

# やり取りにもとづく会話相互作用の組織化にみられる 韻律特徴の分析

小磯 花絵

国立国語研究所 言語行動研究部

伝 康晴

奈良先端科学技術大学院大学 情報科学研究科

## 1 はじめに

会話における話者間相互作用の構造と言語・パラ言語情報との関係を明らかにすることは、相互作用の組織化のメカニズムを知る上で重要である。本研究では発話のやり取りにもとづく相互作用の構造に着目し、韻律特徴との関係を明らかにする。やり取りとは働きかけや応答などの型を持つ発話の一定パターンの連鎖によって特徴付けられるものであり、相互作用上の単位を構成する。このようなやり取り単位の内部と境界という構造上の差異は、韻律特徴に反映されることが予測される。本稿では、実対話の分析を通し、やり取り単位の内部と境界とで F0・パワー・継続長などの韻律特徴に有意な差がみられることを示す。以上の結果を、プランニングコストと談話構造の観点から考察する。

## 2 やり取りにもとづく会話の構造

やり取りにもとづく会話の構造は、発話相互作用の最小単位である交換を基礎にする (Coulthard, 1977/1985)。交換は、(1) にみられるように、質問・依頼・陳述などの働き掛け型、返答・承諾・承認などの応答型、容認・評価などの補足型のムーブの連鎖によって構成される。

- (1) A: えーと大きな松の木ありますか [働きかけ]  
B: あはいあります [応答]  
A: はい [補足]

ここでムーブとは、相互作用上の機能にもとづき定義される行為によって構成される単位である。

また、(2) の 2 番目の発話にみられるように、相手の働きかけに対する応答であると同時に、それ自体も相手に対する働きかけとなるようなムーブも存在する。

- (2) A: 銀の何でしたっけ [働きかけ]  
B: 鉱山ですか [応答兼働きかけ]  
A: 銀の鉱山、はい [応答]  
B: はい [補足]

交換を構成するこれらのムーブは、(1) ある特定の要素が後続するという予測をもたらすかという予測性の観点と、(2) 先行要素によって予測されたものであるかという被予測性の観点から相互に区別される (表 1)。

表 1: 予測性・被予測性の観点からのムーブの分類

	働きかけ	応答兼働きかけ	応答	補足
予測的	+	+	-	-
被予測的	-	+	+	-

このように、働きかけムーブや応答兼働きかけムーブは、ある特定の要素が後続することを拘束する力があり、この拘束力は一定のパターンのムーブ連鎖からなる交換の成立に寄与する。一方、応答や補足は予測的でなく、特定の要素の後続を拘束しない。

このようなやり取りにもとづく会話の構造が以下での分析の基礎となる。

## 3 方法

### 3.1 分析資料

千葉大学地図課題対話コーパス (堀内ほか, 1999) より 8 組の男性被験者 (計 16 人) の対話を 2 対話ずつ (計 16 対話) 選択し、各 5 分 (計 80 分) を分析の対象とした。各被験者ペアの 2 対話は、情報提供者・追従者の役割を入れ替えたものである。

会話音声は DAT で 2 チャンネル独立に収録し、標本化周波数 20kHz で計算機に取り込み分析に使用した。

### 3.2 手続き

分析は以下の手順で行った。

1. 分析資料に対してやり取り構造を付与し、各ムーブ対 (隣接する 2 つのムーブの対) に対して交換内部・境界の属性を与える。
2. 各ムーブの韻律特徴を抽出し、隣接するムーブ間での特徴の変化量をそのムーブ対の韻律特徴とする。

表 2: 交換内部・交換境界におけるムーブ対の韻律特徴の平均値

		N	最小 F0	最大 F0	F0 レンジ	最大パワー	継続長	休止時間	モーラ長	間隙長
全体	内部	659	5.2	-35.4	-40.5	-348.4	-2284.1	-668.3	-4.4	346.0
	境界	400	-10.1	46.8	56.9	463.7	3668.1	1185.4	10.2	559.2
提供者	内部	248	-0.1	-11.1	-11.1	-134.6	-710.8	-109.5	1.0	258.2
	境界	306	-10.6	54.1	64.7	517.3	4312.9	1448.7	15.4	567.7
追従者	内部	411	8.3	-50.0	-58.3	-477.4	-3233.5	-1005.5	-7.7	399.0
	境界	94	-8.4	23.1	31.5	289.0	1568.8	328.2	-6.8	531.4

表 3: ムーブ対の韻律特徴へのやり取り構造の影響

	最小 F0	最大 F0	F0 レンジ	最大パワー	継続長	休止時間	モーラ長	間隙長
全体	$t = 8.2^{**}$	$t = 17.2^{**}$	$t = 19.7^{**}$	$t = 18.5^{**}$	$t = 24.2^{**}$	$t = 17.3^{**}$	$t = 5.1^{**}$	$t = 5.4^{**}$
提供者	$t = 4.1^{**}$	$t = 10.7^{**}$	$t = 12.1^{**}$	$t = 11.4^{**}$	$t = 16.1^{**}$	$t = 10.9^{**}$	$t = 3.8^{**}$	$t = 7.1^{**}$
追従者	$t = 5.1^{**}$	$t = 8.4^{**}$	$t = 10.1^{**}$	$t = 9.7^{**}$	$t = 11.2^{**}$	$t = 7.4^{**}$	$t = 0.2$	$t = 1.6$

(全体:  $df = 1057$ , 提供者:  $df = 552$ , 追従者:  $df = 503$ , \*\*: 1%水準で有意)

3. おのおのの特徴について、交換内部・境界におけるムーブ対の韻律特徴の平均値を比較する。

### 3.3 やり取り構造の付与

ラベリング手順 やり取り構造の付与は、Coulthard (1977/1985) を参考にして設定した基準をもとに、2 人の著者が独立に行った。ラベリングは 100 ミリ秒以上の無音で区切られた間休止単位 (IPU) を基本単位とし、単一の言語行為を遂行していると見なせる範囲をムーブとしてまとめたうえ、各ムーブに対して働きかけ・応答・補足・応答兼働きかけのラベルを付与した。ただし、挿入連鎖 (Schegloff, 1972) も許しており、そこでの働きかけムーブは新たな交換を開始しないものとした。独り言など相互作用上の機能を持たない発話は「その他」とした。

この際、会話の構造を複雑化する一因である重複発話に対処するため、相手の IPU と同時 (200 ミリ秒以内の差) もしくはより後に始まった IPU が、相手の IPU と同時もしくはより先に終了する場合 (あいづちを含む) をラベリング対象から除外した。ただし、両者の IPU が同時に始まって同時に終了する場合は、先行・後続文脈との関連性が高い方を残し、他方を除外した。

一致率 ラベリングの一致を Carletta et al. (1997) と同様の方法で評価した。ムーブ境界の一致率は  $K = .95$  ( $N = 3929$ ,  $k = 2$ ) であり、ムーブ境界が一致したものに關するやり取り構造ラベルの一致率 (上記の基準により除外した IPU を除く) は  $K = .82$  ( $N = 1557$ ,  $k = 2$ )

であった。これは十分に高い一致であり、以下の分析の妥当性を保証する。ムーブ境界およびやり取り構造ラベルが一致しなかったものについては、2 人の著者の相談のうえ、最終判断を行った。

交換内部・境界の属性の付与 各ムーブ対について、第 2 ムーブが働きかけやその他でない場合を交換内部、第 1 ムーブが応答か補足で第 2 ムーブが働きかけの場合を交換境界とした。それ以外の場合は分析対象から除外した。

### 3.4 韻律特徴の抽出

ムーブの韻律特徴として、(1) 最小 F0, (2) 最大 F0, (3) F0 レンジ (= (2) - (1)), (4) 最大パワー, (5) 継続長, (6) ムーブ内休止時間の総和, (7) 平均モーラ長 (= ((5) - (6))/モーラ数) の 7 つの特徴に着目した。隣接するムーブ対の特徴として、これらの韻律特徴の変化量 (= 第 2 ムーブの値 - 第 1 ムーブの値) を用いた。それに加え、(8) ムーブ対間の移行時間 (間隙長) も特徴として用いた。(1)-(4) は音響分析ソフトウェア ESPS/Waves+ で抽出 (メディアフィルターを使用), (5)-(8) はコーパスに付与された時間情報と転記テキストから算出した。

## 4 結果

交換内部・境界におけるムーブ対の韻律特徴の平均値を比較したところ (表 2), 全ての特徴で有意差がみられた (表 3)。それぞれの特徴に関して以下の傾向が指摘できる。(1) 最小 F0: 内部で増加, 境界で減少, (2) 最大

表 4: 内部交替・境界交替・境界継続におけるムーブ対の韻律特徴の平均値

		N	最小 F0	最大 F0	F0 レンジ	最大パワー	継続長	休止時間	モーラ長	間隙長
全体	内部	659	5.2	-35.4	-40.5	-348.4	-2284.1	-668.3	-4.4	346.0
	交替	277	-6.6	47.0	53.7	453.2	3350.8	1079.1	10.1	522.2
	継続	123	-17.9	46.2	64.1	487.3	4382.6	1424.6	10.4	642.6
提供者	内部	248	-0.1	-11.1	-11.1	-134.6	-710.8	-109.5	1.0	258.2
	交替	218	-7.5	54.5	61.9	533.2	3880.9	1292.2	14.8	581.1
	継続	88	-18.5	53.1	71.6	477.9	5383.2	1836.3	17.0	534.5
追従者	内部	411	8.3	-50.0	-58.3	-477.4	-3233.5	-1005.5	-7.7	399.0
	交替	59	-3.6	19.6	23.3	157.5	1392.0	291.7	-7.3	304.2
	継続	35	-16.4	28.9	45.3	510.7	1866.9	389.7	-6.1	914.3

(内部: 交換内部での話者交替, 交替: 交換境界での話者交替, 継続: 交換境界での話者継続)

表 5: ムーブ対の韻律特徴へのやり取り構造と話者交替の影響

	最小 F0	最大 F0	F0 レンジ	最大パワー	継続長	休止時間	モーラ長	間隙長
全体	$F = 40.0^{**}$ $MS_e = 862$	$F = 147.7^{**}$ $MS_e = 5689$	$F = 194.6^{**}$ $MS_e = 6096$	$F = 171.7^{**}$ $MS_e = 478402$	$F = 296.8^{**}$ $MS_e = 15009345$	$F = 151.7^{**}$ $MS_e = 2853103$	$F = 12.9^{**}$ $MS_e = 2052$	$F = 16.4^{**}$ $MS_e = 382629$
	内 > 交 > 継	内 < 交 = 継	内 < 交 = 継	内 < 交 = 継	内 < 交 < 継	内 < 交 = 継	内 < 交 = 継	内 < 交 = 継
提供者	$F = 12.8^{**}$ $MS_e = 894$	$F = 57.5^{**}$ $MS_e = 5065$	$F = 73.3^{**}$ $MS_e = 5403$	$F = 65.7^{**}$ $MS_e = 444770$	$F = 137.4^{**}$ $MS_e = 13095975$	$F = 62.9^{**}$ $MS_e = 2788993$	$F = 7.3^{**}$ $MS_e = 1948$	$F = 25.3^{**}$ $MS_e = 262432$
	内 > 交 > 継	内 < 交 = 継	内 < 交 = 継	内 < 交 = 継	内 < 交 < 継	内 < 交 < 継	内 < 交 = 継	内 < 交 = 継
追従者	$F = 15.4^{**}$ $MS_e = 808$	$F = 35.4^{**}$ $MS_e = 5802$	$F = 52.0^{**}$ $MS_e = 6033$	$F = 50.8^{**}$ $MS_e = 468865$	$F = 63.0^{**}$ $MS_e = 14048568$	$F = 27.3^{**}$ $MS_e = 2496206$	$F = 0.0$	$F = 9.6^{**}$ $MS_e = 496452$
	内 > 交 = 継	内 < 交 = 継	内 < 交 = 継	内 < 交 < 継	内 < 交 = 継	内 < 交 = 継		内 = 交 < 継

(全体:  $df_1 = 2, df_2 = 1056$ , 提供者:  $df_1 = 2, df_2 = 551$ , 追従者:  $df_1 = 2, df_2 = 502$ )

\*\*: 1%水準で有意, 多重比較には HSD 検定を使用)

F0・F0 レンジ・最大パワー: 内部で減少, 境界で増加, (3) 継続長・ムーブ内休止時間・平均モーラ長: 内部で短縮, 境界で伸長, (4) 間隙長: 内部に比べ境界の方が長い.

課題の性質を考慮し, 第 2 ムーブの話者の役割ごとに同様の分析を行ったところ, 追従者の平均モーラ長および間隙長を除き以上の傾向がいずれの役割においてもみられることがわかった(表 2, 3). また, 各韻律特徴の数値を話者ごとの標準得点 (z-score) に変換したデータにおいても同様の傾向がみられた.

交換内部では隣接するムーブで話者が必ず交替するのに対し, 交換境界では同じ話者が継続することもあるため, 以上の傾向は交換内部・境界という構造上の差異ではなく, 話者の交替・継続というより局所的な相互作用上の差異に起因している可能性もある. そこで, 交換境界を話者の交替と継続に分類し, 境界での話者交替(境界交替), 境界での話者継続(境界継続), 内部での話者交替(内部交替)の 3 つの条件で各韻律特徴の比較を行った.

仮に, 話者交替のみが影響を与えているのであれば, 境界交替と内部交替には差はなく, 境界交替・内部交替と境界継続の間に有意差がみられることが予測される. 逆に, やり取り構造のみ関与するならば, 境界交替と境界継続には差はなく, 境界交替・継続と内部交替の間に有意差がみられるだろう.

表 4, 5 の結果から, 多くの特徴に関して境界交替と境界継続の間に有意差がみられないのに対して, 多くの特徴で境界交替と内部交替に有意差がみられることがわかった. この結果は話者交替のみが影響を与えるという仮説を退けるものである. ただし, 話者交替の有無が韻律に全く関与しないということを主張するものではない. 事実, 提供者の最小 F0・継続長・ムーブ内休止時間および追従者の最大パワーなど, いくつかの要素に関しては境界交替と境界継続に差がみられる. つまりこれらに関しては, やり取り構造上の差異に加え, 話者交替の有無が関与しているものと考えられる.

## 5 考察

2 節で指摘したように、働きかけムーブや応答兼働きかけムーブは後続要素を予測するのに対し、応答や補足は後続要素を予測しない。このことから交換内部でのムーブの移行の多くは予測的であるのに対し、交換境界での移行は非予測的であることが指摘できる。

Bull and Aylett (1998) は、本研究と同様にやり取り単位の内部に比べ境界の方が間隙長が長くなる(発話開始のタイミングが遅れる)ことを指摘し、その原因として新しいやり取りの開始時では発話のプランニングコストがより高くなることを挙げている。たしかに、交換境界では次に話すべき内容を少なくともやり取り構造のレベルでは規定しないため、交換を開始する発話のプランニングコストは高くなるものと考えられる。こういった状況下では発話のタイミングの遅れに加え、発話速度の低下や休止などの非流暢的要素の増加が予測される(Brubaker, 1972; Butterworth, 1975)。

今回分析の対象とした特徴のうち、継続長・ムーブ内休止時間・平均モーラ長・間隙長といった時間的な要素において観察された傾向、つまり(1) 交換内部のムーブ対では、第1 ムーブに比べて第2 ムーブは長さが短く、速度も速く、短い休止しが挿入されず、第1 ムーブのあと第2 ムーブがすぐに発話されるのに対し、(2) 交換境界では第2 ムーブは長く遅く、長い休止が挿入され、第2 ムーブは遅れて発話される、という傾向は、まさに上述の予測と一致するものであり、やり取り構造のもつ予測性に起因するプランニングコストの影響をこの点に見ることができる。

一方、F0 やパワーにみられる傾向については、プランニングコストの観点から説明することは難しい。これらの要素については、むしろ談話構造との関係が従来から指摘されている。Hirschberg and Nakatani (1996) は意図にもとづく談話構造と韻律との関係を調べ、談話単位冒頭では最大(平均) F0・パワーが高く、先行休止(間隙長)が長いことを、また談話単位末尾では最大(平均) F0・パワーが低く、先行休止が短く、発話速度が速いことを示している。また、Brown, Currie, and Kenworthy (1980) は、話題の開始部に比べて終結部ではピッチレンジが狭まり、パワーも低くなることを指摘している。本研究の結果はこういった先行研究の結果と完全に一致するものであり、このことは談話構造にみられる韻律の傾向が英語と日本語という言語の違いを越えて広く成立する可能性のあることを示唆する反面、以下のような問題を提起することになる。

Hirschberg and Nakatani (1996) や Brown et al. (1980) の対象とした談話構造は、発話意図や発話内容に着目したものであり、本稿で対象とした相互作用にもとづく構造とは異なる。仮にこれらの談話構造がやり取り単位よりも大きいとすると、交換境界の一部は談話単位境界と一致する。そのため、前節で指摘した韻律上の差が本来的にやり取り構造に起因するものでなかったとしても、統計上はそこに有意差が生じうる。

もちろん、こういった韻律的傾向はいずれか一方から他方が二次的に派生するものではなく、両構造に本質的に存在するという可能性もある。事実、前節でみたように、交換境界における話者交替と話者継続とは韻律特徴に差が生じることもあり(表 5)、同じ交換境界であっても切れ目の強いものと弱いものがあることがわかる。このことは、韻律の差異に対する、やり取り構造より上位の構造からの影響をうかがわせる。このような上位構造が、発話意図や発話内容にもとづく談話構造とどう対応するのかはわからない。今後、本稿でみてきた相互作用的な観点から大局的な談話構造がどのように構成されるのかを検討することで、この点を明らかにし、会話相互作用の組織化のメカニズムに対する理解を深めたい。

## 参考文献

- Brown, G., Currie, K., & Kenworthy, J. (1980). *Questions of intonation*. Baltimore: University Park Press.
- Brubaker, R. S. (1972). Rate and pause characteristics of oral reading. *Journal of Psycholinguistic Research*, 1, 141-147.
- Bull, M., & Aylett, M. (1998). An analysis of the timing of turn-taking in a corpus of goal-oriented dialogue. In *Proceedings of the 5th International Conference on Spoken Language Processing* (pp. 1179-1182).
- Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, 4, 75-87.
- Carletta, J., Isard, A., Isard, S., Kowtko, J. C., Doherty-Sneddon, C., & Anderson, A. H. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23, 13-31.
- Coulthard, M. (1977/1985). *An introduction to discourse analysis*. London: Longman.
- Hirschberg, J., & Nakatani, C. H. (1996). A prosodic analysis of discourse segments in direction-giving monologues. In *Proceedings of the 34th annual meeting of the Association for Computational Linguistics* (pp. 286-293).
- 堀内靖雄・中野有紀子・小磯花絵・石崎雅人・鈴木浩之・岡田美智男・仲真紀子・土屋俊・市川薫. (1999). 日本語地図課題対話コーパスの設計と特徴. 人工知能学会誌, 14(2).
- Schegloff, E. A. (1972). Notes on a conversational practice: Formulating place. In D. Sudnow (Ed.), *Studies in social interaction* (pp. 75-119). New York: Free Press.