

## 自然な発話における言い間違いに関する考察

土井 晃一\*

\*株式会社 富士通研究所 情報社会科学研究所  
〒261 千葉県千葉市美浜区中瀬1丁目9番地3号(幕張システムラボラトリ)  
(043)299-3100  
doy@iias.flab.fujitsu.co.jp

あらまし

ソフトウェアの要求獲得会議では会議参加者の関心のあることをきちんと掘起こすことが重要である。関心のあることを掘起こすためには会議参加者の無意識の構造を知る方法が考えられる。無意識の兆候としては古来から言い間違いが挙げられている。言い間違いを利用するとしても、言い間違いを解釈するやり方は高度の技術を要するようである。そこで本研究では、言い間違い・言い直しを解釈しないで利用する方法を考えることにする。そこでまず、言い間違いの傾向について調べることにする。本論文では、言い直しの時間的推移、言い間違いと発想力の関係、言い直した語と言い直された語との間でどちらに関心が高いかを調べた。

キーワード 言い間違い、言い直し、会議、関心、無意識

## A Study of Some Mistakes in Spontaneous Speaking

Kouichi DOI\*

\*Fujitsu Laboratories, Institute for Social Information Science  
9-3, Nakase 1-chome, Mihama-ku, Chiba-shi, Chiba 261, Japan  
043(299)-3100  
doy@iias.flab.fujitsu.co.jp

Abstract

In software requirements capturing meeting, it is important to find out interesting items of attendants of the meeting. There is a method of finding out interesting items by knowing their structure of unconsciousness. There are some indications of unconsciousness, which is, for example, some mistakes in speaking. To interpret mistakes in speaking seems to need high technic. We aim at the method of not interpreting the mistakes or corrects in speaking. We research the tendencies of the mistakes in speaking. In this paper, we research the time transition of mistakes in speaking, the relationship between the mistakes in speaking and the ability of inspiration, and the relationship between interest and corrects in speaking.

key words mistakes in speaking, corrects in speaking, meeting, interest, unconsciousness

## 1 はじめに

まず、言い間違いの原因について考察してみる。フロイト [1] は言い間違いの原因として身体的理由と精神的理由を挙げている。フロイトは身体的理由として、

1. 気分が悪い・疲れ気味である
2. あがっている
3. 注意が他にそれている

を挙げている。1は確かに身体的理由であるが、2,3はむしろその場の精神的理由である。フロイトが言いたいことは、確かに上記のような身体的理由があるにしろ、言い間違いが生じている時は必ず何らかの深層心理的・無意識の理由があるということである。フロイトは深層心理的・無意識の理由のない言い間違いはありえない、つまり偶然生じる言い間違いはあり得ないと断言している。さらにフロイトは言い間違いで探索すべき概念の範囲として、似た言葉(発音・言語類似・言語連想)と反対の意味の言葉を挙げている。

しかしながら、あまりよく知らない単語であったり、関心が薄い単語であれば言い間違えることが考えられる。

また、ラカンの流れを汲む Nasio は、無意識は相互作用であり、コミュニケーションあるいは精神分析の中でしか無意識は存在しないと述べている [2]。これは精神分析者が被精神分析者の無意識を被精神分析者に示し、理解させ、了解を求めながら精神分析が進んでいくということを意味しているものと思われる。その無意識の兆候の一つとして挙げられるのが言い間違いである。つまり言い間違いのすべてが無意識を顕現化しているものではない。

このような無意識を第三者が観察することで見出すことは可能であろうか？ もし可能であれば、会議支援につながる。会議参加者が意識的には気づいていないが無意識的に重要だと思っていることを会議へフィードバックすることができるからである。しかるに言い間違いは無意識の兆候を示しているのであるから、言い間違いを調べることによって会議支援ができることが期待できる。

しかし、前述のような精神分析の方法は分析者の解釈がどうしても必要であり、かなりの能力が必要となり、誰にでもできるというわけにはいかない。しかも、その解釈にはかなり主観的要素がつきまとう。実際の言い間違いの利用方法には、

1. 解釈しない(客観的)
2. 解釈する(主観的)

の二種類が考えられる。前者は言い間違えた事実だけを客観的に使う方法であり、後者は言い間違いを解釈

して使う方法である。我々は解釈には分析者の主観性が強く現れすぎて、結果が恣意的になると考え、前者の方法を採用する。

言い間違いに関する用語を定義しておく。ソフトウェアの要求獲得会議のコーパスから言い間違いの例を挙げると、

言い淀み: 「電話で何だけ、留守番電話みたいに」  
言い直し: 「たとえば何らかのシステムが出て、出たとしても」  
言い間違い: 「自分の手帳でやってや書くでしょう」

のようになる。言い間違いにはいわゆる言い間違い、言い淀み、言い直しなどが含まれる。本論文では、言い淀みとは不要な語句(感動詞を含む)が挿入された発話を指すことにし、言い直しとは途中で発言が中断され別の語句に発話し直したことを指すことにし、言い間違いとは言い淀み・言い直し以外の言い間違いのことを指すことにする。言い直し以外の言い間違いを利用するためにはどうしても解釈する必要が出てくる。我々は客観的に分析するという観点から、主として、言い直しに限って分析を進める。

また、前述のように言い間違いのすべてが無意識の兆候になっているかどうかは若干の疑念がある。そこで本論文では言い間違いの傾向を統計的に調べることにする。

以下、第2節では会議における言い直しの時間的推移について検討する。第3節では言い間違いと発想力との関連について述べる。第4節では言い直しと関心の高さとの関連について述べる。

## 2 言い直しの時間的推移

本節では、会議の中での言い直しの回数の時間的推移を調べてみる。会議を10分ごとに区切り、言い直しの回数を測定してみる。図1がその様子である。会議は前半(70分)と後半(90分)の二つに分かれ、前半と後半の間には約10分の休憩時間が設けられている。図1中、縦軸は回数(単位は回)、横軸は時間(単位は分)であり、実線が前半、点線が後半の様子を示す。

このデータを取って得た知見をまとめておくと、一口に言い直しと言ってもいろいろな種類があることが分かった。言い直しのなかでも簡単に理由の分かるものとしては、

1. 相手の発言に妨げられたもの
2. 強調しようとしたもの

が挙げられる。しかしこれらをコーパスだけの情報を用いて分類することは困難である。さらに会議に臨席していてもその判断は主観的になる。また長い発言の

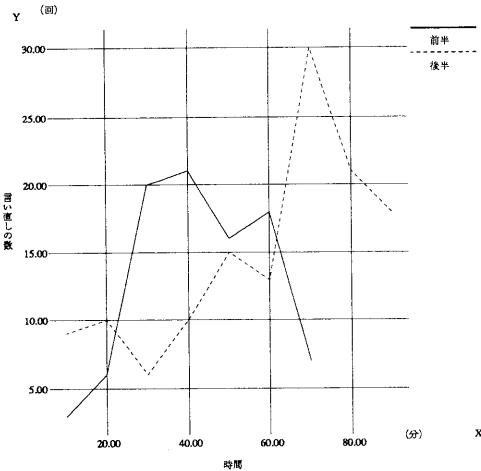


図 1: 第一回目の会議での言い直しの頻度の推移

中の言い直しは短い発言の中の言い直しに比べて選別しにくい。これは言い直しの範囲の大きさが大きくなるからと考えられる。

図 1を見ていて気がつくことは、会議の最初と最後で言い直しの回数が小さく、会議の中頃で言い直しの回数が大きくなる。この原因としてまず考えられることは、特定の個人が集中的に言い直しているためであることが考えられる。そこで個人の言い直しのデータを見てみると図 2 のようになる。

図中縦軸は時間を 10 分毎に区切ったもの、横軸は個人を表す。図中の数字は言い直しの回数を表す。○で囲んだ数字は横の欄でのモードを表す。モードを見る限りでは、特定の個人が全体に影響を与えているようには見受けられない。総発話量の単位はコーパス上でのバイト数を表す。ほぼ発話量に対応するものと思われる。単位発話量当たりの言い直しの値を出してみると ((1)/(2))、発話量の多い人が、言い直しも多く、しかも単位発話量当たりの言い直しの値も大きいことがわかる。このことからわかることは発話量の多い人は間違っているかどうかをあまり考えずに何か発話して、間違っていたら言い直しているのではないかと考えられる。測定はしていないが、言い直しの多い人は発話の速度も早いように見受けられた。無意識に依存する言い直しがあるかどうかはこれだけの結果からだけでは明らかとはならない。

なお総発話数との相関は低かった。このことは相手の発言に妨げられたものの割合がそうでないものに対してそれほど大きくないか、あるいは相手の発言に妨げられたものは時間とは関係なく一定であることが予想される。

さらに総発話量の推移を見てみると、図 3 のようになる。

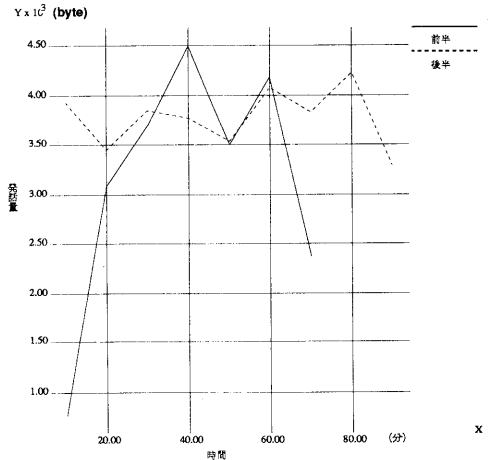


図 3: 第一回目の会議での発話量の推移

図の見方はほぼ図 1 と同じである。違うところは縦軸が発話量 (コーパスで見た時のバイト数) になっているところである。

図 1 と図 3 を見比べて気がつくことは、言い直しのグラフと形が似ていることである。特に前半の 40 分から 70 分にかけてと、後半の 70 分付近は増減の具合と、全体のモード値を取っているという点で似ている。このことから発話量と言い直しとの関係の深さがわかる。

### 3 言い間違いと発想力

言い間違いと発想力の関係を調べるため、会議参加者に簡単な心理学実験を行なった。まず、言い間違いのあった単語をコーパスから見つけ出す。それらの単語を提示する単語とし、各々の単語から連想される単語を書いてもらうということで実験を行なった。各々の単語に対する制限時間は 30 秒である。図 4 に結果を示す。

図中、縦軸は個人名、横軸は提示した単語である。さらに\*は発話誤りのあったところを示す。この会議に関係あるものを書くようにと教示したにも関わらず、関係ないものを書いた人がいたため、関係のあるものだけを集計の際に選別した。図中斜線の左側が実際に書かれた単語数、右側が関係のある単語数である。

30 秒という制限時間が短か過ぎたので、書かれた単語の数が高々 7 しかないが大体の傾向はわかる。

この生データを元に、ある人の最大の発想数を 100 として正規化し、さらに縦の合計を出すすと以下のようなになる。

発想の高い方から並べ直すと、

	U	V	W	X	Y	Z	合計
10	②	1	0	0	0	0	3
20	②	0	②	1	1	0	6
30	1	⑦	3	1	5	3	20
40	1	⑩	6	0	3	1	21
50	⑥	2	⑥	1	1	0	16
60	2	2	2	0	⑦	5	18
70	⑥	0	1	0	0	0	7
小計	20	22	20	3	17	9	91
10	2	1	④	1	0	1	9
20	3	④	1	1	0	1	10
30	1	0	2	0	③	0	6
40	③	0	2	0	③	2	10
50	0	3	4	0	⑤	3	15
60	1	1	0	0	⑧	3	13
70	2	6	12	0	2	⑧	30
80	⑩	0	1	0	0	2	21
90	⑩	2	3	0	0	3	18
合計(1)	60	39	49	5	38	32	223
百分率	27%	17%	22%	2%	17%	14%	100%
総発話量(2)	11189	11646	9793	3673	8836	9921	
(1)/(2)	0.00536	0.00334	0.00500	0.00136	0.00430	0.00322	

図 2: 個人の言い直しの頻度

	留守番電話	息抜き	在席	スケジュール	カメラ	全自動	合計(1)	平均	有効合計(2)	(2)/(1)
V	2/1*	4/2*	4/4*	4/4*	5/1	3/2	22	3.6	14	64%
Y	6/3	7/2	6/5*	6/6*	7/2*	3/1	34	5.6	19	56%
X	3/2	5/1	5/4	4/2*	3/0*	2/0	22	3.6	9	41%
W	3/3	6/6	3/2*	7/7	5/2	4/1*	28	4.6	21	75%
Z	3/3	4/4	5/5*	5/5	3/3	2/2	22	3.6	22	100%
T	1/1	2/2	1/1	1/1	2/2	1/1	8	1.3	8	100%
合計(3)	18	28	24	27	25	15				
平均	3.0	4.6	4.0	4.8	4.1	2.5				
有効合計(4)	13	17	21	25	10	7				
(4)/(3)	72%	61%	88%	93%	40%	47%				

図 4: 言い誤りと連想する単語数

	留守番電話	息抜き	在席	スケジュール	カメラ	全自動
V	25*	50*	100*	100*	25	50
Y	50	33	83*	100*	33*	17
X	50	25	100	50*	0*	0
W	43	86	29*	100	29	14*
Z	60	80	100*	100	60	40
T	50	100	50	50	100	50
合計	278	374	462	500	247	171

図 5: 正規化したもの

有意水準	0.5	0.1	0.01
1	3.841	6.635	10.827

図 7:  $\chi^2$ 検定の有意水準

スケジュール > 在席 > 息抜き >  
留守番電話 > カメラ > 全自動

となる。

関心が高いことと発想が大きいことを同じことと考えると、これは関心の高い順を示す。これは次の会議で話す内容を決めるアンケートとほぼ同じ順序であった。また関心の高いものほど言い間違える人数も大きい傾向があることが言える。

その会議における言い間違いをしたその人の頻度を見てみると、

留守番電話: 1(297)  
息抜き: 2(173)  
在席: 12(13),8(21),10(18),9(18)  
スケジュール: 19(3),8(21),2(49)  
カメラ: 9(16),3(29)  
全自動: 1(250)

数字は度数、括弧の中の数字は順位を表す。関心の高いものほど頻度が高く、順位も高い傾向が見られた。

#### 4 言い直しと関心の高さ

会議参加者の関心度を知るために次のような実験を考える。言い直しの中で、名詞レベルの言い直しに絞り(それ以上のものは解釈に曖昧性が大きいので)、その中でも、二つの言葉の差異が適度に大きいものを抽出した。

図 6 に結果を示す。図中、A は言い直す前の単語、B は言い直した後の単語である。(a) は単語レベル(文節レベルではない)の言い替え、(b) はさらに直後の言い替え(文節レベルの言い直しで、言い直し中に他の発話が入らない場合)を示す。\*は乱数で選んで前後を反転させたものを示す。T,V,W,X,Y,Z は個人名を示す。U は司会者であるので、ここでは測定の対象としなかった。○で囲んだ A,B はその人に言い直しがあったことを示す。一番下の合計欄は、(a),(b) は該当数、括弧の中は A,B のそれぞれの合計を示す。個人の合計は言い直しがあったもので、括弧の中は A,B のそれぞれの合計を示す。

この結果を  $\chi^2$  検定で統計的に確かめる。有意水準を挙げておくと、図 7 のようになる。

まず全体的な A と B の比率は、

$$A : B = 7 : 11$$

であり、

$$(9-7)^2/9 + (11-9)^2/9 = 0.889$$

有意とは言えない。また、(a) も、

$$\alpha; A : B = 4 : 9$$

$$(9-6.5)^2/6.5 + (4-6.5)^2/6.5 = 1.923$$

有意とは言えない。しかし(b)は、

$$\beta; A : B = 1 : 5$$

$$(3-1)^2/3 + (3-5)^2/3 = 5.333$$

$$\chi^2 > \chi_{0.5}^2$$

となり、単語レベルの直後の言い直しは、後に発話した単語に関心があることが分かった。また統計的には有意と言えないまでも、B、すなわち言い直した単語の方に関心が高いことが言える。しかし、A、すなわち言い直された単語の方にも関心の高いものがあることがわかった。

#### 5 おわりに

本論文では、言い間違いの傾向を調べた。言い直しの時間的推移、言い間違いと発想力の関係、言い直した語と言い直された語との間でどちらに関心が高いかを分析した。分析の結果、言い直しの時間的推移に関しては、会議のはじめとおわりに言い直しが少なく、会議の中ほどに言い直しが多いことと、言い直しと発話量に関係が深いことが分かった。言い間違いと発想力の関係に関しては、言い間違えた語句はその会議にとって重要な語句が多く、発想力の高いものに関しては関心も高いことがわかった。言い直した語と言い直された語との間でどちらに関心が高いかに関しては、少なくとも単語レベルの直後の言い直しは、後に発話した単語に関心があることが分かった。

これらの結果をまとめてみると、関心が高いものは発想力が高い。発想力が高いものは発話量が多くなるであろう。発話量が多いところでは言い直しが多い。これらはそれぞれ逆が言えると思われるので、言い直しの場所を見つければ、関心の高いものが発見できることが推察される。

今後はこれらの知見を元にして、会議を支援する方法を考えていきたい。

#### 参考文献

- [1] Sigmund Freud, 安田徳太郎・安田一郎訳. 精神分析入門. 角川文庫, 1980.
- [2] J. D. Nasio, 姉齒一彦他訳. ラカン理論 5 つのレッスン. 三元者, 1995.

	A	B	(a)	(b)	V	Y	X	W	Z	T
1	メモ	電話の取り次ぎ	○	○	B	B	ⓑ	B	B	A
2	コンピュータ	マシン	○	○	A	A	ⓑ	A	A	A
3	机	ドア			Ⓐ	A	A	A	A	A
*4	出社	在席	○	○	A	ⓑ	A	A	A	A
5	ファックス	E-mail	○	○	ⓑ	B	B	B	B	B
*6	いるかないか	外出	○		A	ⓑ	A	B	B	B
7	スクリーニング	フィルタリング	○	○	A	ⓑ	B		B	A
*8	出社	在席	○		B	A	B	A	Ⓐ	A
9	離席	リフレッシュルーム			A	Ⓐ	A	A	A	A
10	翌日	朝	○		B	Ⓐ	A	B	B	B
*11	サービス出社	サービス残業			A	B	B	A	Ⓐ	A
*12	在席	離席	○		ⓑ	A	B	B	B	B
13	一人	依田さん			ⓑ	B	B	B	A	A
14	退社	出張			ⓑ	A	B	B	A	B
*15	リセット	リフレッシュ	○		ⓑ	B	A	A	B	A
*16	バージング	認識	○		A	A	A	A	Ⓐ	A
17	キーボード	キャラクタ	○		ⓑ	A	A	A	A	A
18	故意に	間違っ	○	○	Ⓐ	B	B	B	B	B
計			13(4/9) $\alpha$	6(1/5) $\beta$	8(2/6)	5(2/3)	2(0/2)	0	3(3/0)	

図 6: アンケートの結果