

音声対話コーパスに基づく繰り返し応答の分析

下嶋篤^{*1}小磯花絵^{*1*2}Marc Swerts^{*3}片桐恭弘^{*1}^{*1}ATR 知能映像通信研究所 ^{*2}奈良先端科学技術大学院大学^{*3}IPO: Center for the Research on User-System Interaction

1 はじめに

直前の発話番 (turn) で発話されたテキストの一部を発話する繰り返し応答は、意味論的には新しい情報をほとんど付加しないにも関わらず、日常の音声対話で頻繁に行われる。本稿では、課題志向型の音声対話コーパスに基づいて、対話において繰り返し応答が担う機能を分析する。

一般に、繰り返し応答の機能については、その (1) 社会的、(2) 対話調整的、(3) 情報論的側面に着目したアプローチが考えられる。(1) は、対話参加者の間に特定の社会的状況を生み出したり、変化させたりする機能であり、例えば、Tannen (1994) は既出のテキストを繰り返すことによって生まれる参加者相互の関わり合い (involvement) を指摘し、Norrick (1994) は繰り返し応答が相互作用的に冗談 (joking) を達成すると主張している。(2) は、ある特定の目標に向けて対話の進行を調整する機能であり、従来の研究では、とくに情報の共有という観点から、情報の提示を承認する機能 (acknowledgement) と、情報提示のリペアを促す機能 (repair-initiation) とが指摘されている (Beun, 1995; 中田, 1991; Walker, 1992; 川森・島津, 1996)。

本稿では (3) の情報論的機能に着目し、音声対話の脈絡での繰り返し応答の出現によってどのような情報が運ばれるかを分析する。たしかに、繰り返し応答は、発話テキストが直前の発話番ですでに発話されていることから、進行中の会話の主題についてはほとんど情報を運ばないであろう。しかし、繰り返し応答が、メタ水準の情報、すなわち、会話のトピックではなく会話過程そのものに関する情報 (Gumperz, 1991; Koiso, Shimojima, & Katagiri, 1996) を運ぶことは十分に可能である。とくに、本稿では、開始時期、長さ、平均音高、イントネーション、平均発話速度といった、繰り返し応答の時間的・韻律的特徴に着目し、次のような仮説の検証をめざす。

繰り返し応答の時間的・韻律的特徴は、話者が、繰り返される当の情報を、すでに得ている情報の集合にどの程度まで統合できているかについての情報を運ぶ。

著者らは、先行する研究 (Swerts, Koiso, Shimojima, & Katagiri, 1997) で、繰り返し応答の時間的・韻律的特徴と上の意味での情報統合レートとの間の連関を確認している。本稿では、この分析を一步進め、(1) 厳密にどの範囲の情報統合レートが繰り返し応答の時間的・韻律的特徴によって標示されるのか、(2) それを標示するのは、厳密にどの組み合わせの時間的・韻律的特徴であるのかを調べる。さらに、繰り返し応答の情報論的機能に関する結果が、先に触れた対話調整機能の問題、とくに、Clark and Schaefer (1989) や Traum (1994) の指摘する「基盤化機能 (grounding function)」の問題について何を含意しているかも考察する。

2 方法

データ

分析の対象とした繰り返し応答は、ATR 知能映像通信研究所で収集された音声対話データから抽出された。収録された対話は、防音室の中で対面した二人の話者が、協力して積み木を組み立てるという課題を通して行われた。一方の話者 (指示者) は積み木の完成図を見ながら他の話者 (組立者) に組み立て方を口頭で指示し、組立者はその指示に従って手元の積み木を組み立てる。課題の難度は言語表現や談話構造の多様性に影響を与えるため、完成図は組立者から、積み木は指示者から見えない位置に置くことで、課題に適度な難度を持たせた。

本研究では、このうちの3対話分 (各対話15分計45分)、いずれも親しい話者間で行われたものを分析の対象とした。話者毎に分離して収録された音声データを対象に、無音で囲まれた一連の有声区間、「発話単位 (utterance units)」を抽出し、分析の基本単位とした。

繰り返し応答

ある発話番に属する発話単位の連続 X と、その直後の発話番に属する発話単位の連続 Y について、Y の半分以上を占めるモーラの連続がすでに X に現れているか、もしくは X の一部の意味論的パラフレーズであるとき、Y を X に対する繰り返し応答と呼ぶ。本稿では、直前の発話番に対する応答となっている繰り返しだけに関心があるため、たまたま直前の発話番に出てくるテキストを使いながら、それ自体が情報の提示を開始しているような例は分析から除外した。我々のデータから抽出された繰り返し応答は合計71例であった。

情報統合レートの付与

情報統合レートとは、繰り返し応答の話者が自らが繰り返している情報を、それまでに得られた情報の集合にどの程度統合できたかを、最小1から最大5までの5段階で評価したものである。まず、著者のうち三人が、繰り返し応答前後の音声と書き起こしテキストをもとに、相互に同意が得られるまで各繰り返し応答を検討し、情報統合レートを付与した (consensus labeling)。つぎに、得られた結果の信頼性をテストするため、同じ環境のもとで三人の被験者に個別に情報統合レートを付与してもらい、著者のレーティングと比較した。データは、著者のレーティングにもとづき、各レートから5例づつ計35例を無作為に抽出したものである。各被験者は同一データに対し数回レーティングを行い、最終試行の結果と著者のレーティングとの一致率を算出した。その結果、平均 $\kappa = .58$ で厳密な一致が見られ、平均 $\kappa = .84$ で ± 1 の一致が見られた。 κ は .8 以上で高い一致率があるとされことから (Siegel & Castellan, 1988; Carletta, 1996)、この結果は著者のレーティングが信頼できることを示している。

時間的・韻律的特性

繰り返し応答の時間的・韻律的特性として、次の2つのカテゴリ特性と3つの連続量特性を選んだ。カテゴリ特性は手作業で付与し、連続量特性は自動的に計測した。

繰り返し長さ 繰り返し応答は、被繰り返し応答との関係から、長と短に分類された。短とは、繰り返し応答が被繰り返し応答の一部、あるいはそのパラフレーズである場合を指す。一方長は、繰り返し応答が被繰り返し応答と同一または追加要素がある場合、あるいはそのパラフレーズである場合を指す。

境界トーン J-ToBI(Venditti, 1995)にもとづき繰り返し応答のイントネーションパターンを上昇(H%, L% H%), 下降(L%, L% H L%)の2つに分類した。ラベリングは研究の目的を知らない研究者によって行われた。

平均音高 発話単位毎に F_0 値の平均をもとめ平均音高とした。

発話速度 音素毎の固有の継続長をもとに正規化した平均モーラ長を発話単位毎にもとめ、発話速度の単位とした。

開始時期 被繰り返し応答の終了部と繰り返し応答の開始部との間の時間を開始時期とした。負の値は両発話が重複していることを、また正の値は両発話間にポーズが存在することを意味する。

先行研究(Swerts et al., 1997)では、高い平均音高、速い発話速度、遅い開始時期と情報統合レートの低さとの間に相関が見られたため、本稿では、情報統合の成功・失敗を標示する時間的・韻律的特徴の候補として、それぞれ、表1にあるものを選んだ。ここで X は、 $-0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3$ (正規化後の値)を変域としている。

表1: 標示となる時間的・韻律的特徴の候補

統合	非統合
長さ = 長, 短	長さ = 長, 短
境界トーン = High, Low	境界トーン = High, Low
発話速度 < X	発話速度 > X
開始時期 < X	開始時期 > X
平均音高 < X	平均音高 > X

本稿では、表1中の個々の特徴が話者による情報統合の程度を標示する可能性とともに、複数の特性の値が協調して、(あとで述べる意味で)より正確で包括的な標示を構成している可能性も考慮する。一般に、複数の特性の値の連言は個々の値のそれよりも高い正確さを示し、選言はより高い包括性を示すことが考えられる。以下ではとくに、 $\alpha \vee \gamma$, $\alpha \wedge \gamma$, $\alpha \vee (\gamma \wedge \delta)$, $\alpha \wedge (\gamma \vee \delta)$, $\alpha \wedge \gamma \wedge \delta$, $\alpha \vee \gamma \vee \delta$ という6形式の組み合わせを考慮する。

標示ポテンシャルの評価

一般に、ある特徴 α が別の特徴 β を標示するポテンシャルがあるかどうかは、 α が起こるときにどれくらいの割合で β が起こるか(標示の正確さ)と β が起こるときにどれくらいの割合で α が起こるか(標示の包括性)に依存する。特徴 α が特徴 β を標示する正確さ(ACC)と包括性(COM)は、それぞれ次のように量化される¹:

$$ACC(\alpha/\beta) = \frac{\alpha \text{ と } \beta \text{ が共に起こる場合の数}}{\alpha \text{ が起こる場合の数}}$$

¹ 正確さと包括性という基準は、情報検索の分野で検索プログラムの有効性を評価するために用いられる正答率(precision)と再現率(recall)という基準に合致するが、ここでは、二つの特徴の間に成立する標示関係の問題を、query, 検索プログラム, 検索対象の間に成立する検索効率の問題から区別するために、あえて、正確さ(accuracy)と包括性(comprehensiveness)という新しい術語を用いる。

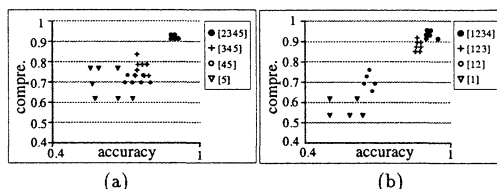


図1: 各統合範囲に対する標示ポテンシャルの分布

$$COM(\alpha/\beta) = \frac{\alpha \text{ と } \beta \text{ が共に起こる場合の数}}{\beta \text{ が起こる場合の数}}$$

バットを振るときは必ずストライクの球が来るときだが、希にしか振らないのでよくストライク球を見送る打者や、ストライク球が来たときは必ず振るがボール球にもよく手を出す打者のように、二つの特徴の標示関係の場合も、完全に正確な標示であっても包括性が極端に低ければ役に立たないし、逆に、完全に包括的な標示であっても正確さが極端に低ければ役に立たない。そこで、本研究では、同じ情報に関する二つの時間的・韻律的特徴の標示ポテンシャルを比較するときに、ACCとCOMの和を比較することではなく、ACCとCOMの低い方の値を比較することにして、正確さか包括性が極端に低いものがその情報を標示する特徴の候補からはずれるようにした。すなわち、 α_1 と α_2 という二つの時間的・韻律的特徴が β という情報を標示するポテンシャルを、次のような式に従って順序づけた:

$$\alpha_1 \supseteq \alpha_2 =_{df} \min(ACC(\alpha_1/\beta), COM(\alpha_1/\beta)) \geq \min(ACC(\alpha_2/\beta), COM(\alpha_2/\beta))$$

3 結果

標示される情報

表2は、統合範囲[1], [12], [123], [1234], [5], [45], [345], [2345]のそれぞれについて、標示ポテンシャルの大きさに応じて、時間的・韻律的特徴を順序づけた結果である。(上位5位までの候補のみ示してある。)

この表では、統合範囲[1234]が、時間的・韻律的特徴によるもっとも正確で包括的な標示の対象であり(91.67%と94.83%)、統合範囲[2345]がそれに続く(91.38%と91.38%)。統合範囲[123]もかなり高いポテンシャルの標示の対象になっている(87.76%と89.58%)。他方、[1], [12], [5], [45], [345]という統合範囲については、もっとも良い候補であっても、標示の正確性と包括性はかなり下がり、それぞれ、61.54%と61.54%, 68.97%と68.97%, 77.27%と73.91%, 66.67%と76.92%, 78.57%と78.57%であるにすぎない。

図1(a), (b)は、それぞれの統合範囲を標示する時間的・韻律的特徴の候補をその正確性と包括性に応じて配置しており、前段で指摘した標示ポテンシャルの隔たりを視覚的に表している。図1(a)は[2345]を標示する候補群と[5], [45], [345]を標示する候補群との間の標示ポテンシャルの明らかな隔たりを、図1(b)は[1234]もしくは[123]を標示する候補群と[1]もしくは[345]を標示する候補群との間の標示ポテンシャルの隔たりを示している。

つまり、我々のデータでは、繰り返し応答の時間的・韻律的特徴によって標示される統合範囲は[1234], [123], [2345]であって、[1], [12], [5], [45], [345]ではない。直観的に言えば、我々が受け取る情報は、話者が自ら繰り返し返している情報を完全に統合できていない([1234]が標示される場合)というのか、話者が情報を完全に統合しそこなっていない([2345]が標示される場合)というのか、話者がうまく情報

表 2: 各統合範囲を表示する時間的・韻律的特徴 (候補)
(LN: 長さ, BT: 境界トーン, TD: 発話速度, DL: 開始時期, TP: 発話速度).

Integration			Disintegration		
		acc. comp.			acc. comp.
[5]	[BT = L%] \wedge [DL < 0.0]	77.3% 73.9%	[1]	[LN = 短] \wedge [TP > 0.2] \wedge [PT > -0.3]	61.5% 61.5%
	[BT = L%] \wedge [PT < 0.1] \wedge [DL < 0.0]	73.9% 73.9%		[LN = 短] \wedge [TP > 0.2] \wedge [PT > -0.2]	61.5% 61.5%
	[BT = L%] \wedge [PT < 0.2] \wedge [DL < 0.0]	77.3% 73.9%		[LN = 短] \wedge [TP > 0.2] \wedge [PT > -0.1]	61.5% 61.5%
	[BT = L%] \wedge [PT < 0.3] \wedge [DL < 0.0]	73.9% 73.9%		[LN = 短] \wedge [TP > 0.2] \wedge [PT > 0.0]	63.6% 53.9%
[45]	[DL < 0.0] \wedge ([BT = L%] \vee [PT < -0.3])	70.8% 73.9%	[12]	[LN = 短] \wedge [TP > 0.2] \wedge [PT > 0.1]	63.6% 53.9%
	[TP < 0.2] \wedge [PT < 0.2] \wedge [DL < 0.2]	66.7% 76.9%		[BT = H%] \vee ([TP > -0.3] \wedge [DL > 0.2])	69.0% 69.0%
	[TP < 0.2] \wedge [PT < 0.3] \wedge [DL < 0.3]	72.7% 61.5%		[BT = H%] \vee ([TP > -0.2] \wedge [DL > 0.2])	66.7% 75.9%
	[TP < 0.2] \wedge [PT < 0.3] \wedge [DL < 0.2]	66.7% 61.5%		[BT = H%] \vee ([TP > -0.1] \wedge [DL > 0.2])	66.7% 75.9%
[345]	[TP < 0.3] \wedge [PT < 0.2] \wedge [DL < 0.2]	58.8% 76.9%	[123]	[TP > -0.3] \wedge ([PT > 0.2] \vee [DL > 0.2])	65.6% 72.4%
	[TP < -0.2] \vee ([BT = L%] \wedge [DL < 0.2])	78.6% 78.6%		[TP > 0.2] \vee [PT > 0.2] \vee [DL > 0.2]	87.8% 89.6%
	[TP < -0.2] \vee ([BT = L%] \wedge [DL < 0.3])	76.7% 78.6%		[TP > 0.2] \vee [PT > 0.3] \vee [DL > 0.2]	87.8% 89.6%
	[TP < -0.3] \vee ([BT = L%] \wedge [DL < 0.3])	75.0% 78.6%		[TP > 0.2] \vee [PT > 0.2] \vee [DL > 0.3]	87.5% 87.5%
[2345]	[TP < -0.1] \vee ([BT = L%] \wedge [DL < 0.2])	74.3% 83.3%	[1234]	[TP > 0.2] \vee [PT > 0.3] \vee [DL > 0.3]	87.5% 87.5%
	[TP < -0.1] \vee ([PT < 0.2] \wedge [DL < 0.2])	74.4% 76.2%		[TP > 0.1] \vee [PT > 0.2] \vee [DL > 0.2]	86.3% 91.7%
	[LN = 長] \vee [TP < 0.2] \vee [PT < -0.3]	91.4% 91.4%		[BT = H%] \vee [PT > 0.1] \vee [DL > 0.0]	91.7% 94.6%
	[LN = 長] \vee [TP < 0.2] \vee [PT < -0.2]	91.4% 91.4%		[BT = H%] \vee [PT > 0.2] \vee [DL > 0.0]	91.5% 93.1%
[2345]	[LN = 長] \vee [TP < 0.2] \vee [PT < -0.1]	91.4% 91.4%	[1234]	[BT = H%] \vee [PT > 0.3] \vee [DL > 0.0]	91.5% 93.1%
	[LN = 長] \vee [TP < 0.2] \vee [PT < 0.0]	90.0% 93.1%		[BT = H%] \vee [DL > 0.0]	94.6% 91.4%
	[LN = 長] \vee [TP < 0.2] \vee [PT < 0.1]	90.0% 93.1%		[BT = H%] \vee [PT > 0.0] \vee [DL > 0.0]	90.2% 94.8%

を統合していない ([123] が標示される場合) というものかであって、いずれの場合も、この標示関係を通じて運ばれる情報はかなり弱いものであると言わなければならない。

標示となる時間的・韻律的特徴

それでは、厳密にどのような時間的・韻律的特徴がこれらの情報を運ぶのであろうか。

表 2 が部分的に示しているように、統合範囲 [123] については、早い発話速度、高い平均音高、遅い開始時期の選言が上位 8 位までを占めており、この統合範囲を標示する時間的・韻律的特徴の一つであることは明らかである。

また、高い境界トーン、遅い開始時期、高い平均音高が選言的に統合範囲 [1234] を標示しているように見える。他方、上位 6 位までの候補の中には、高い境界トーンと遅い開始時期とだけから成るより単純な選言も含まれているため、[1234] の標示のために、高い平均音高が現実には何らかの役割を担っているかどうかは明らかではない。

統合範囲 [2345] については、表 2 に出ているものも含めて、上位 15 位までが長い繰り返し、遅い発話速度、低い平均音高の選言であるため、この組み合わせに強い標示ポテンシャルがあることが分かる。

ただし、いずれの結果も、上位候補の標示ポテンシャルを大きく区別しないため、発話速度、音高、開始時期の閾値を厳密に決定して、それぞれの統合範囲を標示する特徴をさらに詳しく規定することはできない。

図 2 (a)-(c) は、三つの統合範囲 [123], [1234], [2345] について、それぞれを標示する一位の時間的・韻律的特徴が 1 から 5 の統合レート上でどのように分布しているかを示している。太い折れ線はすべての繰り返し応答の分布を示し、細い折れ線はそれぞれの時間的・韻律的特徴をもった繰り返し応答の分布を表している。したがって、各グラフの灰色部分で細い折れ線が太い折れ線に沿って進み、白い部分で下降している程度に応じて、当の時間的・韻律的特徴が優れた標示であると言える。図 2 (a)-(c) では、かなりの程度これが成立しているおり、各統合範囲を標示する特徴について上で行った観察を確証している。

4 考察

繰り返し応答の基盤化機能の問題

Clark and Schaefer (1989) や Traud (1994) は、情報の共有という目的に対して、ある発話単位が担う対話調整機能をその発話単位の基盤化機能と呼んでいる。一方、本稿でこの

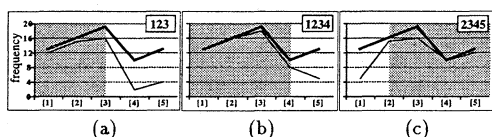


図 2: 最上位標示候補の各統合レート上での頻度

までの分析は、話者による情報統合の成功と失敗に関連して、繰り返し応答が担う情報論的機能に焦点をあてていた。それでは、繰り返し応答の情報論的機能に関するこれまでの分析結果は、繰り返し応答が担う基盤化機能の範囲について何を示唆するのであろうか。

簡単のため、統合範囲 [123], [1234], [2345] を標示する時間的・韻律的特徴を、それぞれ、 S_{123} , S_{1234} , S_{2345} と呼ぼう。(したがって、上記の分析では、 S_{123} は早い発話速度、高い平均音高、遅い開始時期からなる選言、 S_{1234} は高い境界トーン、遅い開始時期、そしておそらく高い平均音高を含んだ選言、 S_{2345} は長い繰り返し、遅い発話速度、低い平均音高からなる選言である。) さて、繰り返し応答が、承認とリベア要求という二つの基盤化機能を担うと仮定すれば、一般に、話者による高い情報統合を標示する繰り返し応答は承認の機能を担い、話者による低い情報統合を標示する繰り返し応答はリベア要求の機能を担うと考えるのが自然に思われる。このため、繰り返し応答の情報論的機能に関するこれまでの分析に基づいて、 S_{123} もしくは S_{1234} という時間的・韻律的特徴をもつ繰り返し応答は承認の機能を担い、 S_{123} という特徴を持つ繰り返し応答はリベア要求の機能を担うと結論できると思われるかもしれない。しかし、実際は、繰り返し応答の基盤化機能は、これほど単純に分類できない。

問題は、前節で有効な標示として結論された 3 つの時間的・韻律的特徴が、それらが標示する情報に関して、互いに重なりあっている点にある。すなわち、 S_{123} と S_{2345} は、どちらも、[23] という統合範囲を標示する統合範囲の中に含んでいるし、 S_{1234} と S_{2345} は、さらに大きな統合範囲 [234] で重なり合っている。このため、1 から 5 の統合レートをどのように解釈しても、常に、これらの時間的・韻律的特徴が統合成功の場合であるとも統合失敗の場合であるとも分類できない数多くの繰り返し応答の例があることになる²。

²たとえば、[123] が統合失敗を表す統合レートの範囲で、[45] が

それでは、これらの繰り返し応答には、どのような基盤化機能があるのか。問題は、自ら繰り返ししている情報を話者が統合できているかどうかという情報がこれらの繰り返し応答の場合には欠けている点である。したがって、承認とリベア要求が、話者による情報統合の成功・不成功が積極的に標示されてはじめて可能になる発話行為であると仮定すれば、承認とリベア要求がこうした繰り返し応答の担う基盤化機能であるとは考えられない。反面、数多くの繰り返し応答が、対話参加者による情報共有の過程に何の影響も及ぼさないと考えるのは直観に反している。

基盤化機能としての展示

そこで、これらの繰り返し応答が担う基盤化機能を捉えるために、展示 (display) という概念を提案したい。展示行為において、話者は、直前の発話番で対話相手が提示したと思われる情報を再生する。このような再生の目的は、相手を当の情報にもう一度直面させ、繰り返しされた情報が相手の意図したものでない場合には訂正させ、意図したものである場合には次の情報の提示に進ませることにある。

このように、承認とリベア要求が、次の情報の提示もしくはリベアという特定の行為に対話相手を積極的に導くものに対して、展示は、再生された情報が正しいかどうかに応じて、相手に適切な行為を選ばせるのであり、承認とリベア要求とは根本的に異なった行為である。したがって、承認とリベア要求の場合に、話者が繰り返しされた情報を統合したかどうかを示すことが対話相手を特定の行為に導くために重要であるのに対して、展示の目的のためには、話者が当の情報とどういう関係にあるかという問題は本質的ではない。

実際、話者が当の情報を統合したかどうかが強く示されるような場合には、繰り返し応答は対話相手に選択の余地を残すという展示の特徴を失い、承認とリベア要求になるであろう。つまり、展示は、話者による情報統合の成功・失敗に関する情報が標示されることを前提しないばかりではなく、そうした情報が不在であっても可能な行為なのである。これは、情報統合が成功した場合とも失敗した場合とも標示されない多数の繰り返し応答の存在を説明するように思われる。

ただし、どのような時間的・韻律的特徴をもつ繰り返し応答がどのような基盤化機能を担うかは、厳密には、1 から 5 までの情報統合レートをもとに解釈するかに依存する。[123] を統合成功の範囲、[45] を統合失敗の範囲と解釈すれば、上記のモデルでは、 S_{123} という時間的・韻律的特徴をもつ繰り返し応答がリベア要求、残りの繰り返し応答が展示の機能を担うことになるし、同様に、[1234] と [5] で分ければ、 S_{1234} という時間的・韻律的特徴をもつ繰り返し応答がリベア要求で残りが展示になる。他方、[1] と [2345] で分割すれば、 S_{2345} という特徴の繰り返し応答が承認で、残りが展示であることになり、最後に、[12] と [345] という分割が妥当であるとすれば、すべての繰り返し応答が展示であることになる。本稿では、繰り返し応答の担う基盤化機能の分布についてこれら 4 つの可能

統合成功を表す統合レートの範囲であるという解釈を取ったとしよう。すると、前節の結果により、たしかに、時間的・韻律的特徴 S_{123} をもつ繰り返し応答は、統合失敗の場合として標示されるが、それ以外の繰り返し応答は統合成功の場合であるとも、統合失敗の場合であるとも標示されないことになる。なぜなら、たとえそれらの一部が特徴 S_{2345} をもっている、したがって、話者の統合レートが [2345] の範囲にあると標示されたとしても、現在探っている解釈では、統合レート 2 と 3 は統合失敗を表し、4 と 5 は統合成功を表すため、それらの繰り返し応答がどちらの場合であるのかの決着はつかない。 S_{1234} という特徴をもつ繰り返し応答の例についても、標示される [1234] という範囲が 1, 2, 3 という統合レートともに 4 という統合レートを含んでいるため、やはり同じように決着がつかない。1 から 5 という統合レートを別様に分割しても、これと同様のことが成立する。

性があることを示すに留め、特定の可能性にコミットすることは避けたい。しかし、どの可能性をとっても、展示が繰り返し応答の担う基盤化機能の一つになることは確実であり、本考察の主旨は、前節までに行った繰り返し応答の情報論的分析の論理的帰結としてこの点を指摘することである。

5 おわりに

標示の正確さと包括性を基準に、日本語音声対話コーパスから抽出された繰り返し応答の情報値の統計的分析を行った。その結果、繰り返し応答の発話速度、音高、境界トーン、長さ、開始時期が、話者が繰り返しされている情報を既得の知識に統合できている程度についての情報を運ぶことが分かった。

他方、話者の情報統合レートを、統合失敗を表す範囲と統合成功を表す範囲とに一意に分割するような時間的・韻律的特徴の対は確認されなかったため、すべての繰り返し応答を、承認の役割を果たす例とリベア要求の役割を果たす例とに二分するような理論は放棄せざるを得なかった。この結果は、話者による情報統合の成功・不成功が積極的に標示されなくとも遂行できる基盤化行為の可能性を示唆しており、本稿では、展示を繰り返し応答の担う機能の一つとして仮定することで、多くの繰り返し応答の基盤化機能を自然に説明できることを示した。

参考文献

- Beun, R. J. (1995). *The function of repetitions in information dialogues* (Tech. Rep. No. 20). IPO Annual Progress Report.
- Carletta, J. (1996). Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, 22, 249-254.
- Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, 13, 259-294.
- Gumperz, J. (1991). Contextualization and understanding. In A. Daranti & C. Goodwin (Eds.), *Rethinking context*. Cambridge: Cambridge University Press.
- Koiso, H., Shimojima, A., & Katagiri, Y. (1996). Informational potentials of dynamic speech rate in dialogue. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*.
- Siegel, S., & Castellan, N. J. J. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw Hill Text.
- Swerts, M., Koiso, H., Shimojima, A., & Katagiri, Y. (1997). *Echoing in Japanese conversations* (Tech. Rep. No. TR-IT-0232). ATR Interpreting Telecommunications Research Laboratories.
- Traum, D. R. (1994). *A computational theory of grounding in natural language conversation* (Tech. Rep. No. 545). University of Rochester.
- Venditti, J. J. (1995). *Japanese ToBI labelling guides* (Tech. Rep.). Ohio State University.
- Walker, M. A. (1992). Redundancy in collaborative dialogue. In *Fourteenth International Conference on Computational Linguistics* (pp. 345-351).
- 川森雅仁・島津明. (1996). 話し言葉における冗長表現の解釈. 信学技報 nlc96-42 (pp. 31-38).
- 中田智子. (1991). 会話にあらわれるくり返しの発話. 日本語学, 10, 52-62.