

音声対話コーパスを用いた言い直しの分析

伝 康晴

奈良先端科学技術大学院大学
情報科学研究科

春木 憂

奈良教育大学大学院
教育学研究科

石崎 雅人

NTT 情報通信研究所
知的通信処理研究部

概要 自然な会話において生じる非流暢性の特徴を調べることは、心理言語学と計算機科学のいずれにおいても重要である。本研究では、日本語の人間同士の対話における言い直しのデータを用いて、Levelt のモニター理論の妥当性を検討する。具体的には、モニター理論から帰結する以下の2つの仮説を検証する。(i) 修復対象の末尾は必ずしも言語単位の境界ではない。(ii) 訂正部分の先頭は言語単位の境界である。分析の結果、モニター理論は概ね支持されたが、注目すべき例外が発見された。(a) 言語単位の境界で言い直りに気付きやすいということはない。(b) 機能語の言い換えは言語単位境界への後戻りを必ずしも伴わない。

A Corpus-based Analysis of Speech Repairs in Japanese

Yasuharu DEN

Graduate School of Information Science
Nara Institute of Science and Technology

Yuu HARUKI

Department of Japanese Linguistics
Nara University of Education

Masato ISHIZAKI

NTT Information and Communication Systems Laboratories

Abstract Investigating the features of speech disfluencies occurring in natural conversations is important both for psycholinguistics and for computer science. The aim of this study is to examine the monitoring theory, proposed by Levelt, using speech repair data in Japanese human-human dialogues. Two hypotheses directly following from the monitoring theory are examined: (i) the end of reparanda do not necessarily match with the end of linguistic units, and (ii) the beginning of repair texts match with the beginning of linguistic units. Notable exceptions for the theory are found concerning the timing of trouble detection and the nonretracing strategy in substitution of function words.

1 はじめに

自然な会話は、言い淀み、言い直し、繰り返し、言い誤りなどのさまざまな非流暢性を含んでいる。これらの非流暢性の特徴を調べることは、人間の言語産出機構の解明と音声対話システムの開発のいずれのためにも重要である。これまで、言い直しに関するさまざまな研究が、心理言語学 (Levelt, 1983; van Wijk & Kempen, 1987)、会話分析 (Schegloff, Jefferson, & Sacks, 1977; Schegloff, 1987)、計算言語学 (Hindle, 1983; Nakatani & Hirschberg, 1994) の分野でなされてきた。このうち、心理言語学者 Levelt によるモニター理論 (monitoring theory) は人間の音声言語産出の理論にもっとも影響を与えたものの一つである。本研究の

目的は、日本語の人間同士の対話における言い直しのデータを用いて、モニター理論の妥当性を検討することである。

Levelt (1983) は言い直しの形態に関して2つの制約を発見した。

Main Interruption Rule (MIR) 発話の中断は言い誤りに気付くとただちになされ、中断位置の言語的な適切性とは無関係である。

Well-Formedness Rule (WFR) 言い直す前の発話と言い直した後の発話が等位構造をなすように、中断前の発話の一部反復される。

MIR は、話者が言い誤りに気付いたときにどのタイミングで発話を中断するかに関するものであり、WFR は、話者が中断のあとどのよう

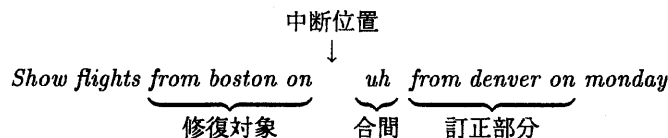


図 1: 言い直しの構造 ((Shriberg, 1994) より)

に発話を再開するかに関するものである。これらに加えて、誤りが句の終末部で気付かれやすいことを、Levelt は見いだしている。

人間の認知機構の一般性から、これらの規則は Levelt が対象としていたオランダ語以外にも当てはまるはずである。これまで広く研究されてきたオランダ語や英語とはまったく異なる言語構造を持つ日本語は、これらの規則の一般性を検討する良い題材である。

本稿では、モニター理論の MIR と WFR から帰結する以下の 2 つの仮説を検証する。

1. 修復対象の末尾は必ずしも言語単位の境界ではない。
2. 訂正部分の先頭は言語単位の境界である。

ここで、修復対象 (reparandum) とは、言い直しによって取り消される部分であり、言い直しの開始位置から中断位置までを指す。また、訂正部分 (repair) とは、修復対象を訂正する部分であり、フィラーや編集語句をはさんで、中断位置から言い直しの終了位置までを指す。図 1 の例では、修復対象の末尾は前置詞 *on* の後であり、これは言語単位 (主要な句) の境界ではない。また、訂正部分の先頭は前置詞 *from* の前であり、これは言語単位の境界である。

訂正部分の先頭が言語単位の境界に一致することは、言い直しにおいて話者が意図的な後戻りを行っている (単に *boston* を *denver* で置き換えるだけでなく、言語的に適切な位置である *from* まで後戻りして発話を再開する) ことを示している。van Wijk と Kempen (1987) は、オランダ語の言い直しの分析によって、このような後戻りに 2 種類の方略があることを見いだした。一つは、主要な統語構成素の先頭への後戻りであり、もう一つは、音韻句の先頭への後戻りである。ところで、日本語においては、これら 2 種類の言語単位は概ね等しく、「文節」と呼ばれる単位に相当する。そこで、本稿では、

van Wijk と Kempen のいう方略の違いに関わりなく、常に言語単位として文節を考える。

以下、まず、2 節で分析に用いたデータと方法について述べ、次に、3, 4 節でそれぞれ、仮説 1, 2 を検証し、その結果を 5 節で考察する。

2 分析方法

2.1 データ

分析に用いたデータは、ATR 経路探索課題コーパスの (異なる 60 話者による) 30 対話である (課題の内容については (石崎, 伝, 土屋, 田本, 中里, 1995) の 2 節を参照)。すべての対話は DAT とビデオテープに収録され、テキストに書き起こされている。

2.2 品詞の付与

言い直しの形態的な特徴を詳細に分析するため、書き起こしテキストを形態素解析し、品詞を付与した。この作業は、形態素解析システム「茶釜」(松本, 北内, 山下, 今一, 今村, 1997) の出力を手で修正することで行い、結果はすべて著者のうちの最低 2 人が検査した。

2.3 言い直しのラベルづけ

言い直しは、品詞付与されたテキストをもとにすべて第 1 著者が抽出し、Shriberg (1994) のパターン・ラベルづけ体系を用いてラベルづけした。なお、本研究では、同一話者の同一発話内で生じた言い直しのみを対象とする。

2.4 要約統計

抽出された言い直しの個数を表 1 にまとめる¹。これらは、フィラーが単独で生じた (FP

¹ 「横にい, い, 横に行って」のように複数の言い直しは複合している場合を「複合的」と呼び、それ以外を「基本的」と呼ぶ。また、言い直しの分類は、ART (音韻誤りを含むもの)、SUB (語句の置換)、INS (語句の挿入)、DEL (語句の削除)、REP (語句の繰り返し)、HYB (2 種類以上のタイプの複合) の 6 種類である。

表 1: 要約統計

単語数	73258
発話数	13630
言い直しの数	1578
- 基本的	1420
- 分類可能	1375
- 分類不可	45
- 複合的	158

表 2: 基本的で分類可能な言い直しの内訳

ART	SUB	INS	DEL	REP	HYB
29	336	138	264	580	28

タイプ) 3050 例を含まない。

このうち、基本的で分類可能な 1375 例の内訳は、表 2 の通りである。以下の分析では、これらの 1375 例を対象とする。

3 仮説 1 の検証

3.1 修復対象の末尾における品詞の分布

仮説 1 の検証のために、修復対象の末尾における品詞の分布を調べる。もし仮説 1 が正しいなら、修復対象の末尾において、言語単位境界になる場合とならない場合の比率は、コーパス全体²の品詞分布から予測されるものと違わないはずである。

この分析においては、各形態素を以下の 2 つのクラスに分類した。(A) 文節末に現れるもの、(B) 文節末に現れないもの。分類の方法は以下の通りである。まず、常に文節末に現れる品詞(副詞、連体詞、活用語の連体形・命令形)を持つものは A 類に、決して文節末に現れない品詞(活用語の語幹・未然形・音便連用形・仮定形)を持つものは B 類に分類する。これで決まらないものは、後続形態素を見る(新たな文節が開始されていたら当該形態素は A 類、そうでなければ B 類)ことで判断を行った。発話末にあり、この判断ができないものについては、活用語の終止形・連用形と助詞は A 類に分類し、名詞は分析対象から除外した。

分析結果を表 3 に示す。期待度数は、コーパ

表 3: 修復対象の末尾における品詞

	A 類	B 類
観測度数	718	579
期待度数	581.9	715.1

表 4: 修復対象の末尾における品詞(タイプ別)

	A 類	B 類	A 類	B 類
ART	8	18	DEL	32
SUB	159	160	REP	287
INS	56	71	HYB	11

ス全体の品詞分布(A類:B類 = 44.9%:55.1%)から算出した。A類に属する場合の数は有意に大きく($\chi^2(1) = 57.7, p = 0$)、修復対象の末尾が文節末に一致する場合が多い。これは、仮説 1 を反証するかのように見える。

ところが、言い直しのタイプごとにこの分析を行うと(表 4, 観測度数のみ)、DELタイプ³だけがこの傾向を示すことが分かった($p = 0$)。他のタイプにおいては A 類が多い傾向は見られない。すなわち、DELタイプを除くすべてのタイプの言い直しで仮説 1 が成り立つ。

3.2 言語単位境界への拡張の有無

前節の結果は、DELタイプを除き仮説 1 が成り立つことを示すものであった。ここでは、修復対象の末尾を言語単位境界へ一致させようという話者の意図があるかどうかをより直接的に検討するために、言語単位境界への拡張の有無を対比させる以下の 2 種類のパターンの言い直しを対象として分析を行う(‘ ’ は中断位置)。

I 型 [東京に・大阪に]行く

II 型 [東京・大阪]に行く

いずれの場合も「東京」が「大阪」に置換されているが、I型では、中断前の発話を助詞「に」まで拡張することによって、修復対象の末尾を文節境界に一致させている。II型では、そのような拡張は見られず、修復対象は文節の途中で終わっている。

²発話末は修復対象の末尾になり得ないので除く。

³DELタイプは、「高知の左があら、あ、下に今下がってます」のような fresh start に典型的に見られる。

表 5: 言語単位境界への拡張の有無

	I型	II型
REP	42	282
SUB	34	158
INS	14	70

本分析では、上例のような内容語の繰り返し、置換、挿入を含む言い直しにおいて、修復対象の末尾を文節境界に一致させるための機能語の追加があるかどうかを調べる。もし修復対象の末尾を文節境界へ一致させようという話者の意図があるなら、I型の場合の数がII型より多いはずである。

タイプ別の分析結果を表5に示す。いずれのタイプにおいても、II型のほうがI型より有意に多い(いずれも $p = 0$)。したがって、修復対象の末尾を言語単位境界へ一致させようという話者の意図はなく、話者は中断位置が言語単位境界であるかどうかに関わらず直ちに言い誤った発話を停止する。

3.3 言い誤りに気付く位置

Levelt (1983) は MIR に加えて、言い誤りが句の終末部で気付かれやすいことを見いだしている。3.1節の結果(修復対象の末尾は必ずしも言語単位境界でない)は、これに反するかのように見える。しかし、これは、単に内容語(文節の途中であることが多い)を言い誤りやすい、という事実を表しているだけの可能性がある。実際、誤り語が明示的にラベルづけされている SUB タイプの言い直し⁴について、誤り語(誤り部分が複数語からなる時はその末尾の語)の品詞を調べたところ、内容語が圧倒的に多かった(内容語: 275 vs. 機能語: 61)。

そこで、内容語の置換を含む SUB タイプの言い直しについて、誤りに気付くまでの遅延距離と中断位置(= 修復対象の末尾)にある品詞との関係を調べた。Levelt の観察が日本語にも

⁴Shriberg (1994) のパターン・ラベルづけ体系では、例えば、図1の言い直しは、以下のようにラベルづけされる。この場合、ラベルから *boston* が誤りと分かる。

from boston on uh from denver on
r s r . f r s r

表 6: 遅延距離と修復対象の末尾の品詞

	A類	B類
同一文節内	95	139
1文節遅延	12	8
2文節遅延	1	1
3文節遅延	1	1

当てはまるなら、遅延距離が長くなっても(あるいは長くなるにつれて)文節末での中断が多いはずである。

分析結果を表6に示す。まず、ほとんどの場合に同一文節内で誤りに気付いて発話を中断していることに注意しよう(同一文節内: 234 vs. 遅延: 24)。このため、遅延が生じている場合の数が著しく少なく、信頼度に問題がないとはいえないが、コーパス全体の品詞分布と比べて、どちらが多いということはない(A類: 14 vs. B類: 10, $p = .22$, フィッシャーの正確確率)。また、遅延の有無と品詞の分布には有意な連関がなく($|\phi| = .1$, $p = .13$)、遅延があるときほど文節末での中断が多いということもない。よって、Levelt の観察は日本語には当てはまらない。

3.4 まとめ

修復対象を言語単位境界まで拡張する傾向は見られず、Levelt (1983) の MIR が支持された。さらに、仮説1も(DELタイプを除く)ほとんどのタイプの言い直しで成立し、言語単位境界で言い誤りに気付きやすいという Levelt の観察は日本語には当てはまらないことが分かった。

4 仮説2の検証

4.1 訂正部分の先頭における品詞の分布

仮説2の検証のために、訂正部分の先頭における品詞の分布を調べる。もし仮説2が正しいなら、訂正部分の先頭において、言語単位境界になる場合とならない場合の比率は、コーパス全体⁵の品詞分布から予測されるものと比べて、前者のほうが多いはずである。

この分析においては、各形態素を以下の2つのクラスに分類した。(C) 文節頭に現れるもの、(D) 文節頭に現れないもの。これらのクラスは

⁵発話頭は訂正部分の先頭になり得ないので除く。

表 7: 訂正部分の先頭における品詞

	C類	D類
観測度数	1093	30
期待度数	515.2	607.8

表 8: 訂正部分の先頭における品詞 (タイプ別)

	C類	D類	C類	D類
ART	28	0	DEL	19
SUB	311	25	REP	572
INS	137	1	HYB	26

それぞれ内容語と機能語のクラスに対応する。分析結果を表 7 に示す。期待度数は、コーパス全体の品詞分布 (C類:D類 = 45.9%:54.1%) から算出した。C類に属する場合の数は有意に大きく ($\chi^2(1) = 1197.3, p = 0$)、訂正部分の先頭が文節頭に一致する場合が多い。また、この傾向は、すべてのタイプの言い直しについて見られる (表 8, 観測度数のみ)。これらの結果は、仮説 2 を支持する。

4.2 言語単位境界への後戻りの有無

前節の結果は、仮説 2 が成り立つことを示すものであった。ここでは、訂正部分の先頭を言語単位境界へ一致させようという話者の意図があるかどうかをより直接的に検討するために、言語単位境界への後戻りの有無を対比させる以下の 2 種類のパターンの言い直しを対象として分析を行う (‘は中断位置)。

III型 [東京に.東京の]上に行く

IV型 東京[に.の]上に行く

いずれの場合も「に」が「の」に置換されているが、III型では、中断後の発話を名詞「東京」まで後戻りして開始することによって、訂正部分の先頭を文節境界に一致させている。IV型では、そのような後戻りは見られず、訂正部分は文節の途中から始まっている。

本分析では、上例のような機能語の繰り返し、置換、削除を含む言い直しにおいて、訂正部分の先頭を文節境界に一致させるための内容語への後戻りがあるかどうかを調べる。もし訂正部

表 9: 言語単位境界への後戻りの有無

	III型	IV型
REP	85	2
SUB	28	24
DEL	11	0

分の先頭を文節境界へ一致させようという話者の意図があるなら、III型の場合の数はIV型より多いはずである。

タイプ別の分析結果を表 9 に示す。REPタイプとDELタイプでは、III型のほうがIV型より有意に多い (REP: $p = 0$, DEL: $p < .001$) が、SUBタイプでは、両者の差は有意でない ($p = .68$)。すなわち、機能語の置換においては、訂正部分の先頭を言語単位境界へ一致させようという話者の意図が働かず、話者は誤りを直ちに訂正する方略を普通に使っている。

4.3 まとめ

仮説 2 は部分的にしか支持されなかった。多くの場合では、訂正部分を言語単位境界まで後戻りして開始する傾向が見られたが、機能語の置換を含む言い直しは例外であった。機能語の置換においては、言語単位境界まで後戻りすることなく、誤りを直ちに訂正する方略が、普通に使われている。これはまた、言い直しが主要な句による等位構造と類似の構造を持つとする Levelt (1983) の WFR に対する例外でもある (文節境界でなければ句境界でないから)。

5 議論

以上の分析結果は、モニター理論が概ね日本語にも当てはまることを示すものであるが、以下の注目すべき例外があることも分かった。

1. 言語単位の境界で言い誤りに気付きやすいということはない。
2. 機能語の言い換えは言語単位境界への後戻りを必ずしも伴わない。

以下では、これらの知見の含意を検討する。

まず、1 に関しては、オランダ語では通常、内容語が言語単位 (主要な句) の末尾に位置するが、日本語では通常、内容語が言語単位 (文

節)の先頭に位置する。したがって、オランダ語では「言語単位境界で中断すること」は「内容語で中断すること」をほぼ含意する。よって、Leveltの観察からいえるのは、「言語単位境界で誤りに気づきやすい」のではなく、「内容語で誤りに気づきやすい」ということかも知れない(あるいは、その両者かも知れない)。確かにいえるのは、「言語単位境界にない機能語では気づきにくい」という点である。我々のコーパスに基づく分析では、「内容語で気づきやすい」という結果は得られなかった(表6と同様の分析を内容語 vs. 機能語で行った)が、少なくとも「言語単位境界にない機能語では気づきにくい」という点は確認された(SUBタイプの言い直しのうち機能語の連続を含む10例について、最初の機能語で中断されなかった場合が7例あった)。

一方、2に関しては、オランダ語では機能語が言語単位の先頭に位置するので、機能語の言い換えは必ず言語単位の先頭から始まることになる。すなわち、機能語の言い換えで言語単位境界から始まらない例は、オランダ語ではその言語構造上観察できないのである(理論上は以前の言語単位の途中まで後戻りの例がありうるが、現実的でない)。そのために、本分析で見いだしたような後戻りなしの機能語の言い換えが見逃ぎされていたのかも知れない。

いずれにせよ、言語構造上の違いによって、オランダ語や英語では観察されない現象が日本語で観察されたり、その逆の場合もありうるので、今後、構造の異なる言語に対する分析を相補的に用いることによって、より良い一般化を追求する必要がある。

6 おわりに

本稿では、日本語の人間同士の対話における言い直しのデータを用いて、Leveltのモニター理論の妥当性を検討した。分析結果は、モニター理論が概ね日本語にも当てはまることを示すものであったが、例外も発見された。それらは、日本語とオランダ語の言語構造の違いに起因するものと思われる。

音声対話システムにおける非流暢性の機械的処理への応用という観点から見ると、本研究の知見は伝(1997)の手法の前提となっている文

節単位の処理の有効性を概ね支持するものである。ただし、伝が問題点としてあげている文節途中での中断や後戻りのない機能語の言い換えなどは、実際のコーパス中で広く観察されることが確認されたので、これらへの対処が非流暢性の機械的処理において今後重要になろう。

謝辞 コーパスへの品詞付与作業に協力していただいた桐山和久さんと山下達雄さんに感謝します。本研究はATR 音声翻訳通信研究所からの受託研究に基づいています。

参考文献

- 伝康晴(1997). 統一モデルに基づく話し言葉の解析. 自然言語処理, 4 (1), 23-40.
- Hindle, D. (1983). Deterministic parsing of syntactic non-fluencies. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, pp. 123-128.
- 石崎雅人, 伝康晴, 土屋俊, 田本真詞, 中里収(1995). 対話課題の分類試案. 情報処理学会研究報告, 95-SLP-9, 77-85.
- Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition*, 14, 41-104.
- 松本裕治, 北内啓, 山下達雄, 今一修, 今村友明(1997). 日本語形態素解析システム【茶釜】version 1.0 使用説明書. テクニカル・レポート NAIST-IS-TR97007, 奈良先端科学技術大学院大学.
- Nakatani, C. H. & Hirschberg, J. (1994). A corpus-based study of repair cues in spontaneous speech. *Journal of the Acoustical Society of America*, 95, 1603-1616.
- Schegloff, E. A. (1987). Recycled turn beginnings: A precise repair mechanism in conversation's turn-taking organization. In Button, G. & Lee, J. R. E. (Eds.), *Talk and Social Organisation*, pp. 70-85. Multilingual Matters Ltd., Clevedon.
- Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 53, 361-382.
- Shriberg, E. E. (1994). *Preliminaries to a Theory of Speech Disfluencies*. Doctoral dissertation, University of California at Berkeley.
- van Wijk, C. & Kempen, G. (1987). A dual system for producing self-repairs in spontaneous speech: Evidence from experimentally elicited corrections. *Cognitive Psychology*, 19, 403-440.