JAIST Repository
https://dspace.jaist.ac.jp/

Title	超低音速空間による「より社会的なコミュニケーション」の実現に向けて
Author(s)	馬場，裕；小林，智也；小倉，加奈代；西本，一志
Citation	インタラクション2012論文集（情報処理学会シンポジウムシリーズ），2012(3)：433-438
Issue Date	2012-03-15
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/10649
Rights	社団法人　情報処理学会，馬場 裕，小林 智也，小倉 加奈代，西本 一志，インタラクション2012論文集（情報処理学会シンポジウムシリーズ），2012(3)，2012，433-438．　ここに掲載した著作物の利用に関する注意：本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 Notice for the use of this material: The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.
Description	

Japan Advanced Institute of Science and Technology

超低音速空間による「より社会的なコミュニケーション」の実現に向けて

馬場　裕[†]　小林　智也[†]
小倉　加奈代[†]　西本　一志[†]

日常行われている会話には，様々な問題点がある．たとえば失言や重要な話の聞き逃しは，人間関係を悪くしてしまう危険性がある．これらは，空気中の音速で声が人間の何百倍の速さで移動する事が要因と考えられる．音速を遅くすることで，音声に追いついたり音声の重畳や参照をすることが可能になる．そこで，超低音速空間 CreepingVoice を構築し，会話の問題解決を試みている．本稿では，超低音速空間 CreepingVoice を実装するための基礎的な実験の結果について報告する．

Toward “further social communications” by a super slow sonic speed space

Yutaka Baba,[†] Tomoya Kobayashi,[†] Kanayo Ogura[†]
and Kazushi Nishimoto[†]

Everyday conversations have various problems. For example. a slip of the tongue spoils human relations. This problem is attributable to the speed of sound. Sound is too fast for us people to catch up. If the sound speed is much slower, we can catch up with our voice, modify it and refer it. We have been constructing a super slow sound space “CreepingVoice” to solve the problems. This paper reports results of elemental experiments for implementing CreepingVoice.

1. はじめに

人類は，秒速約 340m の音速の世界で暮らしている．これは，超人的な身体能力を持つ陸上競技選手である，ジャマイカのウサイン・ボルトが 100m を走る平均速度のおよそ 32.5 倍の速さであり，一般人の日常的行為における身体動作速度をはるかに上回っている．このように音速は人類にとって非常に高速であるため，通常の対面対話のようなたかだか数メートルの範囲においては，音声は実質的に遅延無く瞬時に伝わっているように感じられる．これにより発話のすばやいやりとりが可能となっていることが，対面対話の快適さの一因となっていると思われる．国際電話や衛星中継放送のように，音声伝送に顕著な遅延が生じた場合，我々は違和感を覚え，不便を感じる．

我々は生まれながらに音速 340m/s の世界で生活しているので，このような現実の世界に特段の不便や問題を感じておらず，無条件に高速な音速を容認している．日常生活において人々は，音に追いついたり追い越したりする必要性は無いと考えている．しかしながら，音速が一般人の身体動作速度レベルの低速度であった場合には，以下のようなメリットがあると考えられる．

音声重畳による発話の取消：対面対話において，うっかり失言をしてしまった場合，通常音速の世界では取り消すができない．しかし，もし自分が発した音声に追いつけるならば，その上にマスキングのためのノイズ等を重畳して，失言を物理的に無かったことにすることが可能となる．

発言の遅延聴取による会話へのスムースな途中参加：ある会話の輪に途中から入り込むことは容易ではない．いきなり不用意な発言をすると，会話の流れを壊してしまう．しかし音速が十分に遅い場合，その会話の輪から離れた位置から徐々に接近することによって，過去から現在に向かって早送りするように対話内容を聴取でき，最終的に会話の輪の位置に到着した時点で「今の発言」に追いつくことができるので，会話の流れを把握可能となる．

従来，音速が遅いことを問題としてとらえ，音声を

† 北陸先端科学技術大学院大学
Japan Advanced Institute of Science Technology

電気通信に乗せることによって音速を擬似的により高速にする取り組みと，そこでのコミュニケーションのあり様に関する研究が多数なされてきた．これに対し本研究では，音速が速過ぎることを問題としてとらえ，音速をより低速な「人間的な速さ」とした場合，その世界ではどのようなコミュニケーションがなされるのか，そこにはどのようなメリット・デメリットがあるのかを検討する．

以下，2 章では関連研究について概観する．3 章では，構築した超低音速空間 CreepingVoice の構成について説明する．4 章では CreepingVoice を用いて，対面での対話を不自由なく行える音速の下限を調査する．5 章では，CreepingVoice を用いた音声重畳の実験について述べる．6 章では，CreepingVoice を用いた会話への途中参加実験について述べる．7 章はまとめである．

2. 関 連 研 究

これまで，会話の遅延が人に与える影響についての研究多く行われている．特に，発話に対しての応答の遅延が発話者に大きな影響を与えていることが知られている．

音声のコミュニケーションにおいて，聞き手の反応が話者に影響を与えると言われており，聞き手の反応の遅延からくる影響の軽減についての研究[5] が行われている．人の発した音声を PC で認識して反応を返すシステムで実験を行い，反応に遅延をかけることで人の音声の認識率がどのように変化するかを調査している．遅延があることで認識率が下がり，遅延に相槌のような音声を被せることで影響が軽減されることが検証された．

また，発話交替時の音声遅延による印象への影響についての研究[6] などもされている．音声のみのコミュニケーションで，別室にいる人と協力をして課題を解く実験が行われた．会話における遅延は，相手の印象を悪くすることが分かった．

複数の会話内容の参照を手助けするために，視覚情報と音声の方向を提示することで内容の把握を容易にしている研究[4] が行われている．複数人が参加している 3D チャットルームの中で，いくつかの会話のグループを形成して様々な会話をする．他の会話のグループの会話を参照しやすくするために，話している人の方向と誰が話しているのかを提示することで認知負荷を軽減しいる．

本研究と似た試みを行っている研究として，Snail Light Projector[3] という研究がある．プロジェクタと手で持つことが出来るスクリーンの距離を計測して，再生する映像を変化させている．光速を疑似的に遅くすることで，スクリーンの位置によって映像が変化する．これを利用して新しい画像検索の手法と表現方法を実現している．本研究では，音速を遅くすることで，会話の新しい音声情報検索や表現方法が生まれると考えられる．

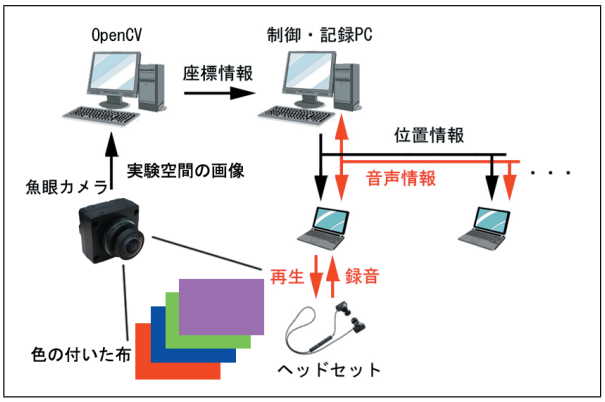

図 1　システム構成

音声の遅延に関する研究は，会話の阻害や印象の悪化についてなどネガティブなものが多い．しかし，音声が遅延することでのメリットについて書かれているものは少ない．また，音声を用いたコミュニケーションを補助するシステムは，仮想空間で行われているものが多く実空間を対象にしているものは少ない．

3. 超低音速空間 CreepingVoice

疑似的な超低音速空間を実現するシステム CreepingVoice を構築した．システムの構成を図 1 に示す．CreepingVoice は，位置計測，録音・再生，防音の 3 つの機能で構成される．

3.1 位置の計測

位置計測の方法は，魚眼レンズの付いたカメラ（(株) ブイシャープ MABEL MB-2000(1280 x 960 pixel)）と OpenCV[2] を用いた画像処理で行う．魚眼カメラを高さ 240cm の天井に設置（図 2）して，実験環境を撮影する．

被験者は，赤，青，緑，黄，紫の五色の布のいずれかをそれぞれに身に着ける．魚眼カメラで取得した画像を，OpenCV でそれぞれ色に分けて領域を抽出する．抽出したそれぞれの色の領域について，最も領域が広いものの重心を，その色をまとった被験者の座標とする．位置検出の画像を図 3 に示す．図中，重心を中心に青色の円を描画している．また，色によって個人を判定して，座標から現在の位置を検出する．検出した位置情報は，制御用の PC を介して全ての被験者

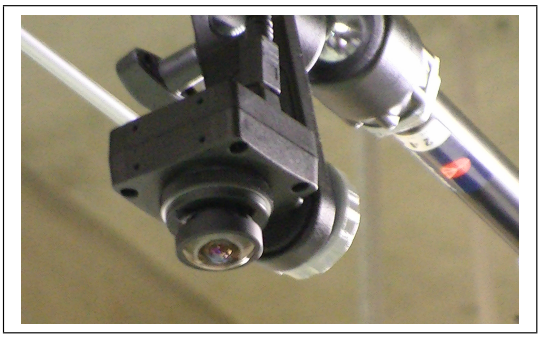

図 2　魚眼カメラ：MB-2000

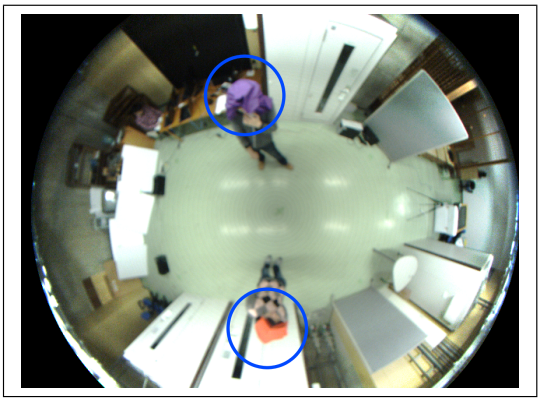

図 3　位置検出

が保持しているクライアント PC にブロードキャストされる．ブロードキャストされた位置情報は，個々のクライアント PC によって受信され，保存される．

3.2　録音・再生

会話音声の録音と再生を行う機材として，Bluetooth ヘッドセット（ロジテック（株） LBT-PCHP04BK）を使用する．ヘッドセットと各クライアント PC 間を通信するための Bluetooth アダプタには PLANEX 社製の BT-Micro3E1XZ を，音声の録音と再生の制御には BASS audio library[1] を用いた．

3.2.1　録　　音

システムが起動されたら，各クライアント PC は常にヘッドセットのマイクから被験者の声を取得し，得られた音声データを全 PC へブロードキャストする．各クライアント PC は，ブロードキャストされた音声データを受信して，どのクライアント PC で録音された音声か分かるように IP アドレスで振り分けながら保存する．さらに全ての被験者の位置情報と音声情報を関連付けしながら保存していくことで，どの被験者が数秒前にどこにいたかが分かるようにする．

3.2.2　再　　生

音声の再生にあたっては，各被験者の位置情報をもとに全被験者間の距離を求め，設定された音速に基づいて音声の到達時間を計算し，各クライアント PC 上で各被験者によって録音された音声の再生タイミングを決める必要がある．今回の実装では，被験者間距離の計算負荷を軽減するために，実験空間を 50cm × 50cm の正方形のメッシュに区切り，各被験者の座標をメッシュ位置で扱うこととした．例えば，音速が秒速 50cm に設定されている場合，あるメッシュ内にあるクライアント PC 上では，1m 離れたメッシュ内のクライアント PC で 2 秒前に録音された音声が再生される．このように，距離に比例した遅延を与えることによって，擬似的に低音速な空間を形成する．

3.3　防　　音

実験では，発話した声が CreepingVoice を経由せず直接相手に伝わってしまうことを防ぐ必要がある．このため，スリーエムヘルスケアのイヤーマフ H10A とカナル型のヘッドセットを使用して，外部の空気振動が耳に入ることを極力防止する．さらに，実験空間にマスキングノイズとして雑踏の効果音を流す．

4.　会話可能音速の調査

通常の空気中における音速と異なれば異なるほど，対面会話の負荷が高くなると考えられる．そこで，CreepingVoice を実装するために会話が可能な音速についての調査を行う．

4.1　実 験 内 容

音速調査実験では，被験者 2 人にシステムを用いて対面会話を行ってもらう．被験者間の距離を 1m に固定して，様々な音速で会話行う．被験者には実験の目的である，会話が可能な音速の調査ということを教示する．会話内容については，特に指定はしなかった．図 4 に実験の風景を示す．音速は，秒速 10cm, 20cm, 30cm, 40cm, 50cm, 60cm の 6 とおりとし，それぞれの条件で 5 分間会話をする．図 5 に，被験者が実験機材を装着した様子示す．

4.2　実 験 結 果

分析のためのデータとして，会話の音声情報，被験者の実験中の映像，実験後のインタビュを収集した．

音速 60cm/sec の条件における会話の音声波形を図 6 に示す．被験者 1 と 2 の音声波形を比較すると，有声部が交互に現れており，話者交代が特に問題なく行われていることがわかる．このように，音速が 60cm/sec と 50cm/sec の時は，普段の会話に比べて発話してから応答までの時間がやや増加するものの，大きな問題はなく会話をすることができていた．

音速が 40cm/sec の条件における会話の音声波形を図 7 に示す．被験者 1 と 2 の音声波形を比較すると，

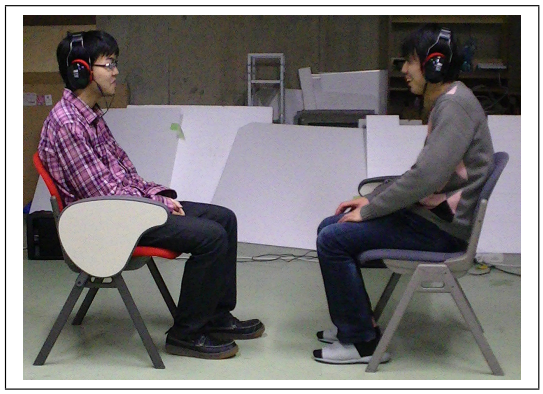

図 4　音速調査実験風景

図 5　イヤーマフとヘッドセットを装着した被験者の様子

図 6 に比べて，有声部の交互出現に乱れが見られ，同時に発声していたり，逆に同時に黙っている状況が多発し始める．このように，音速が 40cm/sec 以下になってくると，会話の衝突や会話中の無音時間が増え，会話の流れが途切れやすくなった．

音速が 20cm/sec 以下になると，相手の応答より早く新しい話題の会話を開始してしまうことが多発し，音速が 10cm/sec では，複数の会話の流れが同時に発生していた．音速が 10cm/sec の条件における会話の音声波形を図 8 に示す．この結果を見ると，音速 40cm/sec の場合よりもさらに有声部の交代が乱れ，特に赤色のハッチング部分では，被験者 1 が被験者 2 の発話と無関係に話し続けている様子が見て取れる．

音速 60cm/sec, 40cm/sec, 10cm/sec の 3 条件について，会話内容を書き起こし，各発話の話題をもとに発話対を求めた結果を図 9 に示す．図中，□の中のアルファベットが話題を示しており，数字は同じ話題に関する何番目の発話であるかを示している．図 9 から，音速が 60cm/sec の時は，話者交代が順当に行われ，

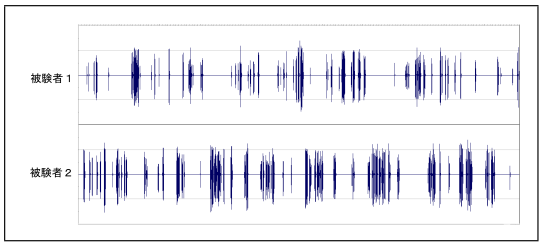

図 6　音速 60cm/sec の音声波形

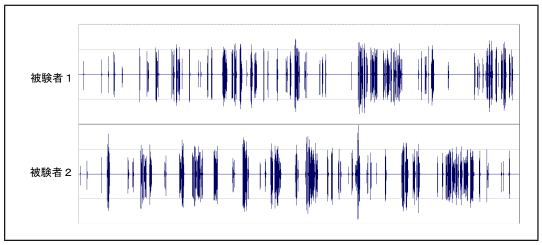

図 7　音速 40cm/sec の音声波形

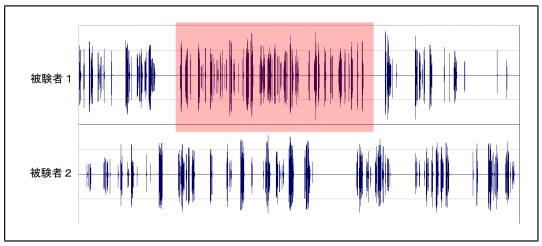

図 8　音速 10cm/sec の音声波形

ひとつの話題について発話のやりとりが連続的に生じていることがわかる．これに対して音速が 40cm/sec の場合，連続して発話がやりとりされる話題もある一方，先行発話への応答が戻ってこないうちに新しい話題の発話が生じる例もいくつか見られるようになる（B1 や C1，D1 など）．さらに音速が 10cm/sec の場合は，発話のやりとりがほとんど生じなくなり，相手の先行発話に対する応答は，あってもごく短い相づち程度のものとなる．大半の発話は，それぞれの被験者が話したいことを一方的に話すだけのものとなり，発話対がほとんど形成されなくなってしまう．このように，音速が遅くなるにつれて話者交代がされにくくなり，会話が成立しなくなる．

この他の特徴として，音速が遅くなるにつれて相手の応答に対して聞いているかどうか確認をする行動が現れること，発話に対する応答の表情やジェスチャなどを見るために会話が中断することなどが観察された．また，1 つの話題に関する会話の開始から終了までの時間が短くなる傾向も見られた．さらにインタビュの結果から，音速 40cm/sec で既に会話負荷が高くなっていたこと，視覚情報と音声情報の不一致に強い違和感を感じることがわかった．

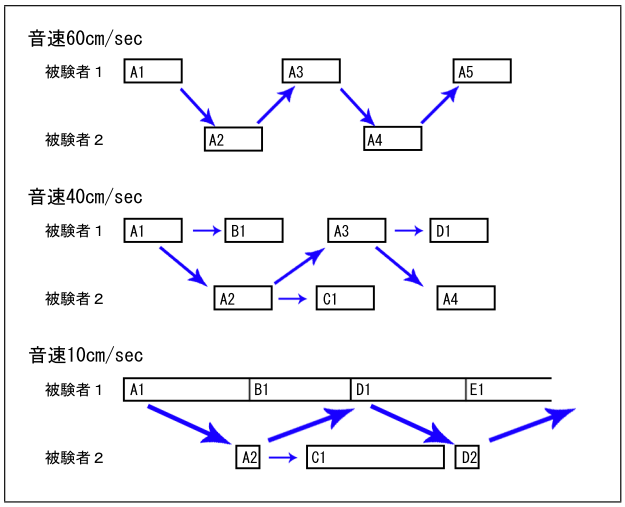

図 9　会話の推移

発話に対する応答は，音声以外にも表情やジェスチャなど視覚情報として受け取ることが可能なものもあるので，音速が遅くなると，音声情報と視覚情報のずれが大きくなる．このずれが会話の混乱を引き起こしていた可能性が考えられる．インタビュ結果に示された違和感も，このことを裏付けている．さらにインタビュの中で，音声の重畳が可能そうかどうかについても尋ねた．その結果，音速が 30cm/sec 以上では，音声に追いついて新しい音声をかぶせてみようとは思わなかったという意見が得られた．

このように，対面で会話可能な最低限の音速と，音声重畳を行える音速との間にはギャップがあることがわかった．会話の負荷が高くならない音速では，速すぎて音声を重畳させることが出来ない．逆に，音声の重畳が可能な音速では，会話の負荷が高くなりすぎる．

以上の結果から，CreepingVoice 空間内では，常時どこでも一定の低音速とするよりは，話者の位置関係や対話状況などに応じて柔軟に音速を可変とすることが必要となると考えられる．たとえば違和感の無い対面対話と，1 章で述べたような会話の輪へのスムースな途中参加との両立のためには，話者間の距離が遠くなるほど音速を遅くすることが求められる．また，非常に発言内容に慎重を期す必要があるような場合は，対面対話の自然さを犠牲にしてでも音声重畳が可能なレベルに音速を遅くすることが望ましいであろう．

5. 音声重畳実験

5.1 実験内容

この実験では，被験者の位置は固定せず自由な移動が可能な状態で音声の重畳を行うことが可能なのかを検証する．被験者２人に，移動が可能な状態での音声重畳を対面会話の中で行ってもらう．被験者に，3.5m × 6m の空間の中で自由に移動をしながら会話をしてもらう．音速は，遅延なしと 50cm/sec，10cm/sec の３パターンを行ってもらう．会話内容は自由で，被験者の要望があった場合のみテーマを与える．会話はそれぞれ 10 分間行う．被験者には，実験前に音声の重畳を目的としていることを提示する．被験者には，イヤーマフとヘッドセット（カナル型），それぞれ異なる色の付いた布を身に着けてもらう．音声重畳の実験風景を図 10 に示す．

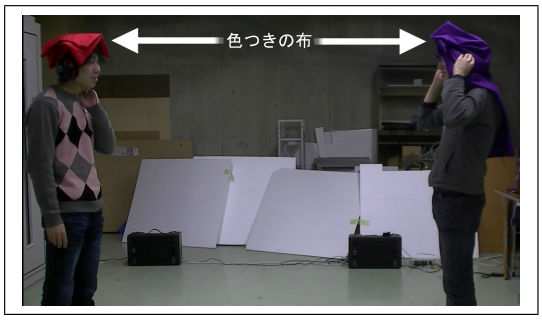

図 10　音声重畳実験風景

5.2 実験結果

分析のためのデータとして，会話の音声情報，被験者の実験中の映像，実験後のインタビュ，魚眼カメラで計測した位置情報を収集した．遅延なし，音速 50cm/sec，音速 10cm/sec の 3 パターンとも，被験者が会話中に移動することは少なかった．また，音声重畳が行われることも少なかった．被験者は相手の発話内容が聞き取りにくかった場合，ジェスチャを使用して聞き直しの意思表示をしていた．

インタビュの中で，音速が遅いと発話しにくい状況になり，発話の回数が減る分，1 回の発話でより多くの情報を伝えうることが必要になることが指摘された．このため，1 回の発話内容により多くの時間をかけて考えるという意見があった．

通常，会話中に音声重畳を行うことはできないので，音速を遅くしても即座に音声重畳できるようになるものではないと考えられる．また，音速の速さが分かりにくいため，音声の重畳が難しいということが分かった．音声重畳を促す方法の１つとして，音声の位置を可視化すると容易になる可能性がある．

6. 会話への途中参加実験

6.1 実験内容

この実験では，被験者が自由に移動可能な状態で，過去の音声の参照を行うことによって会話へのスムースな途中参加が可能かどうかを検証する．被験者は 3

人で，2 人は最初から会話をしているグループとし，残る 1 人が会話へ途中参加する．会話グループの 2 人は，システムを介さず，通常の対面対話で会話を行う．途中参加者はヘッドセットを身に着け，システムを介して会話内容を参照する．音速は，遅延なし，50cm/sec，10cm/sec の 3 条件とする．実験に先立ち，被験者全員に研究の目的を説明する．また，各条件における実験中，会話グループには話している話題を途中で変えないように教示した．

実験を開始する前に，途中参加者には防音室に入ってもらい会話グループの会話が聞こえないようにする．会話グループが 3 分間会話をしたところで，途中参加者が実験空間に入り，3m 離れたところから会話グループに接近を開始する．その後，会話グループは 5 分間会話を続ける．途中参加者には，この 5 分間に会話グループの会話内容に追いつき，会話に参加することを求めた．

6.2 実験結果

分析のためのデータとして，会話の音声情報，被験者の実験中の映像，実験後のインタビュ，魚眼カメラで計測した位置情報を収集した．

遅延無しでは，途中参加者が会話グループそばに移動するまでに 11 秒，途中参加者に対する過去の会話内容の説明に 20 秒かかった．その後 3 人での会話になった．音速 50cm/sec では，途中参加者が会話グループそばに移動するまでに 27 秒，その後，途中参加者は実験環境を 53 秒間動き回り会話に参加した．会話グループからの途中参加者に対する過去の会話内容に関する説明はほとんどなかった．音速 10cm/sec では，途中参加者が会話グループそばに移動するまでに 24 秒，その後 20 秒間は，発話をしないで内容を聞いていた．さらに会話グループからの途中参加者に対する過去の会話内容に関する説明が 50 秒間行われた．

遅延無しの場合，会話グループから過去の会話内容を説明して貰う以外に今までの会話を知る方法が無いのは当然である．音速 10cm/sec では，音速が遅すぎるため，離れた位置では今の会話に参加するにはあまり有用ではない過剰に古い会話内容が再生される．このため，途中参加に必要な過去の会話内容を聞くためには，会話グループに非常に接近しなければならない．結果として，必要な過去の会話内容を把握できていない段階で，会話グループが形成する F 陣形の P 空間[7]に入り込んでしまう．このため，元々の会話グループはこの途中参加者を無視できず，途中参加者に対して過去の会話内容を説明することになるものと思われる．一方音速 50cm/sec では，途中参加者が F 陣形の R 空間（あるいはそのさらに外側）にいる状態で必要な過去の会話内容を参照できる．途中参加者が 53 秒間実験環境を動き回っていたのは，P 空間に入り込まずに必要な過去の会話内容を参照していたものと思われる．こうして，必要な会話内容を参照したのちに P 空間に入るため，スムースに会話に参加することができたと考えられる．

7. まとめ

本研究では，超低音速空間 CreepingVoice を構築し，超低音速での対面コミュニケーションについて検討した．対面会話が可能な音速と音声の重畳が可能な音速の調査をして，両者を両立させることが難しいことが分かった．CreepingVoice を用いて音声重畳を試みたが，通常音速の世界では音声重畳をしないため，本システムによって可能になったとしても，実際に行うことは容易ではないことがわかった．また，会話への途中参加実験を行った結果，音速 50cm/sec で会話へスムースに途中参加可能となることが分かった．今後は，音声重畳や会話の輪への途中参加をさらに直感的に行えるように CreepingVoice 再構築し，これを用いたより一般的な会話実験を行い，超低音速空間での人の対話行動についての分析を行う予定である．

謝辞

本研究の一部は，平成 21 年度（財）栢森情報科学振興財団の研究助成を受けて実施された．ここに謝意を表する．

参考文献

1) BASS.http://www.un4seen.com/
2) OpenCV.http://opencv.jp/
3) 松崎 圭佑 , 甲田 春樹 , 岩井 大輔 , 佐藤 宏介 :Snail Light Projector, インタラクション 2009.
4) 大橋 純, 広渕 崇宏, 河合 栄治, 藤川 和利, 砂原 秀樹:覚情報により強化された 3D サウンド場による共有型多人数音声チャットシステムの設計と実装, 情報処理学会研究報告. CSEC, [コンピュータセキュリティ] 2006(26), 227-232, 2006-03-16.
5) 大西 仁, 望月 要:伝送遅延コミュニケーションに与える心理的影響とその軽減：非分化性音声応答と視覚刺激の効果比較, 電子情報通信学会技術研究報告. TL, 思考と言語 106(23), 13-18, 2006-04-14
6) 小川 一美 , 斎藤 和志:発話交替時の音声遅延による印象への影響, 日本心理学会第 66 回大会発表論文集, 121. 26-Sep-2002
7) Kendon, A: Conducting interaction: Patterns of behavior in Focused Encounters, Studies in International Sociolinguistics 7, Cambridge University Press, 1990.