

あいづちの有無の影響を知るための主観データに基づく会議の質の 評価法と評価

土井 晃一*, 大森 晃**

*株式会社 富士通研究所, **東京理科大学
〒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, ohmori@ms.kagu.sut.ac.jp

和文抄録

ソフトウェアの要求獲得会議での、了解・情報の共有に対するあいづちの効果を知るための実験を行ない、その評価法を提案し、評価を行なう。評価対象として、会議参加者の会議に対する主観評価と会議参加者の発言における終助詞「ね」の頻度を選び、評価を行なう。

主観評価の評価には、単に平均値を取るよりも合理的な傾向一致度という新しい評価法を考案し、実際に評価を行なう。

その結果、あいづちを入れると、会議参加者は会議がどうも盛り上がり上がらないと感じるが、伝達に対する不安は感じていないことがわかった。つまり、あいづちの量をパラメータにして、会議の質と伝達に対する安心感がトレードオフの関係にあることがわかった。

Evaluation Method and Evaluation for Quality of Meeting Based on the Subjective Data in order to Know the Effect of Chime-in

Kouichi DOI*, Akira OHMORI**

*Fujitsu Laboratories,
9-3, Nakase 1-chome, Mihama-ku, Chiba-shi, Chiba 261, Japan
043(299)-3100
Akira OHMORI**
**Science University of Tokyo

Abstract

In order to know the effect of chime-in for understanding and sharing information in a software requirements capturing meeting, we experiment with meeting. We propose the method of evaluation, and evaluate the meeting. For the objective of evaluation, we choose the subjective evaluation of the attendants at the meeting and the frequency of the particle "ne" in their utterance.

For the evaluation of the subjective evaluation, we propose the more reasonable method, the method of coincidence degree of tendency, than the method of average. We evaluate with the method.

As the result of the experiment, attendants feel bad at the meeting and do not feel the anxiety of communication when the coordinator uses frequently uptake. We can find the trade-off relation between the quality of the meeting and the confidence of communication.

1 はじめに

ソフトウェアの要求獲得会議では、会議の性質上、そこで要求される内容に関して、要求獲得者と情報提供者との間の了解・情報の共有が不可欠である。しかし、例えば要求獲得者の了解・情報の共有の過程と程度は、情報提供者には見えにくい。逆も然りであり、第三者にいたってはなおさら見えにくい。もしも何らかの兆候によってそれを知ることができれば、さらなる了解・情報の共有を助長させる手だてを工夫できるかもしれない。これは、要求獲得会議の「質」を向上させるであろう。

了解・情報の共有の兆候は、ひとつには、情報獲得者のあいづちの量によって現れるであろう。そこで、あいづちの量を統制条件とした実験を行ない、情報提供者の会議に対する主観評価を行なった。

また、我々は要求獲得の試行会議から、了解・情報の共有の兆候は、(1)同意・確認・強調など対話者とコミュニケーションを密にしようという意図、を反映した言語現象として現れることを推察した。さらに聞き手側は(2)Turn-taking、話し手側は(3)助詞「ね」、という現象で観察できることが推察された。すなわち、助詞「ね」はコミュニケーションを密にし、コミュニケーションを密にしようという意図は助詞「ね」に現れるものと推察された。これは無統制な試行会議からの推察であり、この結果だけからでは、助詞「ね」の頻度とTurn-takingの増減とコミュニケーションを密にしようという意図との関係がわからない。そのような関係は統制された実験から得られるものと思われる。そこで、この試行会議とは別途行なっていた実験のコーパスをもとにして、最低限のTurn-takingであるあいづちと助詞「ね」との関連についても分析した。

2 実験条件

実験の条件は以下の通りである。一グループは学生二人（以下会議参加者と呼ぶ）と先生一人（以下コーディネータと呼ぶ）からなり、「～はどうかあるべきか」という題で、題はあらかじめ学生に決めさせ、話す内容についても考えてこさせるという条件で行なった。全部で八グループからなり、実験は10分づつ、五日連続して行なった。最初の二日は実験になれるためとし、データとしては使わなかった。会議参加者の発言はコーディネータに対する挙手で始めるように教示し、会議参加者同士のコミュニケーションは許さないことにした。グループは二つに分け、四グループは発言が途切れた時だけにコーディネータがあいづちを入れるグループ（以下I群と呼ぶ）とし、残りの四グループは可能な限りTRP（Transition Relevance Place）にあ

いづちを入れるグループ（以下II群と呼ぶ）とした。この条件は会議参加者には知らせなかった。この結果全部で24セッションのデータが得られた。ここで、一グループに対する10分間の実験が1セッションである。実験の様子はオーディオ・テープとビデオ・テープで録音・録画し、書き起こしをおこない、コーパスを作成した。会議参加者とコーディネータの配置は図1の通りである。

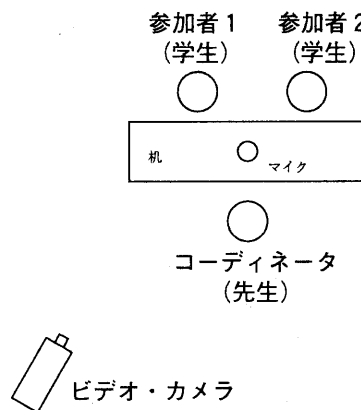


図1: 実験の見とり図

会議参加者がコーディネータとのコミュニケーションを密にし、会議参加者同士のコミュニケーションを疎にするために、会議参加者同士を向かい合わせずに、二人ともコーディネータの方向を向かせる図のような配置にした。

3 統計についての考察

本節では統計についての一般的な考え方について述べる。

一般に統計を使う際には細心の注意が必要である。ここでは、仮定（何を等確率とみるか、あるいははみたいか）と統計量の意味をきちんと考えなければならない。機械的に公式に当てはめても、ある程度の傾向を知ることにはできるかもしれないが、正確な結果は得られない。

例えば、一番良く使われる百分率でも生データが等確率で生起することを確認することが必要となる。また、平均は分布が左右対称のときのみ有効となる。もし、分布に偏りが見られる時にはラジアン（中央値）を用いるべきである。これは、図2を見てもわかるように、ラジアンの方がより代表値としてはふさわしい。つまり、平均値は僅かなノイズがかけ算で効いてくるため、特に、山の右側にある時、悪影響を及ぼす。さらに、標準偏差は正規分布に従う

ときにのみ用いるべきである。こういう使い方をすれば、いわゆる「統計の嘘」から逃れることができる。

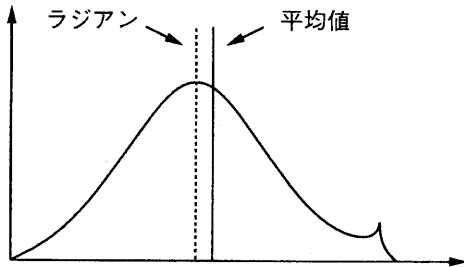


図 2: 平均値とラジアンの違い

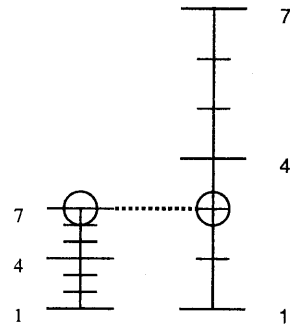


図 3: 個人のばらつき

4 傾向一致度を用いた評価法

本節では、前節の考え方を元に、傾向一致度を用いた、7段階尺度法の新しい評価法を提案する。

まず、会議の質の評価について考察してみる。会議の質を定量的に測るためには、例えば、会議参加者による会議の質の主観評価を7段階尺度法によって測定する方法が考えられる。本実験では会議参加者は二人であるので、以下、一会議当たりのデータは、例えば、(3,7)のように表記する。

互いに参加者が異なる複数の会議を評価しようとする二つの問題が生じる。一つは会議のばらつきをどう考えるかであり、もう一つは個人のばらつきをどう考えるかである。まず、会議のばらつきに対しては、会議の質をみる絶対的基準が存在するかという問題がある。我々は会議のばらつきに対して絶対的基準が存在するという立場を取る。次に個人のばらつきについて考えてみる。個人のばらつきが起きる原因として考えられるものは、

1. 個人によって基準点が異なる
2. 個人によって間隔が異なる
3. 何回か行なうと、実験の最初と最後で、ゆれが生じる
4. 会議の全体を見ていない

1に関して、個人の間隔が同じものとして図示すると図3のようになる。

図3は同じ会議に対して、一方が7と答え、他方が3と答えた場合である。図からもわかるようにこのデータを同一のものとして扱うのは抵抗がある。

また、4に関しては、他の会議参加者のことを忘れている、あるいは、考えていないことが予想される。

本研究では、会議のばらつきを見たいので、個人のばらつきは可能な限り押えたい。そこで個人のばらつきのあまりに大きなものは捨てることにした。具体的には、二人の傾向が一致しないものは捨てることにした。つまり、(7,3)のように一方がプラスの評価を出し、他方がマイナスの評価を出したものはデータとして採用しないことにした。

ここで、傾向一致度という概念を導入する。傾向一致度には、プラスの傾向一致度とマイナスの傾向一致度がある。プラスの傾向一致度とは、会議参加者の主観データが共に5以上のときだけ"1"を計量し、全セッション数で割ったものである。マイナスの傾向一致度とは、会議参加者の主観データが共に3以下の場合だけ"1"を計量し、全セッション数で割ったものである。以下、I群のプラスの傾向一致度をI+、マイナスの傾向一致度をI-、II群のプラスの傾向一致度をII+、II群のマイナスの傾向一致度をII-と表記する。

これは、単に平均値を取る場合と比べると、以下のような違いがある。次の表1のようなデータセットを考えてみる。

機械的に平均値を取った場合と、本方法を比べてみると、表2のようになる。

要するに、データセット1と3の(7,3)という悪いデータの影響をなくしたい。一つだけ(7,3)というデータがある場合、代表値に対してどのような影響があるかを示してみた。データセット1に対しては、平均値法では代表値を引き下げようように働く。一方、本評価法では、代表値に対して影響を与えない。次に、データセット2に対しては、平均値法では代表値を引き下げようように働く。一方本評価法では、代表値を引き上げようように働く。次に、データセット3に対しては、平均値法では代表値を引き上げようように

働く。一方本評価法では、代表値に影響を与えない。最後に、データセット4に対しては、平均値法でも、本評価法でも代表値を引き上げるように働く。

つまり、(7,3)という悪いデータがある場合は、代表値に影響を与えない(データセット1と3)。また、反対の傾向を持つ場合は、それなりの影響を代表値に与える(データセット4)。また、傾向に対して正しく強調されるという意味で妥当である(データセット2)。

このように、単に平均値を用いるより、傾向一致度を用いた方が、より妥当的な代表値を与えることができる。

DS_1	DS_2	DS_3	DS_4
(7,3)	(5,5)	(7,3)	(5,5)
(6,6)	(6,6)	(3,3)	(3,3)
(6,6)	(6,6)	(3,3)	(3,3)
(6,6)	(6,6)	(3,3)	(3,3)
.	.	.	.
.	.	.	.
.	.	.	.

表1: データセットの例

5 傾向一致度を用いた評価

本節では、前節までの議論に基づいた評価法での評価を行なう。

まず、会議の質を評価するための代用特性を以下の5つ選びだした。

評価項目1 面白かった

評価項目2 話しがはずんだ

評価項目3 内容にまとまりがあった

評価項目4 内容が充実していた

評価項目5 議論に疲れた

それぞれについて、会議参加者に7段階で評価してもらった。

評価結果を表にしてみると、表3ようになる。

データセット	何に?	どう効く?
DS_1	平均値法	-
	本評価法	0
DS_2	平均値法	-
	本評価法	+
DS_3	平均値法	+
	本評価法	0
DS_4	平均値法	+
	本評価法	+

表2: 悪いデータの寄与

	I	II	I+	II+	I-	II-
評価項目1	4 < 6	5 < 9	4 > 0			
評価項目2	3 > 2	5 < 8	7 > 3			
評価項目3	2 < 3	4 = 4	5 > 3			
評価項目4	1 < 2	2 < 7	6 > 2			
評価項目5	5 > 4	3 = 3	4 < 7			
合計	15 < 17	19 < 31	26 > 15			

表3: 評価項目ごとの一致度と傾向一致度

表中の数字は一致度(I,II)・傾向一致度(I+,II+,I-,II-)の値を表す。見易さのため、12で割っていない。一致度では有意さは見えないが、傾向一致度では、ほぼすべての評価項目で、プラス側でもマイナス側でも有意になる。評価項目5で結果が悪いのは、両義性を持つためと考えられる。つ

まり、「議論に疲れた」という認識は、会議に対して、良い場合と悪い場合の両方が考えられるからである。

この結果からいえることは、コーディネータがあいづちを入れると、会議参加者の会議に対する主観評価が下がることである。

本評価法を、会議参加者が三人以上の時にどう適用するかについて考察してみる。全員がプラス側、あるいはマイナス側に付けた時のみ評価の対象にすべきであろう。

6 「ね」の頻度

6.1 試行会議

ソフトウェアの要求獲得会議のコーパスをもとにして、助詞「ね」とその後接現象について分析した。総計 90 分からなるコーパスをもとにして助詞「ね」の直後に現れる言語現象を観察した。直後に現れる言語現象は対話者との関係で大きく三種類に分類できる。すなわち、Turn-taking がおきていないもの、Turn-taking がおきているもの、間(ま)が生じるものである。これにより、総計 479 個の助詞「ね」の内訳は次のようになった。Turn-taking がおきていないものは 279 個で、そのうち感動詞などが続くもの 43 個、倒置が生じているもの 43 個、繰り返しがおきているもの 11 個、言い間違いがおきているもの 1 個、その他こうした特徴的な現象を伴わないもの 181 個であった。Turn-taking がおきているものは 187 個、間があいているものは 4 個であった。

助詞「ね」が発せられると、話者のいずれかが続けて話すことが大勢を占め $((279 + 187)/479 = 99.2\%)$ 、とりわけ Turn-taking がおきている傾向にあった $(187/479 = 39.0\%)$ 。Turn-taking がおきていないところでも、感動詞などが続くもの・繰り返しがおきているものがあつた $(43 + 11/479 = 11.3\%)$ 。この現象は話者が Turn-taking を期待しているものと推察される。つまり、Turn-taking がおきる・期待されるものが、約半数を占めた $((43 + 11 + 187)/479 = 50.3\%)$ 。このことから、助詞「ね」が出現するところでは、Turn-taking が生じやすいものと判断できる。

Turn-taking が生じやすいということは、コミュニケーションを密にしようという意図の現れであると推察される。

一方、コミュニケーションを密にしようという意図の現れとして、Turn-taking がおきると推察される。Turn-taking の総数は 1322 個であった。このうち助詞「ね」の占める割合は 14.1% と大きく、特徴的であった。

以上のことから、助詞「ね」はコミュニケーションを密にし、コミュニケーションを密にしようという意図は助詞「ね」に現れ、したがってコミュニケーションを密にしよう

という意図を調べるためには、助詞「ね」を調べればよいことが推察される。しかし、これは無統制な試行会議からの結果であり、これだけでは助詞「ね」の頻度と Turn-taking の増減とコミュニケーションを密にしようという意図との関係がわからない。

そのような関係は統制された実験から得られるものと思われる。ただし、実験で統制できる条件は Turn-taking の頻度しかない。また複雑な条件を取り除くために、Turn-taking の種類としてあいづちだけを取り上げ、その頻度を統制した実験が望ましい。

ここでは、そうした条件を満たすような別途行っていた実験を利用して、助詞「ね」の頻度を観察することにする。

6.2 実験結果

コーパスをもとにして、各参加者の 3 回の実験を通した総発話に現れる助詞「ね」の数を、各参加者の 3 回の実験を通した総発話時間(沈黙の時間を含めない)で割ったものを計量した。総データ数は 16 であり、あいづちのあるなしのそれぞれの群に対して 8 個のデータがある。ただし、単位は回/分である。分布の概要は図 4 のようになる。データの分布が対称ではないので、代表値としてメジアン、散布度として四分位偏差を用いた。

	あいづちなし	あいづちあり	全体
メジアン	1.13	0.69	0.90
第一四分位	0.57	0.16	0.40
第三四分位	1.91	1.97	2.02
四分位偏差	0.668	0.903	

図 4: 分布の概要

個人差はあるものの、参加者個人を「ね」の頻度が少ないもの(0 以上 0.40 未満)、普通のもの(0.40 以上 2.00 未満)、多いもの(2.00 以上)の三グループに分けた。この分類は、あいづちあり・なし双方をあわせたデータセットから求めた第一四分位と第三四分位を参考にして設定したものである。この分類を階級としてヒストグラムを描くと、図 5 のようになる。

横軸が階級、縦軸が人数を表す。あいづちありのグループがあいづちなしのグループに比べて、「ね」の頻度が少ない傾向にある様子が見てとれる。言い換えると、あいづちなしのグループがあいづちありのグループに比べて、「ね」の頻度が多い傾向がある。

この分析結果から、コーディネータがあいづちを打たないと、参加者は助詞「ね」を多用する傾向にあると判断

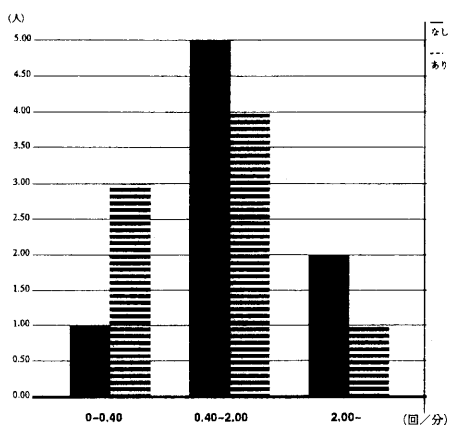


図 5: あいづちのあるなしに対する助詞「ね」の頻度の差

される。

助詞「ね」の頻度は発話者の伝達に対する自信を表すものと推察される。コーディネータがあいづちを打たないことにより、参加者は伝達に対する不安が生じ、その結果、助詞「ね」を多用するものと推察される。

6.3 評価結果

ソフトウェア要求獲得の無統制な試行会議から、(1) 助詞「ね」と(2) Turn-taking と(3) 同意・確認・強調など対話者とコミュニケーションを密にしようという意図、との間に以下のような関係を推察した。すなわち、(a) 助詞「ね」が出現しているところでは Turn-taking がおきる、(b) Turn-taking がおきるということは、同意・確認・強調など対話者とコミュニケーションを密にしようという意図の現れである。また一方、(c) コミュニケーションを密にしようという意図の現れとして、Turn-taking がおき、(d) そこでは助詞「ね」の占める割合が大きく、特徴的であった。しかし、これだけでは(A) 助詞「ね」の頻度と(B) Turn-taking の増減と(C) コミュニケーションを密にしようという意図との関係がわからない。そこで、別途行っていた統制された実験からこの三者間の関係を調べてみた。その結果、コーディネータがあいづちを打たないと、参加者は助詞「ね」を多用する傾向にあると判断された。

助詞「ね」の頻度は発話者の伝達に対する自信を表すものと推察される。コーディネータがあいづちを打たないことにより、参加者には伝達に対する不安が生じ、その結

果、助詞「ね」を多用するものと推察される。

会議参加者に伝達に対する不安を生じさせないためには、コーディネータには適度にあいづちを打つことが望まれる。さらに伝達に対する不安を生じさせないことによって、会議参加者は了解・情報の共有に専念できるものと推察される。また助詞「ね」の頻度が大きい時は、その発話者は聞き手が了解しているという自信をあまり持てないものと推察される。

7 会議での使い方

本節では、この実験結果の会議への応用を考察する。会議が円滑に進んでいない時、すなわち会議参加者が

1. 伝達に対する不安を感じている気配がするとき
2. 会議がどうも盛り上がらないと感じている気配がするとき

コーディネータが反省する材料になる。つまり、1のときはあいづちを増やしてやることによって対応する。2のときはその逆に、あいづちを減らしてやることによって対応することになる。

この結果は、心理学実験の結果が常にそうであるように、必ず成り立つわけではない。必ず逆に感じる人がいることは確かである。つまり、あいづちを打たれると、じっくり聞いていると思って、より熱心に話すこともありうる。また、逆に、あいづちを打つとうるさく感じることもありうる。その時には場合に応じた運用が必要になる。

8 おわりに

ソフトウェアの要求獲得会議での、了解・情報の共有に対するあいづちの効果を知るための実験を行ない、その評価法を提案し、評価を行なった。評価対象として、

1. 会議参加者の会議に対する主観評価
2. 会議参加者の発言における終助詞「ね」の頻度

を選び、評価を行なった。

その結果、あいづちを入れると、会議参加者は会議がどうも盛り上がらないと感じるが、伝達に対する不安は感じていないことがわかった。つまり、あいづちの量をパラメータにして、会議の質と伝達に対する安心感がトレードオフの関係にあることがわかった。

今後は、会議の発言中に現れる概念の数を計測することによって、会議の質に対する、あいづちの効果を調べていきたい。