]) 345, (345, [復唱]) 6789, (6789, [復唱])	[未知情報応答]
A: はい, ありがとうございます。	[了解]

(C : Customer, A : Agent)

図 1: タグ付けした対話例

除き, Response は Initiate よりも発話速度は若干遅いことが分かる。これは Response は Initiate よりも重要な情報を含むことが多く, 発話者は対話相手に対してより確実に情報を伝えようとするために, 発話速度が遅くなったと考えられる。話者継続時の発話速度は Initiate よりも Response の方が速くなっているが, Response に属する重要な情報を含む発話意図である「未知情報応答」だけを対象とすると, Initiate の発話速度の方が速くなっている。

最後に Follow-up に関してだが, 「対話全体」・「話者交替時」・「話者継続時」の全ての場合において, Follow-up の発話タイミングは Response とほぼ同じであるが, これは, Response の場合と同じ理由であると考えられる。しかし一方で Follow-up は Initiate・Response に比べて F0 が低く, 発話速度は明らかに早い。この理由は Follow-up は省略されることも多い付加的な発話であり, 新たなやりとりを始めるための発話である Initiate や対話相手からの働き掛けに対する反応である Response に比べて発話内容自体の重要度が低いためと考えられる。

3.2 各発話意図間の比較

各発話意図の対話リズムに関する分析結果を, 図 2, 3, 4 に示す。まず対話リズムに大きな影響を与えると考えられる話者交替時について述べる。話者交替時における発話タイミングは (図 2 の ■ を比較すると), 個々の発話意図では, 「Initiate に属する発話意図」 > 「Response に属する発話意図 (「肯定・受諾」を除く)」 > 「Follow-up に属する発話意図」のようにきれいな傾向があり, 思考にかかる時間と発話タイミングの強い関係を示している。このとき「肯定・受諾」の発話タイミングだけが特に早い理由としては以下のように考えられる。「肯定・受諾」のほとんどは, 「確認」に対する応答であり, 確認する箇所 (内容) は本タスクではある程度決まっているため, 「肯定・受諾」の発話者にとっては, 対話相手が確認することをあらかじめ予想でき, 実際

に確認された際にはすぐに応答ができると考えられる。そのため, 「肯定・受諾」の発話には, 思考時間がほとんど必要なく, オーバーラップも他の発話意図に比べ多い。また各発話意図のオーバーラップに注目すると, 新情報や重要な情報を扱う「未知情報要求」や「依頼」, 「情報伝達」では, オーバーラップはあまり起きていない。反対に, 頻繁にオーバーラップが起きている発話意図には, 「確認」や「真偽情報要求」などがある。これら発話意図は, タスクにおける問題解決のための典型的な発話が多く, 思考時間が少なくとも十分であるためと考えられる。

F0 に関しては, 各発話意図のほとんどの組み合わせにおいて有意な差が見られた ($p < 0.01$)。 「確認」や「肯定・受諾」, 「了解」と言った予想しやすい (されやすい) 発話意図や発話自体に重要な情報が含まれていない発話意図に関しては, F0 は低くなり, 重要な情報を含んでいる場合や予想しにくい発話意図の場合は, F0 は高くなる傾向 (強調傾向) があつた。

発話速度に関しては, 新情報や予想しにくい情報, 重要な情報を含む発話意図の発話速度はゆっくりになり, 対話リズムを整える働きのある発話意図の場合には発話速度が非常に速くなる傾向があつた。

なお, 「否定・拒否」の思考時間は, 「肯定・受諾」とほぼ同じか若干長い程度であると考えているが, 実際の「否定・拒否」の発話タイミングは, 「肯定・受諾」よりもかなり遅く, また F0 はかなり高く, 発話速度はかなりゆっくりである。この理由として, 「否定・拒否」の発話は相手の予想 (期待) とは異なるという意味では重要な発話であるため, 発話者は意図的に強調していることが考えられる。

次に Response に属する発話意図に関して, 「重要度が高い発話」 = 「発話タイミングが遅い発話」と仮定し, 発話タイミング・F0・発話速度の各間での相関係数を求めると, 発話タイミングと F0 では $r=0.84$, 発話タイミングと発話速度では $r=-1.00$, 発話速度と F0 では $r=-0.84$ とそれぞれの相関は非

表 2: 対話全体・話者交替時・話者継続時における対話構造ごとの平均値と標準偏差

		対話全体				話者交替				話者継続			
		I (1329)	R (1212)	F (329)	A・R (1391)	I (870)	R (1015)	F (234)	A・R (1383)	I (459)	R (197)	F (95)	A・R (8)
発話 タイミング	平均値	1.06	0.44	0.50	0.20	0.94	0.34	0.40	0.20	1.28	0.95	0.74	0.82
	標準偏差	1.00	0.74	0.81	0.30	0.95	0.70	0.80	0.29	1.04	0.69	0.77	0.38
平均 log(F0)	平均値	4.71	4.73	4.61	4.63	4.72	4.74	4.62	4.64	4.71	4.67	4.59	4.56
	標準偏差	0.24	0.29	0.20	0.22	0.26	0.30	0.22	0.22	0.19	0.26	0.13	0.11
相対 発話速度	平均値	-0.21	-0.39	0.33	0.68	-0.25	-0.45	0.13	0.69	-0.14	-0.06	0.82	-1.39
	標準偏差	2.08	2.78	3.51	3.21	2.11	2.76	3.70	3.21	2.01	2.87	2.92	1.22

(I:Initiate, R:Response, F:Follow-up, A・R:相槌・復唱)

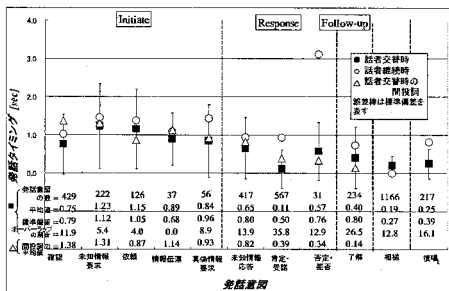


図 2: 各発話意図の発話タイミング

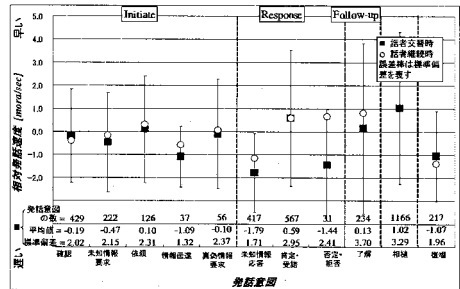


図 4: 各発話意図の相対発話速度

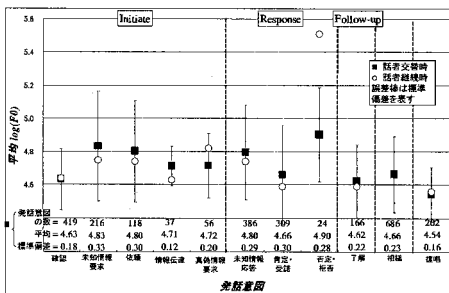


図 3: 各発話意図の平均 log(F0)

常に高いことが分かった。すなわち応答においては、重要度の高い発話は、発話タイミングが遅く、F0は高く、発話速度はゆっくりになり、しばしば強調される傾向がある。Initiateに属する発話意図に関しては、発話タイミングとF0には強い相関があったが($r=0.97$)、その他の組み合わせでは相関が見られなかった。

最後に、話者継続時の発話タイミング・F0・発話速度(図2, 3, 4の○)に関してだが、各発話意図間の関係は、話者交替時のものと同じ傾向であった。しかし、同一発話意図における話者継続時と話者交替時の値を比較すると、話者継続時の発話タイミングは話者交替時よりも遅く、F0に関しては低く、発話速度に関しては速くなっている。

3.3 相槌と復唱に関する分析

「相槌」と「復唱」の発話タイミングに関しては、共に他の発話意図に比べ発話タイミングが早く、さらに標準偏差も非常に小さい。この理由として、「相槌」と「復唱」は対話リズムを整える働きが強く、発話タイミングに関しては他の発話意図よりも厳しい制約があると思われる。そのため、人間は音響的な特徴などを使いほとんど反射的に応答を行っていると考えられる。なお「相槌」の発話タイミングに関しては、市川らの、相槌は0.4sec以内に打たれる[10]という結果と一致する。

また、「復唱」の発話速度は、かなりゆっくりであるが、F0は他の発話意図のF0よりもかなり低い。この理由として、「復唱」はキーワードを繰り返すことによる暗黙的な確認であるが、相手の発話を邪魔しないように、発話者が意図的にF0を低く抑えていると考えられる。このように、対話リズムを調整する働きのある発話に関しては、相手に伝達すべき情報を含んでいる場合(「復唱」など)は、情報を伝達できかつ相手の発話を邪魔しないように、発話速度はゆっくりになるが、F0をかなり低くする。反対に、相手に伝達すべき重要な情報を含んでいない場合(「相槌」など)は、相手の発話を邪魔しないように、発話速度を速くし、さらにF0を若干低くする。このようにして「相槌」や「復唱」の発話者は、相手の発話の邪魔をせず、対話のテンポを良くしながら対話を進めていると考えられる。

3.4 間投詞の発話タイミングの分析

発話交替時の間投詞の発話タイミングに関して分析したところ、全体の平均値と標準偏差はそれぞれ0.94, 0.97であった。また発話意図ごとの発話タイミングと間投詞の発話タイミングとの関係について分析したところ(図2中の■と△)、興味深い結果が得られた。それは、発話意図ごとの発話タイミングと間投詞の発話タイミングには強い相関が見られたことである($r=0.73$)。ある発話意図の発話タイミングを x 、その発話意図の間投詞の発話タイミングを y とした場合の回帰直線は、 $y = 0.92x + 0.14$ であった。

思考にかかる時間がほぼ同じである「肯定・受諾」と「否定・拒否」は間投詞の発話タイミングがほぼ同じであることや、各発話意図における間投詞の発話タイミングはその意図の発話タイミングの直後であることが多いことから、発話意図ごとに思考にかけられる時間が存在し、その時間内に発話内容を決定できない場合や決定できないと予想した場合に間投詞を発話することが示唆される。

4. 自然でスムーズな人・機械間の対話実現を目指して

前節で述べた結果より、我々は2話者によるタスク指向対話において対話リズムを以下のように考えている。各発話意図には思考にかけられる時間(許容できるポーズ)が存在し、その時間は、発話意図の種類(exchange構造内の種類を含む)や対話構造の状態(話者交替/話者継続)により決定される。また発話者はその思考にかけられる時間内に発話内容を決定できない場合や決定できないと予想した場合に間投詞を発話する。実際の各発話タイミングは、「思考にかけられる時間(許容できるポーズ)」と「発話意図、発話内容の重要度、発話内容の決定度(発話内容がどの程度決まっているか)」の関係により決定されると考えている。そのため、発話タイミングは、「対話相手の先行発話の予想しやすさ」、「対話相手の発話内容の理解しやすさ」、「発話意図」、「発話内容の重要度」、「対話相手が予想(期待)している発話意図と決定した発話意図との差」、「情報検索の時間」などが強く影響すると考えている。

また発話全体のF0と発話速度の決定には、「発話意図」、「発話内容の重要度」、「対話相手が予想(期待)している発話意図と決定した発話意図との差」などが強く影響すると考えている。

さらに聞き手も各発話意図(発話内容)の平均的な発話タイミング、F0、発話速度、さらに可能ならば対話相手の発話タイミング、F0、発話速度をモデル化しており、そのモデル化された対話リズムと実

際の発話の対話リズムの差から、聞き手は様々な非言語情報を取得していると考えられる。そのため、スムーズで自然なタスク指向型の音声対話システム実現には、平均的な発話意図と発話タイミング、F0、発話速度との関係について人間同士の対話を用いてモデル化し、音声対話システムがそのモデルを用いてユーザの状態とシステムの状態を考慮した対話リズムを実現することと、さらに実際のユーザ発話とモデルの違いから非言語情報の取得ができる必要があると考えている。

5. まとめ

本稿では、タスク指向対話を収集し、「対話構造、発話意図、相槌、復唱、間投詞」と「対話リズム(発話タイミング、F0、発話速度)」の関係について分析を行った。その結果、スムーズで自然な対話の実現に関して重要な知見が得られた。

今後の予定としては、タスク指向対話における対話中の対話リズムの推移と同調傾向の分析、先行の発話意図を考慮した当該発話意図と対話リズムに関する分析を行う予定である。またこれらの分析を行った後は、そこで得られた知見を用いて、システムのさらなる改良を行う予定である。

参考文献

- [1] 藤原 敬記, 伊藤 敏彦, 荒木 健治, 甲斐 充彦, 小西 達裕, 伊東 幸宏, “認識信頼度と対話履歴を用いた音声言語理解手法”, 電子情報通信学会論文誌, Vol. J89-D, No. 7, pp.1493-1503.
- [2] 伊藤 敏彦, 山田 真也, 荒木 健治, “音声認識・言語理解性能や状況の違いによるタスク指向音声対話の言語的・音響的特徴の比較”, 日本音響学会誌, Vol. 63, No. 5, 2007(採録決定).
- [3] Kitaoka, N., Takeuchi, M., Nishimura, R., and Nakagawa, S., “Response Timing Detection Using Prosodic and Linguistic Information for Human-friendly Spoken Dialog Systems”, Journal of JSAI, Vol. 20, No. 3 SP-E, pp. 220-228(2005).
- [4] 東海林 圭輔, 高橋 美佳, 井原 誠也, 伊藤 敏彦, 荒木 健治, “対話に関するリズムや同調作用を考慮した音声対話システム”, 情報処理学会研究報告, 2006-SLP-61, pp.43-48(2006).
- [5] 長岡 千賀, 小森 政嗣, 中村 敏枝, “対話における交替潜時の対話者間影響”, 日本人間工学会誌, Vol. 38, No. 6, pp.316-323(2002).
- [6] スコギンズ・リーバイ, 川嶋宏彰, 松山隆司, “間の合った発話タイミング制御を目的とした漫才の動的構造の分析”, インタラクション 2005, D-404(2005).
- [7] 荒木 雅弘, 伊藤 敏彦, 熊谷 智子, 石崎 雅人, “発話単位タグ標準化案の作成”, 人工知能学会論文誌, Vol. 14, No. 2, pp.251-260(1999).
- [8] 藤江 真也, 福島 健太, 柴田 大輔, 小林 哲則, “FSTと韻律情報を用いた相槌・復唱機能を有する対話ロボット”, 人工知能学会研究会資料, SIG-SLUD-A401-03, pp.15-20(2004).
- [9] Software manuals of ESPS/waves+ with EnSigTM. 1997. Entropic Research Laboratory, Inc..
- [10] 市川 熹, 佐藤 伸二, “対話理解に対する抑揚情報の役割”, 情報処理学会研究報告, 1994-SLP-2, pp.51-58(1994).