

## 遠隔会議における発話の衝突と 精神的ストレスの関係

玉木秀和<sup>†</sup> 東野豪<sup>†</sup> 小林稔<sup>†</sup> 井原雅行<sup>†</sup>

Web 会議に代表される小規模な遠隔会議システムは利用や導入の手軽さがあるが、映像、音声の品質に制限があるため、大規模で高精細な遠隔会議システムに比べ、円滑な会話がし辛い。小規模な遠隔会議システムを用いて、司会者のいない創造会議を実施すると、参加者同士の発話が衝突してしまうことが多い。我々は発話の衝突を低減することを目標に研究を進めているが、発話衝突の許容量、基準は明らかになっていない。そこで本稿では、音声会議において、音声遅延量を変化させることで発話衝突確率を変動させる実験を実施し、それが精神的ストレスと、会議の生産物であるアイデアの数に及ぼす影響を調査した。

### Relationship between Speech Contention and Psychological Stress in Distributed Conferences

Hidekazu Tamaki<sup>†</sup> Suguru Higashino<sup>†</sup>  
Minoru Kobayashi<sup>†</sup> and Masayuki Ihara<sup>†</sup>

In Web conferences, we sometimes cannot recognize when other participants begin speaking. So two or more participants often start to speak at the same time. This depresses the participants' motivation and wastes time. We are trying to overcome this problem but we have not known a limit that a speech contention rate should fall inside. So we did an experiment to investigate the relationship between speech contention and psychological stress.

### 1. はじめに

「ちょっと集まって議論しましょう」「まずは集まってざっくばらんにアイデアを出し合しましょう」このようなシーンはオフィスの中でよく見かけるが、これらは Face-to-Face だからこそ不自由なく実現できる、というのが現状である。

遠隔会議システムの研究開発が重ねられ、技術が発展してきた。しかし、高臨場感を生む遠隔会議システムは、性能を求める程大規模なものとなり、部屋単位で作りこまなくてはならないものも存在する[1]。これに対して使用場所の制約が少ない Web 会議システムは、インターネットに接続できれば、ノート PC から使用することも可能である[2]が、前述のシステムと比較すると映像や音声の質は低く、円滑な会話を行うことは難しい[3]。

冒頭に挙げたような、複数人が集まって行う、司会者のいないフリーディスカッションや創造会議中の会話では、話者交替が頻繁に起こる。このような会議を、Web 会議に代表される小規模の会議システムで実施する場合、二人以上の会議参加者が同時に話し始める「発話衝突」が頻繁に起こり、問題になると我々は考えている。我々はこれまでに Web 会議において発話衝突を低減する手法を提案してきたが[4]、司会者のいない創造会議を実施する際、発話衝突確率をどの程度に抑えるべきかという基準は、見出されていない。

これまで遠隔会議システムにおいて音声遅延をどの程度に収めるべきかという観点での実験はなされているが、1対1の会話を対象にしているものや、実験のタスクとして、しりとりや数字読み上げゲームなどを行っていた[5]。タスクの内容により話者交替の頻度は変化し、それにともない発話衝突確率や、発話衝突に起因する精神的ストレスも異なることが予想される。

そこで我々は、複数拠点間で行う、司会者のいない創造会議において、発話衝突確率の許容量を求めるために実験を行った。音声会議において、音声遅延を変動させて、発話衝突確率を変化させた。そして精神的ストレスと、会議中に創出されたアイデア数を測定したのでその結果を報告する。

### 2. 遠隔会議システム

これまでに規模や特徴が様々な遠隔会議システムが作られ、よりコミュニケーションしやすくなるよう、工夫がなされてきた。

複数人で行う会議は Picture in Picture といわれる、升目状に区切った領域に会議参加者の映像を並べて表示するシステムが広く普及した。参加者間で、誰が誰を見てい

<sup>†</sup>日本電信電話株式会社 NTT サイバーソリューション研究所

NTT Cyber Solutions Laboratories Nippon Telegraph and Telephone Corporation

るか、という情報を表現可能にしたものとして Hydra がある[6]. Hydra は参加者 1 人につき 1 台ずつのディスプレイ、カメラ、スピーカ、マイクを用意し、これが参加者の顔、目、口、耳の役割を担う。視線と対人距離を表現し、誰が誰に注目しているかを認識することができる。

映像圧縮技術の発展や、ネットワーク速度の向上を背景に、複数拠点間を、あたかも同じ部屋にいるかのように繋ぐ試みもでてきている[1]. しかし、高臨場を目指すほどシステムの規模は大きくなり、据え置き型、もしくは部屋単位で設計しなくてはならないため、利用場所が限定される。これに対し小さい規模のシステムで利用可能な Web 会議は、インターネットに接続できる環境であれば、ノート PC からでも参加できるため、場所の融通が効くが、映像や音声の質は下がり、会話しづらくなる[3].

Web 会議などで伝えることのできない非言語情報を伝達するために、アバタを利用した例もある。石井らはアバタを介した遠隔コミュニケーションシステムにおいて、次に発話することを促す視線を表現する手法を提案した[7].

### 3. 話者交替

対面会議の話者交替については、これまでに詳細な研究が行われている。まず基本的な用語を整理する[8].

#### ターン(turn)

ある話者が順番を取って発話をするための発話権。発話権を得て話された 1 人の話者の一続きの発話もターンと呼ぶ。

#### ターン構成単位(turn constructional unit; TCU)

ターンを構成する最小の単位であり、文、節、句、単語が含まれる。ターンは 1 つ以上の TCU から構成される。話し手がこの単位を用いて話すなら、ターンを取得することが可能になる。ただし、「相槌」などはこの単位に相当せず、話し手は「相槌」によってターンを取得できない。

#### 話者交替適格場(transition-relevance place; TRP)

各 TCU の末尾に存在する話者交替にふさわしい場所。

#### 無標な重複(overlap)

TRP 付近に生じる音声的重複。この重複は terminal overlap と呼ばれ、聞き手が完結可能点を予測したうえで話出した結果生じるものとされるもので、ターン配分規則に従っている。

#### 妨害(interruption)

TRP 以前の TCU に対してなされる重複。相手の発話権への侵害となる。ターン配分規則に違反している。

小磯らは、TRP において、現話者が発話を継続する、もしくは他の参加者が発話を

開始する場合を話者交替の成功とし、複数の参加者が発話を開始する場合に起こる衝突と、誰も発話せずに沈黙することを話者交替の失敗とした[9].

#### 発話衝突(simultaneous talk / speech contention)

本稿では、TRP において、1 つ以上の TCU が重複すること、と定義する。

Sacks は話者交替のルールを次のように定めた[10].

1. 最初の TCU (ターン構成単位: turn constructional unit) における最初の TRP (話者交替適格場: transition-relevance place) において,
  - a) 「現話者による次話者選択」というテクニックを含んだ TCU が構成されているなら、選択されたものが次のターンを取る権利と義務を得る。他の三余社はその権利と義務を得ず、移行はここで行われる。
  - b) 「現話者による次話者選択」というテクニックを含まない TCU が構成されているなら、次の話者についての自己選択が可能になるが、これは義務的ではない。最初に話始めた者がターンを取る権利を得て、移行はここで行われる。
  - c) 「現話者による次話者選択」というテクニックを含まない TCU が構成されており、他の者が自己選択も行わないなら、現話者が話し続けることが可能であるが、義務的ではない。
2. 最初の TRP においてルール 1a) や 1b) が作動せず、1c) に従って現話者が話し続けているならば、以降の TRP において、1a) ~ c) が話者交替が生じるまで循環的に再適用される。

### 4. 遠隔会議における話者交替の研究

話者交替の観点から遠隔会議システムの遅延量を評価した研究がある。

鎧沢らは、2 拠点間の映像コミュニケーションを行う際に、早い話者交替を必要とする会話は往復遅延 250msec までとし、特別早い話者交替を必要としなければ往復 500msec まで音声遅延が生じても問題ないとした[11]. ITU-T 勧告 G.114 では、往復 300msec を超えると話に支障をきたすと報告している[12].

タスクや会議の性質によって求められるシステムの要件は変わると考えられる。周知会議や、司会者のいる統制された会議ではある程度遅延のある会議システムでも許容されるが、司会者のいない創造会議などは、話者交替が頻繁に起こる[13]ため、ある程度の質が求められると推測する。さらに 2 つの拠点間の会話であれば、会話のパスが単純であるが、3 つ以上の拠点間で行う場合は会話のパスが複雑になり、問題が顕著になる。司会者のいない創造会議が Web 会議などで快適に実施可能になれば、便利であるが、そのためにどの程度の要件が求められるかは明らかにされていない。そこで我々は、この指針を立てるために実験を実施した。

## 5. 実験

複数拠点間で、司会者のいない創造会議をし、遅延量と発話の衝突確率、精神的ストレス、生産されたアイデアの数の関係を調べた。

### 5.1 実験システム

図1のように、音声会議環境を構築した。発話の衝突や精神的ストレスには、様々な要素が複雑に影響することが予想される。そのため、今回の実験では音声のみを伝達しあう音声会議を選択した。各被験者のマイクから入力された音声はディレイマシンに送られた。ディレイマシンは音声遅延を0から1秒まで調整することができた。ディレイマシンから出力された音声は、入力元以外の被験者のスピーカへ送られた。つまり、それぞれの被験者は自分以外の被験者の音声をスピーカを通して聞くことができた。被験者はそれぞれ防音室の中に入った。

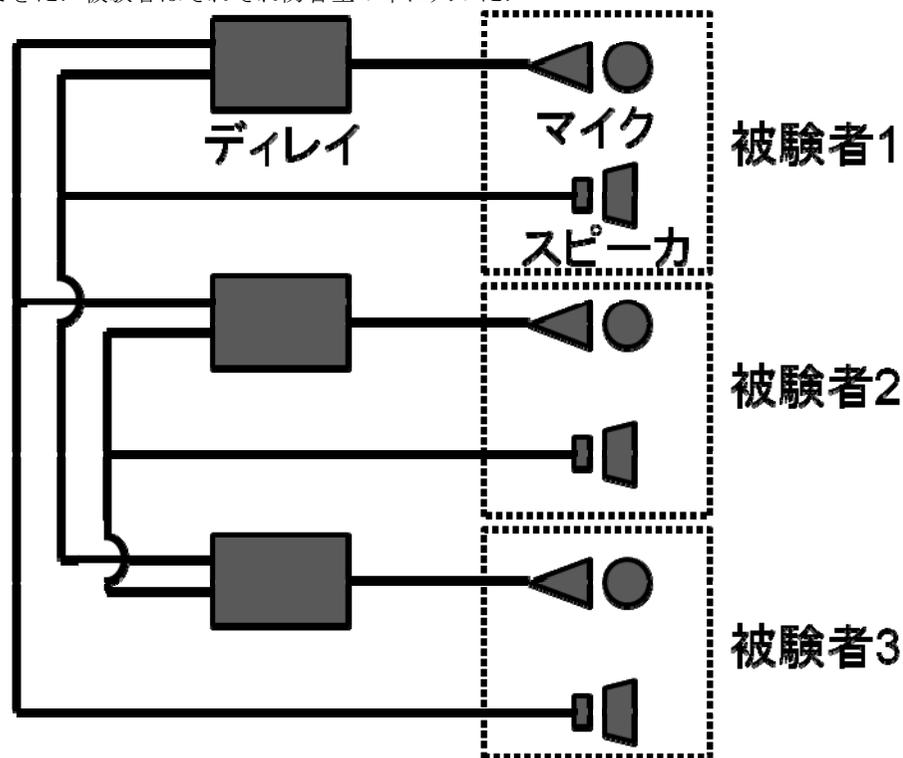


図1 音声遅延会議システム

### 5.2 手順

3人ひと組で、決められた時間内に、できるだけ多くのアイデアを考え出す会議を行わせた。

往復音声遅延を0msec, 200msec, 400msec, 600msec, 800msec, 1000msecの6種類用意し、それぞれ順序を入れ替えて実施した。1つの条件ごとに7分間実施し、時間内にできるだけ多くのアイデアを出させた。3人のうち1人書記役を決めさせ、会議に参加しながらグループで考えたアイデアを記録させた。会議のテーマは、「携帯電話の新機能を考える」とし、6つの条件の間で継続させた。それぞれの条件で会議を実施した後、質問紙によるアンケートに答えさせた。アンケート項目は下記の通りである。5段階で、項目ごとに当てはまると思うほど高い点数をつけさせた。

- ・ 音声遅延をストレスに感じる
- ・ 誰が話しているか分からないことをストレスに感じる
- ・ 誰が話し始めそうかわからないことをストレスに感じる
- ・ いつ話し始めていいかわからないことをストレスに感じる
- ・ 話し始めが他の人とぶつかることをストレスに感じる
- ・ 他の参加者が自分の話を聞いているかわからない
- ・ 沈黙が多い
- ・ 話に割り込めないことをストレスに感じる
- ・ 盛り上がらない
- ・ 相手の存在感が薄い

会議中の被験者の映像、音声をビデオに記録し、分析した。

### 5.3 結果と考察

#### 5.3.1 発話衝突確率

ビデオを解析し、1つのグループの条件ごとの発話回数、発話衝突回数を計測した。なおこのグループは、遅延量を0msecから1000msecまで順に増加させて実験を行っていた。

発話回数を図2に示す。遅延量が増えるほど発話回数が減少しているが、同じテーマでアイデアを出す会議を行ったため、時間が経過するほど発話回数が減った可能性も考えられる。これに関しては今後他のグループでの発話回数も計測し、傾向を比較する必要がある。

発話した回数を母数とし、そのうち発話衝突が起こった確率を発話衝突確率として図3に示す。1グループの集計結果のみであるが、音声遅延量が増加するにつれ、発話衝突確率も増加する傾向がある可能性がある。

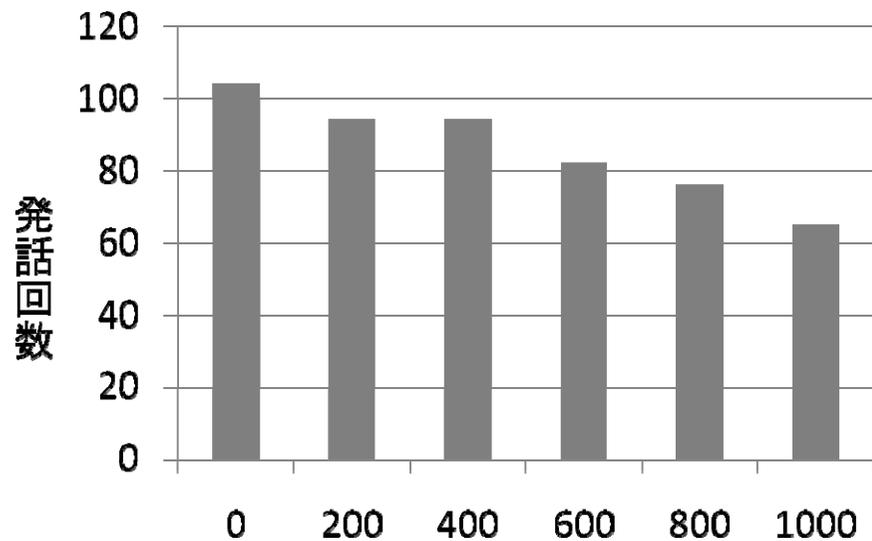


図2 発話回数

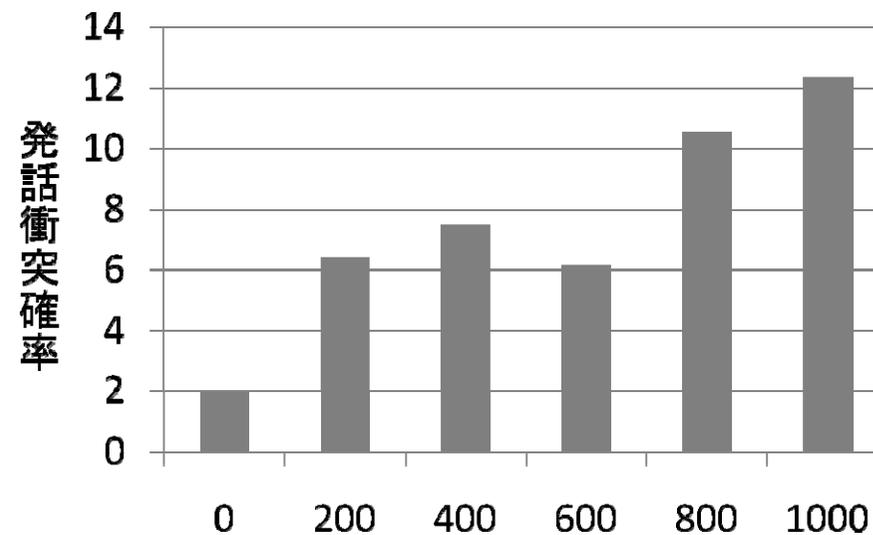


図3 発話衝突確率

### 5.3.2 主観評価の結果

質問紙によるアンケートの結果を表1にまとめる。Friedmanの検定を用い、音声遅延量の変化がアンケート結果に影響あるという結果が得られたのは「音声遅延をストレスを感じる」「話し始めが他の人とぶつかることをストレスを感じる」の2項目である。

表1 Friedmanの検定結果

アンケート項目	Friedmanの検定結果
音声遅延をストレスを感じる	有意水準 1%で有意差あり
誰が話しているか分からないことをストレスを感じる	有意差なし
誰が話し始めそうかわからないことをストレスに感じる	有意差なし
いつ話し初めていいかわからないことをストレスに感じる	有意差なし
話し始めが他の人とぶつかることをストレスを感じる	有意水準 5%で有意差あり
他の参加者が自分の話を聞いているかわからない	有意差なし
沈黙が多い	有意差なし
話に割り込めないことをストレスを感じる	有意差なし
盛り上がらない	有意差なし
相手の存在感が薄い	有意差なし

この2項目について多重検定を行い、条件ごとの比較を行った結果を表2, 3に示す。「音声遅延をストレスに感じる」というアンケート項目では、往復音声遅延量が800msec周辺に閾値がある可能性がある。

表2 多重比較「音声遅延をストレスに感じる」

	0	200	400	600	800	1000
0	-	なし	なし	なし	****	****
200	-	-	なし	なし	*	****
400	-	-	-	なし	なし	****
600	-	-	-	-	なし	**
800	-	-	-	-	-	なし
1000	-	-	-	-	-	-

\*: 有意水準 10%, \*\*: 有意水準 5%, \*\*\*: 有意水準 1%, \*\*\*\*: 有意水準 0.5%未満

多重比較「話し始めが他の人とぶつかることをストレスに感じる」

	0	200	400	600	800	1000
0	-	なし	なし	なし	なし	なし
200	-	-	なし	なし	なし	**
400	-	-	-	なし	なし	なし
600	-	-	-	-	なし	なし
800	-	-	-	-	-	なし
1000	-	-	-	-	-	-

\*: 有意水準 10%, \*\*: 有意水準 5%, \*\*\*: 有意水準 1%, \*\*\*\*: 有意水準 0.5%未満

「話し始めが他の人とぶつかることをストレスに感じる」というアンケート項目では、往復音声遅延量が1000msec周辺に閾値がある可能性がある。前述の通り、発話衝突確率は、音声遅延量1000msecのとき、発話衝突確率12%であった。このため、今回の分析まででは、往復音声遅延量を1000msecにし、発話衝突確率が12%を超えたときに、会議参加者は発話衝突をストレスに感じる可能性がある、ということがいえる。今回の実験では音声遅延を変化させることにより発話衝突確率が変動し、その結果のストレスを主観評価として測定した。そのため発話衝突の起こる原因を音声遅延以外にすれば参加者の感じるストレスも変化すると考えられる。しかし発話衝突確率12%というのは1つの指針となりうる。例えば会議参加者を増やすだけでも発話衝

突確率は増加することが考えられる。そうすればより少ない音声遅延量であっても会議参加者はストレスを感じるかもしれない。条件を変えて実験することで、汎用的に使用できる指針として、司会者のいない創造会議において許容される発話衝突確率を求めていく。

### 5.3.3 アイデア数の比較

条件ごとに、記録されたアイデアの数を集計し、7分間で創出されたアイデアの数を表4にまとめる。この種類の生産性は変わらなかった。およそ100秒に1つ程度のアイデアが創出されたことが分かるが、このペースの生産活動は、数百ミリ秒単位の遅延量の違いによっては影響を受けなかったと考えられる。

表3 アイデア創出数

遅延量(msec)	アイデア数
0	4.4
200	4.4
400	4.3
600	4.9
800	4.2
1000	4.1

## 6. おわりに

複数拠点間で司会者のいない遠隔会議をする際に許容される発話衝突確率の基準を見出すため、音声遅延が可変な会議システムで実験を行った。3人一組のグループで会議を行い、音声遅延を変化させることにより発話衝突確率を変動させた場合、発話衝突確率が12%を超えると会議参加者が精神的ストレスを感じるという結果が得られた。また、今回のようなタスクにおいては、音声遅延による発話衝突の度合いにより、アイデア創出数に差が出ないことも明らかになった。

今後は、音声遅延以外の原因による発話衝突においても実験を行い、司会者のいない創造会議において許容される発話衝突確率として、汎用的な基準を求めていく。

## 参考文献

- 1) <http://www.cisco.com/web/JP/solution/video/index.html> (2011年2月現在)
- 2) <http://www.meetingplaza.com/index-j.html> (2011年2月現在)

- 3) 徳勲, 友保康成, 渋谷雄, 田村博: テレビ会議技術の課題と利用法についての考察, 8th Symposium on Human Interface, pp.207-212, 1992
- 4) 玉木秀和, 中茂睦裕, 東野豪, 小林稔: 人のコミュニケーションリズムに着目した Web 会議円滑化手法, IEICE Technical Report MVE2009, pp.101-106
- 5) 鎧沢勇, 滝川啓, 大久保榮, 渡辺義郎: 衛星通信を利用した画像会議におけるエコー及び伝搬遅延の影響, 電子通信学会論文誌(B), Vol.J64-B, No.11, pp.1281-1288, 1981
- 6) <http://www.cisco.com/web/JP/solution/video/index.html> (2011年2月現在)
- 7) 石井亮, 宮島俊光, 藤田欣也: アバタ音声チャットシステムにおける会話促進のための注視制御, ヒューマンインタフェース学会論文誌 Vol.10, No.1, 2008.
- 8) 榎本美香: 日本語における聞き手の話者移行適格場の認知メカニズム, ひつじ書房, 2009.
- 9) 小磯花絵, 伝康晴: 円滑な話者交替はいかにして成立するか: 一会話コーパスの分析にもとづく考察一, 認知科学, 7, 93-106, 2000.
- 10) Harvey Sacks, Emanuel A. Schegloff, Gail Jefferson: A Simplest Systematics for the Organization of Turn-Taking for Conversation, Language, 50, 4 Pt 1, 696-735, Dec 74
- 11) 鎧沢勇, 滝川啓, 大久保榮, 渡辺義郎: 衛星通信を利用した画像会議におけるエコー及び伝搬遅延の影響, 電子通信学会論文誌(B), Vol.J64-B, No.11, pp.1281-1288, 1981
- 12) ITU-T Rec. G.114, "One-way transmission time", 2003.
- 13) 井上智雄, 岡田謙一, 松下温: テレビ会議における映像表現の利用とその影響, 情報処理学会論文誌, Vol.40, No.10, pp.3752-3761, 1999.