

言い直しと音声強調の関連分析

7R-1

内藤 昭三

NTT ソフトウェア研究所

1 はじめに

協調作業などで交わされる人間同士の自然な対話では、コミュニケーション修復などのためにしばしば言い直しが使われる。その際に、言い直しモードであることを聞き手に伝達し、聞き手の理解を補助するために、音声強調が効果的に使われていると考えられる。本稿では、自然な対話の理解/生成モデルの構築に向けて、音声対話における言い直し現象とそれに伴う音声強調との関係について分析する。以下、2節で、分析データを説明し、3節で言い直しタイプと音声強調に関する分析結果を述べ、むすびでは言い直しと発話生成モデルとの関係について考察を加える。

2 分析データ

分析対象とした対話データは、2名の対話参加者のそれぞれに、情報提供/問い合わせ、などの目標を与え、その目標が達成できるまでに交わされた15の電話対話である。対話の被験者は、学生アルバイト、研究同僚、子どもなどである。各話者の発話は、2チャンネルを使い個別に録音されている。電話対話を分析データとしたこと理由は、録音条件を良くできることに加え、視覚情報を極力排除し、目的達成の手段は音声言語に限られるという状況では、話し手は、音声強調を含む音声言語の能力を最大限に活用しようとするだろうと考えたからである。また、今回の分析の音声特徴量としては、主にパワーパラメータを使用し、同語反復個所に対しては時間長を補助的に用いた。その理由は、単語固有度の高いピッチや時間長に比べて、パワーの使用は強調に特徴的と考えられるからである。今回の分析では、パワーの音韻固有性や被験者の年齢、性別、出身地域による方言性などは考慮に入らなかった。

(言い直し対話例1)

A: うん

B: これはあの一 えーと一 えー あ 土曜日はですね
あの一 ゴルフ接待ゴルフがはいっております
この日はえと一日じゅう千葉のほうに
でかけている予定でございます

A: あらあらあら

各対話の主題、時間長などのデータは表1の通りである。対話目標の明確度は対話データによって異なっており、いくつかの対話(たとえば、「子どもとの対話(学校で)」)では、対話参加者の一方には、必ずしも明確な目標は与えられていない。1対話あたりの平均時間長は、214秒、平均ターン数、106(「はい」のような応答も、相手発話の無音区間に発せられている場合には、独立した1個のターンとして数えた)、書き起こしテキスト(一文字2バイト、句読点も含む)の平均バイト数、5389バイトである。

¹ Analysis on the relationship between self-repairing and acoustic stress in dialogues
Shozo NAITO
NTT Software Laboratories

3 分析結果

3.1 言い直しのタイプ

言い直し個所は、83個であった。これらを、同語反復の有無、同一の構文/意味カテゴリの有無、および構文構造の観点から、以下の12タイプに分類した[1]。

1. タイプ1: 同語反復がある
 - (a) タイプ1-1: 構造A
 - (b) タイプ1-2: 構造B
 - (c) タイプ1-3: 構造C
 - (d) タイプ1-4: 構造D
2. タイプ2: 同一の構文/意味カテゴリの反復がある
 - (a) タイプ2-1: 構造A
 - (b) タイプ2-2: 構造B
 - (c) タイプ2-3: 構造C
 - (d) タイプ2-4: 構造D
3. タイプ3: 反復がない
 - (a) タイプ3-1: 語中に中断が入る
 - (b) タイプ3-2: その他

ただし、構造A, B, C, Dは以下の通りである。中断とは、ポーズ(短い場合もある)あるいは挿入表現(あー、えー、えーと、など)および相手の応答を、Rは反復個所を意味している。

1. 構造A: X 中断 Y(YがXの同語反復の場合を含む)
2. 構造B: R X 中断 R Y
3. 構造C: X R 中断 Y R
4. 構造D: A R X 中断 B R Y

対話例1では、同語反復があり(R = 「ゴルフ」)、X = nil, Y = 「接待」と対応づけることができるので、タイプ1-3に分類できる。83個の言い直しのタイプ分類を表2に示す。

3.2 音声強調

それぞれの言い直しの構造パターンにおいて、どの要素が音声的に強調されているかを分析した。2人の対話者それぞれに対する強調の閾値レベル(db値)は、以下の手順により決定した。まず、各対話者に対して、対話中でのパワーの最大値を見つけ、その値より、6db低い値を強調の閾値レベルとした。出現頻度の比較的に多かった同語あるいは同一の構文/意味カテゴリの反復のある4タイプ、1-1, 1-2, 1-3, 2-1の全77個所に対して、反復個所の前後どちらがわが強調されているかを調べた。調査結果を表3に示す。

3.3 考察

- 構造が複雑で、入れ子に言い直しが行なわれている例もみられた。以下の対話例2では、「左右対称に」という描く位置の指示が反復して行なわれており、その間に描くべき対象である「とんがった三角」に対する記述が言い直しを伴って行なわれている。

表 1: 分析データ

番号	話題	時間長 (秒)	ターン数	書き起こしテキスト長 (バイト)
1	打ち合せ時間変更	98	62	2836
2	秘書採用電話面談	188	89	4720
3	マーチャン大会のスケジュールリング	154	66	4048
4	英会話テープの売り込み	220	106	5756
5	花屋への道順	420	251	10930
6	お絵描きの指示 (ネコ)	244	70	4980
7	お絵描きの指示 (カラス)	245	65	4822
8	オーディションへの応募 (女性)	207	135	5548
9	チケットの申し込み	189	112	5646
10	オーディションへの応募 (男性)	271	84	5866
11	子どもとの対話 (学校で)	184	78	3512
12	子どもとの対話 (公園で)	82	27	1480
13	もし大金を手に入れたら	216	125	6048
14	役と体童	244	169	7472
15	風邪の予防策	245	149	7172
計		3207	1588	80836
平均		214	106	5389

表 2: 言い直しのタイプ分類結果

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	計
1-1	0	0	0	0	2	1	2	0	0	1	2	5	2	0	1	16
1-2	0	0	1	2	12	4	1	2	2	1	3	0	1	1	0	30
1-3	0	0	2	1	6	7	1	0	0	2	0	0	0	0	1	20
1-4	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	4
2-1	0	0	0	2	4	1	0	1	0	1	0	0	0	0	2	11
2-2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
2-3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2-4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
3-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
計	0	1	3	7	26	13	4	4	2	5	5	5	3	1	4	83

表 3: 反復構造の強調 / 非強調

	前方	後方	両方強調	両方非強調	計
1-1	8	7	0	1	16
1-2	12	6	2	10	30
1-3	12	0	2	6	20
2-1	2	5	0	4	11
計	34	18	4	21	77

(言い直し対話例 2)

A: はいはい

B: それから えー その一 楕円形の えー また左右対称に
えー とんがった三角を えー ちょうどこの 卵の三分の
一ぐらいの 大きさぐらいに とんがった三角を左右対称
に描いていただきたい

A: それは真中に

- 同語反復のある場合には、相対的に前の方が強調され、一方、同一の構文 / 意味カテゴリの反復のある場合には、後の方が強調される傾向にある。同語反復の場合は、新たな情報の付加はなく、単なる念押しの場合が多く、同一の構文 / 意味カテゴリの反復の場合には、新たな情報が加わり、そのことにより同一事実の理解に役立てようとする意図していることがこの差異の理由として考えられる。また、同語反復の場合に、後の出現に強調がないときには、同時に後の反復部分の発話が早口になることの多いことが観察されている。
- 言い直しに関しては、今回の分析で用いた構造パターンによる分類とは別に、機能的な分類として、間違えて言い直す、情報を付け加えて詳しく言う

ために言い直す、他の表現にして言い直す、確認のために同じ情報を繰り返す、というようなパターンが認められた。日本語の書き言葉では、情報的に重要な目的語や動詞などが、構文的には後ろの方に出現する。一方、話し言葉では、まず重要な情報を提示しようとする傾向にあり(特に緊急な場合には)、その後で時間的な余裕があれば、再度構文などの情報的には必ずしも重要ではない要因も含めた精密な処理が行なわれると考えられる。

4 むすび

音声対話における言い直し現象に着目し、音声強調との関係を分析した。言い直しは、人間の発話生成における漸進的な処理過程が表出したものと考えられる。目標を達成するための文生成が完了した後で、音声発話が行なわれるわけではなく、部分的な文生成と音声発話がバイブライニングに進んでおり、その途中で前段の文生成へのフィードバック / 再処理が起こることなどが言い直しの原因と考えられる。言い直しと音声強調との関係分析は、このような発話生成における漸進的な処理モデルの設計 / 検証においても役立つと考えられる。

謝辞 対話データを提供して頂いた小坂直敏氏、分析に協力して頂いた波山万紀子氏に感謝致します。

参考文献

- [1] Yuji Sagawa, Noboru Ohnishi and Noboru Sugie: Repairing Self-Repairs in Japanese, *Natural Language Processing Pacific Rim Symposium (NL-PRS'93)* pp.191-198, 1993.