

## 解 説

## イグノーベル賞受賞研究 SpeechJammer

## SpeechJammer, Awarded Ig Nobel Prize

栗原 一貴  
Kazutaka Kurihara

産業技術総合研究所  
National Institute of Advanced Industrial Science and Technology.  
k-kurihara@aist.go.jp, <http://unryu.org>

塚田 浩二  
Koji Tsukada

科学技術振興機構 さきがけ  
Precursory Research for Embryonic Science and Technology, Japan Science and Technology Agency.  
tsuka@mobiqutous.com, <http://mobiqutous.com/>

**Keywords:** delayed auditory feedback, artificial speech disturbance, 2012 Ig Nobel prize.

## 1. はじめに

SpeechJammer という研究は、数奇な運命をたどり今日に至っている。我々はその渦中において、現代における研究の在り方について大いに考えさせられた。論文とは何だろう。学会とは何だろう。研究と社会の接点とは何だろう。これらは社会の情報化と価値観の多様化が進む今、改めて考えるべきテーマではないだろうか。

本稿では、通常の論文の形式をあえて避け、論文の辺縁にあった出来事や我々の思惑などを盛り込み、この稀有な体験を伝えることを試みる。まず、研究の経緯を述べる。これはいわゆる論文の「研究背景」に相当する整理されたものではなく、研究の着想について時系列的に述べたものである。次に、SpeechJammer の紹介として、唯一本研究が学術的に受け入れられた WISS'10<sup>\*1</sup> の論文 [栗原 10] を再編し記述する。そして最後に、その後イグノーベル賞の受賞までに何が起こったのか、そして何を我々は学んだのかについて総括する。

## 2. 研究の着想

その発端は、2010年の夏に遡る。当時第一著者はプレゼンテーショントレーニングシステム「プレゼン先生」を開発していた [Kurihara 07]。これはプレゼンテーションをしているユーザをカメラで撮影、録音し、それを音声認識技術と画像処理技術を用いて分析することで、基本的なプレゼンテーションの良し悪しをユーザにフィードバックするというものだった (図1)。指標として取り扱うのは、話速度、言いよどみの有無、声の抑揚、顔の向き (聴衆のほうを向いているか) などである。あら



図1 プレゼン先生システム

かじめ設定したしきい値を超える挙動が確認された場合は、「話速度を落としてください」、「言いよどんでいます」、「前を向いて話しましょう」などの警告をディスプレイに表示するという設計だった。

「プレゼン先生」を試してみると、確かに自分のプレゼンテーションに「ダメ出し」をしてくれて便利なのだが、警告を受けたときに「では、どうすれば問題を改善できるのか」についてヒントが得られない点が不満であった。それに対し、二つの改善策を考えた。

- 「どう問題を改善すればよいか」について専門家のノウハウをアドバイスとして組み込む。
- 何らかの技術を用いて、そもそも問題が起こらないようにユーザを制御する。

「話速度が速すぎる」という問題でこれらを考えてみると、前者は例えば「時間は十分あります。心にゆとりをもって、まるで絵本を子どもに読み聞かせるように話してみてください」のような話速度低下が期待できそうな具体的アドバイスを提示することであり、後者は、人間の「しゃべる」という行為に直接影響を与えるような何らかの認知的法則を応用し、話速度をシステムが制御できるようにすることに相当する。

これらの改善策の実装方法を模索している中、休日

\*1 日本ソフトウェア科学会第18回インタラクティブシステムとソフトウェアに関するワークショップ。査読付き国内会議。

<http://www.wiss.org/>

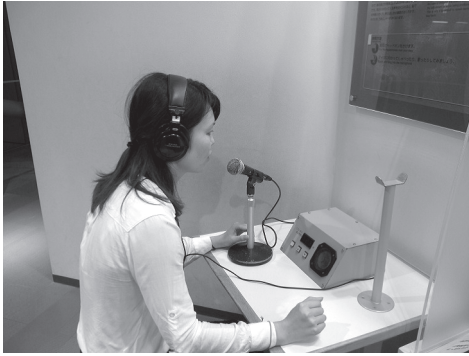


図2 日本科学未来館における聴覚遅延フィードバックの展示

に訪れた日本科学未来館にその答えを見つけた。聴覚と認知についての小さい展示ブースだったと記憶しているが、単純にマイクとヘッドフォンが備え付けてあり、自分の声を数百 ms 遅らせて自分に聞かせる聴覚遅延フィードバックという仕組みについての、非常に古典的な科学館展示の一つがあった(図2)\*2。この装置により、人間は正常にしゃべることが難しくなってしまうという。著者自身が体験してみると、実際におもしろいようにしゃべれなくなってしまった。そこでアイデアが閃いた。この展示は単に「人間の聴覚にはこのようなおもしろい性質があります」と教えてくれるにすぎないが、この現象を用いれば、人間の話速度をシステム(外部)からある程度制御できるのではないかと。単純には、しきい値を超える話速度が検出された場合にブレーキをかける装置として、そして、発展的には微細な時間単位でこの現象を用いることで、連続的に話速度の制御が可能になるのではないだろうか? 調べてみると聴覚遅延フィードバックは吃音症の改善に用いられた事例はあるものの、このように健聴者の発話・コミュニケーション制御に積極的に用いられた事例はないようだった。

ここからは、我々研究者であれば日常的に行っている推論過程に入る。ある問題【(1)より良いプレゼンテーショントレーニング】に対して、ある解決法【聴覚遅延フィードバック】を得た。これは抽象化すると、「人々に対し外部からその発話を制御もしくは妨害する」ソリューションであると考えられる。このソリューションが有効であるような、別の問題はあるだろうか。思索の末、二つ思いついた。一つは(2)会議や議論の場で、参加者の公平な発言を制御する議長の仕事を支援すること。もう一つは、(3)もし装置が小型化できたならば、日常生活のうえで会話に対する無作法をしている人々を取り締まる携帯機器とすることである。この三つの問題を取り扱う、「聴覚遅延フィードバックを用いた発話制御の応用」という仮研究タイトルと、「発話を適切に制御することにより人々のコミュニケーションをより豊かにする」という研究のミッションが決まった。

\*2 なお、この展示は現在終了している。

そして、さらに第一著者はあることを思い出した。以前参加した国内学会 WISS'09 において、第二著者が「指向性スピーカ」を用いた自作の遠隔音声通信機器を遊撃的にデモしていた風景である。これは、拡声器のような形状をしているが操作者の声が拡散せず直線的に伝達され、遠隔地の特定の人物にのみ聞き取れる機能をもっていた。この遠隔通信機器を拡張すれば、携帯機器として遠隔地から特定の話者の発話のみを制御することが可能になり、難しいと思われた(3)が実装できるのではないだろうか。すぐに第二著者に連絡を取り、快諾を得て共同研究が始まった。

かくして約1か月の試作・執筆期間を経て、書き上げたのが WISS'10 で発表した論文「SpeechJammer: 聴覚遅延フィードバックを利用した発話阻害の応用システム[栗原 10]」である。実際にインフォーマルな実験をしてみると個人差と慣れの影響が強く、実用化に至るまでにはまだまだ詳細な研究が必要であるとわかり、本来到達したい連続的な話速度制御にも至れなかったが、核となる研究コンセプトとプロトタイプングの経緯を中心に盛り込んだ。次章はその論文の再編版である。実際の研究経緯と、論文中的研究背景の構成の違いなどにも注目して読んでいただきたい。

### 3. SpeechJammer (WISS'10 再編版)

#### 3.1 はじめに

音声および文字を用いた他者との対話は、言語によるコミュニケーションの基本的な手段である。このうち、音声による対話は、身体のみで言語を表現でき、1対多の情報伝達が容易であるなどの優れた特徴があるため、人類が文字を発明して以降も絶えることなく用いられて今日に至っている。

しかし対話による紛争の平和的解決が重要視されている現代社会において、音声のもつ負の特徴によってその精神が脅かされたり、また健全な社会生活に害悪をなしているケースが後を絶たない。

我々はそのような音声の負の特徴のうち、不可避、占有、重畳無効化、調整困難の四つを以下のように定義し、これに注目する。

- 不可避: 音声が一方向的に話し手から開始・継続でき、聞き手が制御できない。
- 占有: 音声対話が通常、同時にただ一人だけの発話を許容する。
- 重畳無効化: いっときに二人以上の発話があった場合、すべての発話が無効化し理解困難になる。
- 調整困難: 話し手が意図的に話速度などの音声プロパティを聞き手にふさわしく調整するのにスキルを要する。

これらの特徴により、代表的に以下の3種類の問題が指摘される。

(1) 議論における「言った者勝ち」問題

対話による紛争の解決には、公正な議論が不可欠である。音声による対話は不可避、占有、重畳無効化の特徴があるため、適切に発言権の交代ルールを規定し、それを遵守しなければならない。しかし、議論の内容よりも自身の存在を（マスコミなどを通じて）アピールすることを目的とする勢力が存在する場合などに、いたずらに自身の発言を長引かせたり、他者の発言の合間に割り込んだりといった行為が現れるようになる。また、他者の発言中に野次を飛ばすなどして、発言を無効化するなどの戦略が取られる。これは不可避、占有、および重畳無効化の濫用であり、結果的に「言った者勝ち」という風潮がまかり通っている。

(2) 会話が不適切である場における会話を止めにくい問題

公的な場の一部では過度の会話が不適切であると規定されていることがある。例えば図書館や電車内での騒がしい会話はすべきではないとされている場合が多い。これは音声のもつ不可避な特徴から、聞くことの拒否が難しい以上、他者が不快に思うレベルの騒音は発生させないようにすべきだという社会的総意から生まれた規定であろう。

しかしこのような場において、すでに発声している会話を止めることは難しい。なぜならば会話が不適切であることを会話者に説得し止めさせるためには、説得者も会話に参加せざるを得ない可能性が非常に高いからである。これは説得者自身も不適切行為に身を染めることであり、心理的負担が大きい。また、意を決して説得に乗り出したとしても、(1)のように公正な議論を行う同意が得られなければ占有、重畳無効化の濫用により対話が無秩序化し、会話を止めるという目的からますます遠ざかる可能性を秘めている。

結果的に現状では積極的に会話禁止に乗り出すのではなく、ヘッドフォンを着用するなどして音声の不可避を断ち切り、会話と共存を図るなどの消極的解決手段に頼らざるを得ない様子がしばしば観察される。

(3) プレゼンテーションにスキルが必要な問題

現代社会においてプレゼンテーションは広く普及が進んでいる1対多のコミュニケーション手段である。ここにおいても音声は主要な情報伝達手段として用いられているが、その運用は普段行っている1対1もしくは1対少数の音声コミュニケーションとは異なる。例えば通常の1対少数の音声コミュニケーションでは、しばしば聞き手の理解が難しかった点をその場その場で問いただすことでインタラクティブに解消できるが、通常のプレゼンテーションではそのような機会は極めて少ない。したがって話者自身が内容、および話速度などの音声プロパティについてある程度「最適なパフォーマンス」を想定、実施し、それを維持することが必要である。プレゼンテーションの初心者がこれを行うのは困難であり（調整困難）、習得には訓練が必要である。

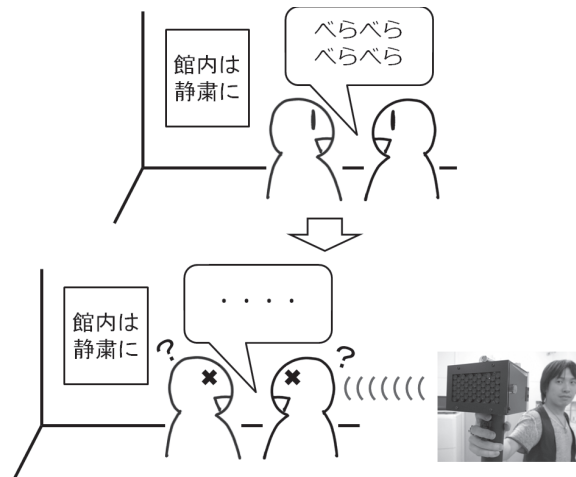


図3 SpeechJammer システム概要

栗原ら [Kurihara 07] はそのような様式の自習のためのシステムを開発したが、音声の問題を視覚で指摘するなど、指摘対象と指摘のモダリティが異なっており、強制力が弱かった。

我々は、外部からの働きかけにより話者の発言を制御する研究を行っている。その第一段階として本論文では、人間の聴覚特性の応用である聴覚遅延フィードバックという手法を用いて遠隔地から強制的に発言を阻害するシステムを構築する（図3）。本手法は肉体的苦痛を伴うことなく発言を阻害することができ、また発言をやめればただちにその認知的な影響が消失し、また話者のみに作用するためそれ以外の周囲の人達には無害であるという優れた性質も持っている。発言者の外部から強制的な発言の阻害が可能になることで、例示した問題の根源である音声の負の特徴が改善される可能性がある。すなわち、不可避性は緩和され、その結果適切に話者交替を行うことで占有性を制御でき、かつ発言の重畳がなくなることで重畳無効化性も問題となくなる。またプレゼンテーショントレーニングシステムと組み合わせることで、調整困難性を改善できる。

本論文は、まず関連研究を列挙する。その後実装したシステムについて詳細を示す。そして今後の評価実験計画について検討する。

3.2 関連研究

§1 聴覚遅延フィードバック

我々は発声時、単に発声という運動出力を行うだけでなく、実際に耳から聞こえた発声音（聴覚フィードバック）を脳内で活用していると考えられている[庄司]。ここで、人工的に自己の発声音を聞かせるタイミングを遅延させると、脳内の処理系に何らかの作用が働き、以降の発声に支障をきたす。これが聴覚遅延フィードバック(DAF)による発言阻害である。DAFは吃音症(どもり)と関係が深い。健常者ではDAFにより人工的などもり(発言阻害)が誘発される形となるが[庄司]、逆に吃音

症患者は DAF により吃音が矯正されることが知られており [Stuart 97], 医療用の DAF 機器も市販されている [SmallTalk].

我々は DAF を活用し, 話者の意思とは関係なく外部から話者の発話を阻害するデバイスを開発する. システム構成によっては通常の DAF 用システムとは異なり, マイクとスピーカが話者から離れた場所に設置される点特徴的である.

§2 議論支援

本研究はデバイスによる音声コミュニケーションの制御を扱ったものである. これは広義には議論の支援研究とも考えられる. 議論の支援に関する関連研究として, 議論内容を記録し再利用性を研究した長尾らによる [Nagao 04] や, 議論記録から有意義なコミュニケーションパターンをボトムアップに構築することを目指した中田らによる [中田 10] などがあげられる.

またテキストチャットにおいてコミュニケーションにあるルール・制約を適用したり, 議論に関する何らかの情報を可視化することで議論を活性化させたり, 集団意思決定などのある共通の目的を達成することを支援する研究も盛んである [Leshed 10, 村田 04, 西田 09]. 本研究で開発するシステムは, 音声による議論に「強制的な発話の妨害措置」という新たな強い制約を課すことが可能である. この制約を活用することにより, 議論のコミュニケーションパターンに有意義な変化をもたらすことができ, これら議論に関する研究を支援・拡張することが展望される.

3.3 システム設計

我々は話者の発話を阻害するシステム, 「SpeechJammer」を設計するうえで, その適用局面を検討する必要がある. なぜならば, 発話の場にマイクとスピーカがどのように置かれているかによって, ハードウェア設計に違いが生じ, また音波の伝達時間に違いにより, DAF のパラメータが変化するからである. 我々の想定する本システムの適用局面は,

- (1) 会議室における発言権の制御
- (2) 携帯発話阻害銃
- (3) 発話者自身による発表練習支援

の三つであり, これらは 3.1 節であげた三つの問題点にそれぞれ対応している. 図 4, 図 5, 図 6 に, それぞれの局面におけるシステムの構成を示す. 図中, 実線は電気信号の伝達を, 波紋は空気中の音波の伝達を表す.

表 1 は, これらをマイクとスピーカの設置状況をもとに分類したものである. 会議室における発言権の制御には, それぞれの発話者の近傍に設置されたマイクと, 会場に備え付けられたスピーカをインフラとして活用できる (図 4). この場合, 空気中の音波の伝達はスピーカ・聴衆間の一方のみである. 携帯発話阻害銃は, 聴衆の手許にマイクとスピーカを備えたデバイスを用意し, 生

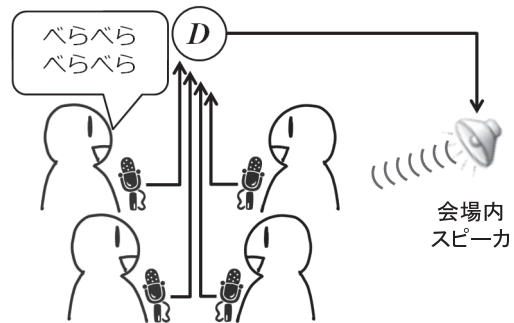


図 4 会議室における発言権の制御

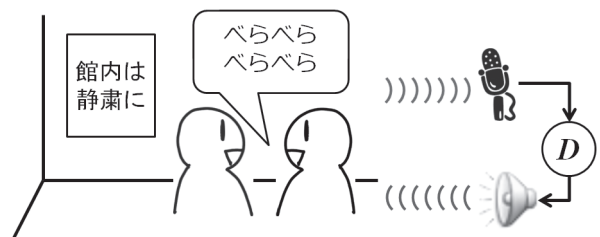


図 5 携帯発話阻害銃



図 6 発話者自身による発表練習支援

表 1 発話阻害の使用が想定される局面のマイクおよびスピーカ設置状況による分類

	話者近傍 スピーカ	遠隔 スピーカ
話者近傍 マイク	発話者自身による 発表練習支援	会議室における 発言権の制御
遠隔 マイク	(なし)	携帯発話阻害銃

活上で遭遇する不適切な発話に対し発話阻害を行う (図 5). この場合, 空気中の音波の伝達は話者・聴衆間の往復分である. 発話者自身による発表練習支援では, マイクとスピーカを発話者近傍に設置できるヘッドセットを活用することができる. この場合, 空気中の音波の伝達はほぼ無視できる (図 6). このセッティングのもとでプレゼン先生システム [Kurihara 07] と組み合わせることにより, 話速度がしきい値を超えて早口になっているとシステムが判断したとき, 発話阻害機能を発動させる.

本論文では, これらのうち最も複雑な構造をもつ携帯発話阻害銃タイプの SpeechJammer のプロトタイプを

開発し、検証する。このデバイスさえあれば、ほかの二つの用途（発表練習支援、会議の発言権制御）の検証にも活用することができる汎用性をもっているからである。

§ 1 遅延の計算

システムを実現するためには、話者とデバイスの間の距離に応じた音声の伝達時間を考慮し音声遅延を発生させる必要がある。ここで図7のように話者とデバイスとの距離を  $x$  [m] とし、気温を  $t$  [°C]、DAF で発話を阻害するために必要な遅延を  $D_{daf}$  [s] とすると、空気中 (1 atm) の音の伝達速度を加味したうえでシステムに生じさせる遅延  $D$  [s] は

$$D = D_{daf} - 2x/v \tag{1}$$

で表せる。ただし、音速  $v$  [m/s] は

$$v = 331.5 + 0.61t \tag{2}$$

である。

上式によると、例えば  $D_{daf} = 0.2$  s のように話者に与える遅延を固定した場合、気温 20°C において  $D \geq 0$  となるのは  $x \leq 34.37$  m である。これは気温 20°C において本手法による最大射程が約 34 m であることを示しており、一般的な室内環境であれば実用上十分であるといえる。

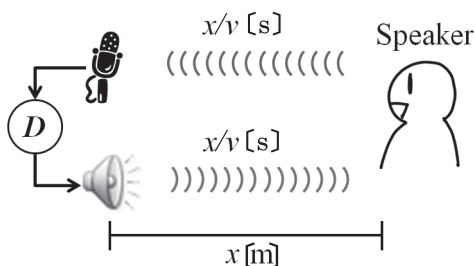


図7 必要な遅れ  $D$  の計算

この最大射程内の任意の標的に対し常に発話阻害を引き起こすためには、標的とデバイス間の距離  $x$  が固定され事前に知られている場合は固定値  $D$  を計算すればよく、 $x$  が未知、もしくは変動する場合は距離センサにより逐次的に  $x$  を測定し  $D$  を求めればよい。

一方、システムを簡素化し固定の  $D$  により発話阻害を発生させる場合の考え方は以下のとおりである。DAF による発話阻害が  $0.1 \leq D_{daf}$  の音声遅延で発生するとすると、式 (1) より  $D = 0.1$  s とすれば、 $0 \leq x$  となり、固定の  $D$  によって任意の距離での発話阻害が可能となる。ただし、 $D_{daf}$  による発話阻害の程度の違いや  $D_{daf}$  の有効な上限値は我々の調査した範囲では明らかになっていないため、実験により与える必要がある。また実際はマイクとスピーカの性能により  $x$  の上限（最大射程）は規定される。

3.4 実装

我々は、SpeechJammer の2種類のプロトタイプを実装した。

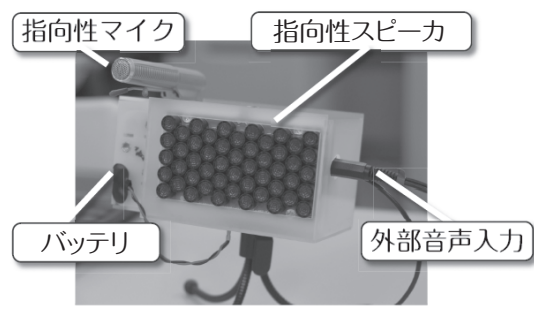


図8 SpeechJammer プロトタイプ1の概要

プロトタイプ1はSpeechJammerの基礎的な機能を備えた最初のプロトタイプである(図8)。指向性マイク (Sony ECM-Z60) と指向性スピーカ (トライステートパラメトリックスピーカキット) を一体化し、アクリル製のケースに取り付けた。ケースには三脚用のねじを切っており、一般的な三脚に固定して角度などを調整できる。指向性マイクの出力はホスト PC のマイク入力に、指向性スピーカの外部入力にホスト PC のヘッドフォン出力に接続される。ホスト PC 上では、Pure Data (以下、PD) で記述された音響処理プログラムが動作しており、指向性マイクの入力に任意の遅延をかけたうえで、指向性スピーカに出力することができる。なお、AC ラインに起因するノイズを防ぐため、電源として単三電池×8本を組み込んでいる。

上記プロトタイプ1での基礎的な動作を確認したうえで、スタンドアロン型のプロトタイプ2を実装した。指向性マイク (Sony ECM-CZ10) と指向性スピーカ (トライステートパラメトリックスピーカキット) を中心に、レーザポインタ、距離センサ、各種スイッチ、およびメイン基板を一体化して、独自に製作したアクリル製ケースに組み込んだ(図9、図10)。レーザポインタは、SpeechJammer の大まかな照準として利用し、背面のスイッチで任意にオンオフできる。距離センサは、ターゲットとの距離を計測し、遅延時間を調整するために利用する。

メイン基板には、マイコン (Microchip PIC18F-452)、デジタルディレイ IC (Rohm BU9262AFS)、プリアンプ、メインアンプ、および周辺回路が配置されている。

指向性マイクの入力は、プリアンプを通して、メイン基板のデジタルディレイ IC に接続される。デジタルディレイ IC は、シリアルインタフェースを介して、遅延時間を 9.2 ~ 192 ms まで変更できる。遅延時間は、背面のロータリースイッチを用いて8段階で選択する\*3。デジタルディレイ IC の出力は、メインアンプを通して、指向性スピーカに接続される。プリアンプとメインアンプは通常はミュートされており、トリガスイッチが

\*3 ロータリースイッチの設定次第で、距離センサを用いてターゲットとの距離を計測し、遅延時間を自動的に調整したり、任意の範囲で周期的に遅延時間を変更する仕様も準備している。

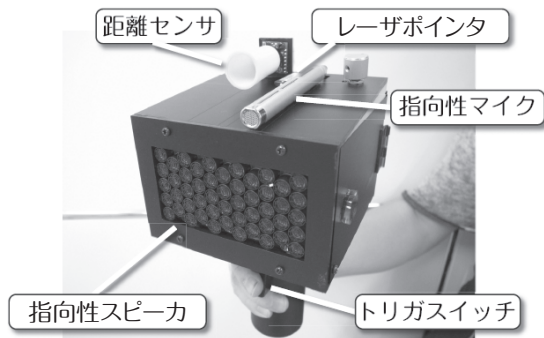


図9 SpeechJammer プロトタイプ2の概要, 前面

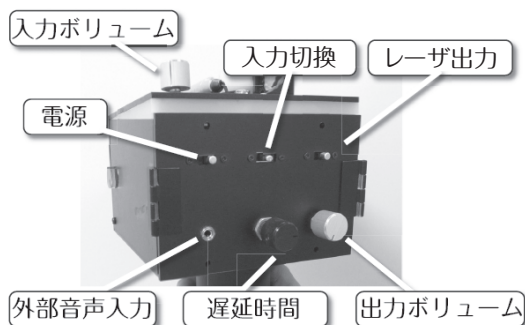


図10 SpeechJammer プロトタイプ2の概要, 背面

押された場合だけオーディオ入出力が行われる仕様としている。また、入出力音量はケース上面と背面のつまみで個別に制御できる。マイコンは、デジタルディレイIC、トリガスイッチ、ロータリースイッチ、距離センサなどを制御するために利用する。

このように、ユーザはSpeechJammerをターゲットに向け、拳銃のようにトリガスイッチを引くだけで、容易に発話阻害機能を発動することができる。

なお、SpeechJammerのプロトタイプは2012年以降も進化を続けており、スマートフォンベースの3号機、普及をにらんだ安価でシンプルな構成の4号機、携帯性を犠牲に出力(≒効果)を大きく向上させた5号機などを実装している(図11)。今後、第二著者のWebサイト上で回路や筐体などの細かい仕様を公開するとともに

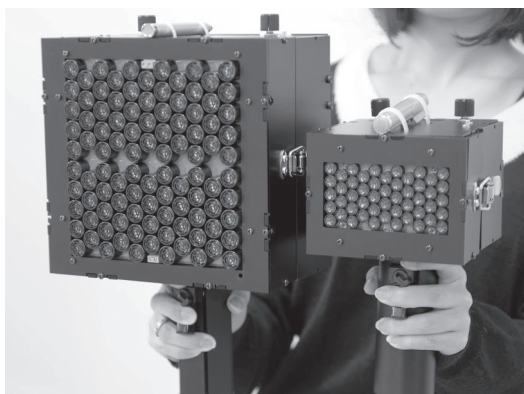


図11 SpeechJammerの新しいプロトタイプの一例。  
左：出力重視の5号機、右：普及重視の4号機

に、キット化や玩具化などの展開を見据えている。

### 3.5 まとめ

我々は音声によるコミュニケーションを円滑にするために、不可避、占有、重畳無効化、調整困難といった音声の負の特徴を緩和するシステム、SpeechJammerのプロトタイプを実装した。SpeechJammerは聴覚遅延フィードバックの仕組みを用いて、肉体的な苦痛を伴うことなく話者の発話を阻害することができる。今後はさまざまなパラメータを変化させ、パラメータと発話阻害効果の関係を調査し、応用方法を検討していく。

## 4. イグノーベル賞への軌跡

前述のWISSという国内会議は、荒削りでも将来を切り拓くようなシステム開発研究を積極的に取り入れるという先進的な試みを実践している。例えば、WISS'10では論文フォーマットに「未来ビジョン」というコラム欄を用意し、研究の延長上にある未来像を研究者自身が想像し、長期的展望を語らせている。しかし、世界に目を向けると、こうした傾向はまだ一般的なものではない。我々の分野の国際学会は採録率が低いものが多く<sup>\*4</sup>、研究や論文としての充実度・完成度に重きを置く苛烈な競争が繰り返されている。その価値観からするとpremature(未熟)と言わざるを得ないSpeechJammerは国際学会において不再録を繰り返し、1年を浪費してしまった。我々は疲弊し、国際展開を諦めた。かくなるうへは研究の『成仏』を願い墓に埋めるような気持ちで、「せめてこの研究を我々が行ったことを世界に示そう」と合意し、英語論文[Kurihara 12]と英語デモビデオ[Youtube]のインターネットでの公開に踏み切った。

すると、驚くべきことが起こった。突如として海外の大手を含む無数のメディアからの取材が殺到した。Youtubeに投稿したビデオの再生数は1週間で50万回を超え、その後100万回を突破した。これは、Youtubeの週間動画ランキングで2位になるほどの異常事態である。厳密な拡散の経緯は定かではないが、聞くところによると、とある海外の有名なブロガーが公開された論文を発掘し「これで我々の言論の自由は終わった」のような扇情的なトーンで紹介したことが一因だという。かくして、賛否両論が激しく展開された。中東の軍事産業関係者や言論統制下国家のジャーナリストからの問合せなど、騒動は我々の想像を超えた範囲に広がっていった。そのうねりはWired.jpなどを通じて逆輸入され、日本でも話題になった。

このような取材に対応するため、我々は以下のような、ともすると滑稽なメディア対応テンプレートを準備せざ

\*4 トップカンファレンス(CHI, UIST, Ubicompなど)でおおむね10~20%。

るを得ない状況に陥った。

#### 【悪用される懸念に対する考え】

- 本研究について、仮に遠隔地からの第三者による発話阻害の効果が確実に得られる状況が将来達成できたとき、悪用されることにより言論の自由が人々から奪われるのでは、というご指摘をいただいている。それに対して我々は以下のように考えている。
- 原点として、我々は言論の自由は人々に平等に与えられるべきものであり、「声の大きい人が勝つ」と俗に言われるような、特定の人物だけに言論が占有される不公平を払拭したいと考え、本研究をスタートさせた。
- しかし「どういう言論が不公平であるか」という判断が、SpeechJammer 使用者の倫理観に依存している点は現状の課題である。ともすると本技術が乱用され、みだりに人々の言論が封殺されてしまう事態も起こりかねないことは認める。
- しかし、本技術は刀や銃などによる戦闘と比べて、「お互いに使用し合っても破滅せず、話し合いによる解決の余地を常に残せる」という性質をもっている点が重要である。もしも誰かが不本意にも SpeechJammer をあなたに使用したうえで会話を占有し始めたら、あなたもその人に SpeechJammer を使用し、冷静な話し合いが行えるまで待てばよいのである。
- 一方で、組織や国家が無人の SpeechJammer を至る所に配置し言論を封殺するようなディストピア像を想像する人もいる。この場合は上記のような話し合いは期待できない。
- しかし、本技術の効果は「耳栓」をすれば容易に回避できる程度のものであるのでそのような心配はない。
- 本技術の応用が適切なのは、「自分自身のプレゼンテーショントレーニング」、「参加者全員が会話のルールについて納得している会議や公共スペース」などの、条件次第では発話阻害されることを当事者が了承しており、耳栓をするような回避手段をとることが想定しにくい、非難されるような日常の局面である。このテンプレートの作成においては、前述の WISS'10 における「未来ビジョン」の経験が役に立った。我々は SpeechJammer の論文執筆にあたり、究極的な研究の将来像として「世界平和と兵器」という壮大なテーマで考察を深めていたのだ。当時のメディアの過熱具合は、それほどの壮大な議論が求められるものであった。

メディア対応にも疲弊していたある日、一通のメールが届いた。端的に言えば、「イグノーベル賞を受賞できますけど、しますか?」という内容である。我々はイタズラメールの類ではないかとあやしんだ。しかし、やり取りしてみるとどうやら本物のようである。イグノーベル賞は皮肉を込めた授賞もあるため、怒る人もいる。だ

からやけに低姿勢の物言いになるらしい。謹んでお受けした。

意外に知られておらず、我々も実際知らなかったことだが、イグノーベル賞は自己推薦も含む推薦によってエントリーされる。我々は自己推薦をしなかったので、どこかの国の誰かが推薦してくれたということだ。その人物については謎のままである。しかし、それまでの地球規模での各種メディアでの騒動が推薦の背景にあることは想像に難くない。

イグノーベル賞の授賞式については、過去に2回もイグノーベル賞を受賞されている偉大な中垣氏による[中垣 09]が詳しいので、そちらを参照されたい。荘厳な雰囲気のあるハーバード大学サンダースシアターにおいて、ノーベル賞受賞者を招いての笑いの絶えない催しであった。賞金も交通費も出なかったが、あらゆる意味でプラスストレスな思い出となった。

## 5. おわりに

最後に、我々が SpeechJammer の数奇な軌跡から得た教訓について以下の観点からまとめたい。

- (1) 研究発表手段の多様化
- (2) 学会の価値の変化
- (3) ネット時代の「わかりやすさ」
- (4) 「議論を呼ぶ」ことの意義

まず(1)研究発表手段の多様化について述べる。論文を書き学会で発表することだけが研究の発表手段ではなくなっている。インターネットを使えば、個人であっても、研究者でなくとも自分の創造した価値を世の中に問えるようになった。SpeechJammer が、国際会議での連戦連敗の後に、インターネット上で爆発的に評価されたことは一つ示唆的であったと思う。すなわち、今までは研究者にとってほぼ唯一の価値観であった学会というのが、世界に林立する価値観の一つになったともいえるだろう。もちろん、その中で歴史と伝統と実績をもつ学会の価値はなくなることはないだろうが、我々は最適な手段を「選べる」立場に立ったのかもしれない。

その中で、(2)学会の価値も変わるだろう。研究成果をより多数の人々に伝達するという意味では、学会よりもインターネットのほうが圧倒的に優れている。今や、論文を Web で公開することは普通であるし、学会の様子をインターネット中継することも広まりつつある。この時代に、研究へのお墨付き、権威付けを与えるという機能以外に、学会で重要なことは何だろうか。我々は、「専門家同士の緩やかな交流の場」がやはり重要な機能として残るように思える。SpeechJammer プロジェクトは、もともと学会で遊撃的に第二著者が技術をデモンストレーションしていたことを第一著者が思い出したことに端を発する。このような専門性と共通の関心をもつ仲間との出会いはなかなかインターネットでは得られなさそうだ。

次に、インターネットでの情報発信が無視できない強力なツールとなったことを前提にしたとき、不特定多数の人々に研究成果を伝えるうえで重要なことは何だろうか。SpeechJammerの事例から学べることは、脱言語と脱コンテキストによる(3)「わかりやすさ」の重要性である。SpeechJammerは、英語で発信し、また偶然にも特定の文化圏や趣味に依存せず、幅広く全人類が話題にできるテーマを扱っていた。さらに、その実現方式も一般人にも容易に説明することができた。インターネットを使うということは、全人類を相手にするということであり、こうした「脱言語」、「脱コンテキスト」のわかりやすさが求められてくる。

そのうえで、(4)議論を呼ぶことも重要だ。SpeechJammerの拡散の際には先述のとおり、「これで我々の言論の自由は終わった」のような扇情的で否定的な意見が発言力のある情報発信源から飛び出し、「おもしろい」、「なんて便利だ」、「仕組みが痛快だ」などの肯定意見とともに両論が激しく展開されたい。むしろ我々はねらってそうしたわけではないが、近年では「炎上マーケティング」という言葉もあるくらいである。一筋縄ではいかない、人々の情感に多方面から訴えかける要素というのは、無視できない影響を及ぼすのだろう。

「人々を笑わせ、そして考えさせる」。これはイグノーベル賞の受賞要件である。一見わかりやすいが、よくよく考えると深く、熱い研究者の思いがある。若輩ながら我々のこれまでの研究哲学を肯定されたようで、受賞を誇りに思う。インターネットがインフラ化した21世紀における研究活動は20世紀のそれとは変わらざるを得ないだろう。我々は今回のできごとを胸に刻み、この混迷の時代で次の一步を踏み出していく。本稿を通じて読者各位になにかインスピレーションを差し上げることができたなら幸いである。

#### 謝辞

本研究の一部は、科学技術振興機構さきがけプログラムの支援を受けた。展示写真と研究インスピレーションを提供いただいた日本科学未来館に感謝する。

#### ◇ 参考文献 ◇

- [Kurihara 07] Kurihara, K., Goto, M., Ogata, J., Matsusaka, Y. and Igarashi, T.: Presentation sense: A presentation training system using speech and image processing, *Proc. ICMI'07*, pp. 358-365 (2007)
- [栗原 10] 栗原一貴, 塚田浩二: SpeechJammer: 聴覚遅延フィードバックを利用した発話阻害の応用システム, *WISS2010* 論文集, pp. 77-82 (2010)
- [Kurihara 12] Kurihara, K. and Tsukada, K.: SpeechJammer: A system utilizing artificial speech disturbance with delayed auditory feedback, <http://arxiv.org/abs/1202.6106> (2012)
- [Leshed 10] Leshed, G., Cosley, D., Hancock, J. T. and Gay, G.: Visualizing language use in team conversations: Designing through theory, experiments, and iterations, *CHI EA'10*, pp. 4567-4582 (2010)
- [村田 04] 村田和義, 川口 修, 倉本 到, 渋谷 雄, 辻野嘉宏: 遅延時間を用いたチャット参加者の発言制御, *ヒューマンインタフェース学会論文誌*, Vol. 6, No. 4, pp. 411-422 (2004)
- [Nagao 04] Nagao, K., Kaji, K., Yamamoto, D. and Tomobe, H.: Discussion mining: Annotation-based knowledge discovery from real world activities, *Proc. PCM'04 Part 1*, pp. 522-531 (2004)
- [中垣 09] 中垣俊之: イグノーベル賞授賞式顛末記, *Unicorn Journal*, Vol. 69, pp. 31-33 (2009)
- [中田 10] 中田篤志, 角 康之, 西田豊明: 非言語情報の出現パターンによる会話状況の特徴抽出, *インタラクション'10* 予稿集 (2010)
- [西田 09] 西田健志, 山崎敬一: 集団意思決定への参加感を高めるトーナメント型議論, *グループウェアとネットワークサービスワークショップ2009* 論文集, Vol. 2009, No. 8, pp. 123-128 (2009)
- [SmallTalk] SmallTalk, <http://www.casafuturetech.com/Cata-log/smalltalk.shtml>
- [Stuart 97] Stuart, A., Kalinowski, J. and Rastatter, M. P.: Effect of monaural and binaural altered auditory feedback, *J. Acoust. Soc. Am.*, Vol. 101, No. 6 (1997)
- [庄司] 庄司壽一, 伊福部達: 吃音軽減のための音声フィードバック装置について, <http://mito.cool.ne.jp/stutstudent/ronbun.html>
- [Youtube] Youtube, <https://www.youtube.com/watch?v=USDI3wnTZg>

2013年2月●日 受理

#### 著者紹介



栗原 一貴 (正会員)

2007年東京大学大学院情報理工学系研究科コンピュータ科学専攻博士課程修了。PhD。日本学術振興会特別研究員(DC2)を経て、同年、産業技術総合研究所に入所。現在メディアインタラクション研究グループ研究員。Human-Computer Interaction全般に興味をもつ。



塚田 浩二 (正会員)

2005年慶應義塾大学大学院政策・メディア研究科博士課程修了。同年、産業技術総合研究所研究員。2008年4月より、お茶の水女子大学特任助教。2012年4月より、科学技術振興機構さきがけ研究員(専任)。生活環境に適したユーザインタフェースの研究・開発に従事。博士(政策・メディア)。