

発話速度を考慮した発話タイミング分析*

◎芦村 和幸 (JST/CREST) ニック・キャンベル (ATR, JST/CREST)

1 はじめに

対話音声データを用いて、話者交代に着目した発話タイミングの分析を行ない、発話間 overlap の有無に応じて、異なる基準により発話タイミングが決定されることを示した [1]。しかしながら、発話タイミングは、「話し手発話」と「聞き手発話」との相対的な関係であるため、各話者の発話速度を考慮する必要がある。本稿では、話し手発話開始と聞き手発話開始との相対的な関係を調べるため、大規模対話音声データにもとづき話し手の発話速度を分析したうえで、聞き手の発話タイミングについて検討した。

2 音声試料

音声試料として、現在、JST/CREST 発話様式プロジェクトで作成中の、電話対話データベース [2] を用いた。このデータベースは、10 週間にわたって、10 人の被験者の対話音声を収録したものであり、毎回の収録においては、2 人一組の被験者による 30 分の自由対話を収録してある。収録音声ファイルに対しては、スペクトログラムの視察にもとづき、以下のような情報が付与されている。

- 発話内容の書き起こしテキスト
- 発話開始・終了の時間情報
- 発話内ポーズ長
- 発話の重複箇所
- 発話機能 (質問-応答ペア、あいづち、など)

今回の分析対象データとしては、「収録にある程度慣れたが、まだ友人というほどは親しくない話者同士」という観点で、第 2 週目の収録結果 30 分を採用した。対象被験者は、22 歳男性 (話し手) および 42 歳男性 (聞き手) の二人である。

3 発話タイミングの定義

本稿では、「発話のタイミング」という用語を、「話し手発話の開始時刻」と「聞き手発話の開始時刻」との相対的かつ時間的な位置関係としてとらえる。この関係は、発話間 overlap の有無に応じて、大きく 2 つに分けられる。

overlap がない場合 聞き手が、話し手の発話を最後まで聞き、意図を理解した上で、内容や状況に最

適なタイミングまで待ってから発話を開始する。この場合、話し手発話と、聞き手発話の間にポーズが知覚され、このポーズ長により聞き手の発話タイミングが定義される。

overlap がある場合 聞き手が、話し手の発話完了を待たずに発話を開始する。話し手発話の途中で聞き手発話が始まるため、単なる発話開始時間ではなく、「話し手発話のどこで聞き手発話が始まるか」という overlap 発生位置により聞き手の発話タイミングが定義される。

4 発話タイミング分析

まず、「発話間 overlap の有無」、および「話し手発話内のポーズの有無」により分類した上で、「話し手発話」と「聞き手発話」の発話間ポーズ時間長を求めた (表 1)。全データの標準偏差が、その平均に比べて非常に大きいものに対して、「発話間 overlap の有無」および「話し手発話内のポーズの有無」による分類後は標準偏差が比較的小さくなっており、overlap および発話内ポーズによる分類は妥当と考えられる。しかしながら、分類後も平均に対する標準偏差の比率が大きいので、さらなる分析が必要と考えられる。

表 1 発話間ポーズのばらつき

overlap	話し手発話内のポーズ	件数	平均 (秒)	標準偏差 (秒)
なし	なし	189	0.319	0.343
	あり	49	0.368	0.310
あり		247	-0.498	0.710
全データ		485	-0.092	0.695

※注意: overlap がある場合については、参考情報として、overlap 発生位置を、話し手発話完了位置を基準とした負のポーズ長として示す。

次に、発話タイミングの傾向を詳しく調べるために、overlap がない場合については、発話間ポーズ長を話し手の発話速度¹により正規化 [3] した上で、ポーズ長の分布を求めた。また、overlap がある場合については、overlap 発生位置と各発生位置での発生件数を求めた。

*Analysis of utterance timing based on speech rate, by Kazuyuki ASHIMURA (JST/CREST) and Nick CAMPBELL (ATR, JST/CREST)

¹先行研究 [3] では、発話時間長を発話モーラ数で除した「発話テンポ」を用いているが、本稿では、逆に、発話モーラ数を発話時間長で除したものとして「発話速度」を定義する。

4.1 overlap がない場合

4.1.1 発話内ポーズがない場合

発話間ポーズ長は 1~12 モーラに分布しており、0~1 モーラにピークがみられた (図 1)。したがって、聞き手は、話し手の発話完了を待って、できるだけ早く発話を開始しようとしていると考えられる。

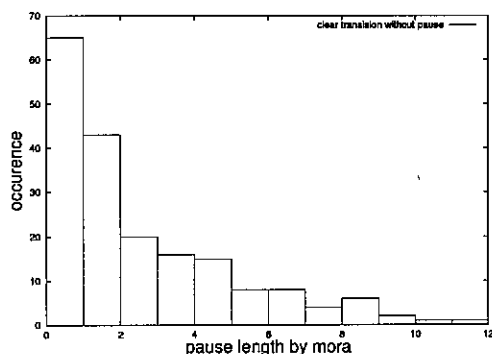


図 1 overlap がなく発話内ポーズがない場合のモーラポーズ長分布

4.1.2 発話内ポーズがある場合

発話間ポーズ長は 0~9 モーラに分布しており、1~2, 3~4, および 7~8 (いずれも、モーラ) という、3 つのピークがみられた (図 2)。各ピークを構成する発話の機能は、表 2 に示すとおりであり、1~2 においては「あいづち (単なるあいづち)」, 3~4 においては「発話継続」が特徴的であることがわかった。また、7~8 においては特徴的な機能はみられなかった。

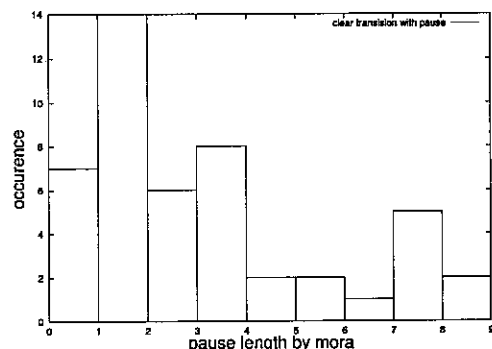


図 2 overlap がなく発話内ポーズがある場合のモーラポーズ長分布

4.2 overlap がある場合

overlap 発生位置と各発生位置での発生件数を分析した (表 3)。表 3 より、音声的観点からは、「発話末尾付近」・「発話先頭付近」・「母音、/N/, /s/ の引き延ばし」・「末尾、先頭以外の句境界付近」という 4 つの位置で overlap が発生するケースが多く、言語的観点からは、「呼応の各要素の間」[4] で overlap が発生するケースが多いことがわかった。

表 2 3 つのピークを構成する発話の機能内訳

ピークのモーラポーズ長	聞き手発話の機能	件数
1~2	あいづち (単なるあいづち)	6
	発話継続	4
	あいづち (受理の表明)	3
	あいづち (応答)	1
3~4	発話継続	4
	質問-応答ペア	2
	あいづち (受理の表明)	1
	あいづち (単なるあいづち)	1
7~8	新情報提供	1
	話し手の質問に対する応答	1
	質問	1
	発話継続	1
	あいづち (応答)	1

表 3 各 overlap 発生位置における overlap 発生頻度

overlap 発生位置	件数
音声的観点で見た発生位置:	
発話末尾付近	126 (30)
発話先頭付近	52 (12)
母音、/N/, /s/ の引き延ばし	80 (19)
末尾、先頭以外の句境界付近	63 (15)
促音付近	14 (3)
笑い付近	14 (3)
発話内ポーズ付近	11 (3)
言語的観点で見た発生位置:	
呼応の各要素の間	55 (13)
複文の接続詞/接続助詞の前後	4 (1)
題述構造の前項と後項の間	1 (0.2)
倒置された述語の後	3 (0.7)
総数	422 (100)

※注意: 件数の右隣の () 内の数値は、各件数の総数 422 件に対する割合 (%)。

5 おわりに

対話音声データにおける発話タイミングについて、話し手の発話速度に着目して分析した結果を報告した。今後は、今回分析した発話タイミングと、各話者の発話内容のもつ機能との関連性を詳細化するとともに、発話音声の音節構造と発話速度の関連性について検討を進めたい。

参考文献

- [1] 芦村, キャンベル: 日常会話における発話タイミング分析, 情報研報, 2002-SLP-40-19, pp.109-114, 2002
- [2] 芦村, キャンベル: JST/CREST 発話様式プロジェクトの電話対話データベース, 人工知能学会全国大会論文集, 3C5-11, 2002
- [3] 海木, 句坂: 局所的な句構造によるポーズ挿入規則化の検討, 信学会論文誌 D-II, Vol. J79-D-2 No.9, pp.1455-1463, 1996
- [4] 木田, 乾, 神崎, 高梨, 井佐原: 構文論から見た対話-円滑な話者交替を可能にする構文構造-, 人工知能研資, SIG-SLUD-A102-6(11/06), pp.33-38, 2001