情報処理学会論文誌 Vol.54 No.6 1794–1806 (June 2013)

国際電話会議の負担を軽減する手法の提案と評価

山下 直美[1,a)] エチェニーケ アンディ[2] 葛岡 英明[3]
石田 亨[4] ハウタサーリ アリ[4]

受付日 2012年12月28日, 採録日 2013年3月1日

概要：非母語で多人数音声会議に参加することは，大変な困難をともなう．対面会議ではほぼ問題なく会議に参加し議論できていても，多人数音声会議になると，理解度が低減し，発言権を得ることが困難になってしまう人が大半である．このような状況が続くと，非母語者の多人数音声会議における存在感が急激に低下してしまう．本研究では，電話会議中に非母語者が複数の高度な情報処理を並列に行わなければならない過負荷な状態に陥っていることに注目し，非母語者に付加的なリソースを与えることを考える．具体的には，非母語者にだけ母語者間の対話交替時に0.2〜0.4秒の無音区間を人工的に付与することによって非母語者の電話会議の参加に対する負担軽減を試みる．評価実験の結果，録音済みの音声会議に無音区間を挿入した場合，非母語者の会話に対する理解度が有意に増加することが分かった．しかし，リアルタイムの電話会議では，無音区間の挿入によって非母語者に一定のメリットを生むものの，会話に不自然さが生じるといった負の効果も生じることが分かった．

キーワード：多人数会話，電話会議，非母語者，通信遅延，会議支援

A Method That Ease the Burden of Non-native Speakers' Participation in Conference Calls

Naomi Yamashita[1,a)] Andy Echenique[2] Hideaki Kuzuoka[3]
Toru Ishida[4] Ari Hautasaari[4]

Received: December 28, 2012, Accepted: March 1, 2013

Abstract: If audio conferencing is difficult for native speakers, then it is extremely taxing for non-native speakers (NNSs). Audio conferencing is particularly difficult for non-native speakers because they have to cope both with imperfect audio signals and imperfect language skills. Previous research has shown that non-native speakers need to pay extra time and effort when recovering from missed conversation contexts under adverse audio conditions. In this paper, we focus on such extra time required for non-native speakers and examine whether inserting short silent periods between conversational turns (as short as 0.2–0.4 seconds) helps their communication abilities. First, we manually inserted short silent periods in a previously taped meeting and found that these intervals significantly helped non-natives understand the conversation better. Second, we inserted the short silent periods in a real-time audio conference by controlling the transmission delay between native and non-native speakers (i.e. adding artificial delay only among native speakers) and found that there is an optimal range of delay that benefits the non-native speakers without hampering the native speakers.

Keywords: group communication, conference call, non-native speaker, delay, communication support

1. はじめに

電話会議の利用は，増加の一途をたどっている．電話会議を用いることによって，チームのメンバが複数地点に散

[1] 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, Kyoto 619–0237, Japan
[2] カリフォルニア大学アーバイン校
University of California Irvine, Irvine, California, USA
[3] 筑波大学大学院情報工学研究科
Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki 305–8577, Japan
[4] 京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University, Kyoto 606–8501, Japan
[a)] naomiy@acm.org

© 2013 Information Processing Society of Japan

らばっていても，素早い意思決定を安価に行うことができるためである．特にメンバが複数の国にまたがるような国際チームでは，全メンバが一堂に会することが困難なため，電話会議は重要なコミュニケーション手段となる．以前は，電子メールを用いて時間をかけて議論されることも多かったが，最近は電話会議を用いて短時間で意思決定がなされるケースが増えている．

このように地理的に分散したメンバ全員の合意を得ながら短時間で意思決定を行えることは電話会議の大きな利点である．しかしその一方で，電話会議が非母語者に大きな負荷をかけていることは見逃すことはできない [13], [15], [16]．電話会議のように音質の良くない会話環境において，聞こえなかった箇所を頭の中で補完しながら理解し，それと同時に会議に貢献するための発言内容を考えることは，非母語者にとっては大きな負担である [13], [16]．Rogers らによると，バイリンガルであっても雑音下における会話の理解度は母語者に劣ることが報告されている [14], [17]．このような場で重要な意思決定を行うことは，非母語者にとって大変に不利である．

そこで本研究では，対面会議では十分な会話能力を持っている非母語者に焦点を当て，彼らの電話会議参加を支援することを目指す．これまで，電話会議に関する様々な研究がなされてきたが [3], [4]，非母語者の電話会議参加を支援する研究はほとんどなされていない．特に，国際的なグループに属している非母語者は，対面での会議にはほとんど問題なく参加しているように見えることが多いにもかかわらず，実は電話会議では大きな負荷がかかっていることは見落とされがちである．

2. 予備調査

本論文ではまず，予備調査として，わが国の情報分野の第一線で活躍中の研究者がどの程度，国際電話会議の参加に際して負担を感じているかを調べるために簡単なアンケート調査を実施した．情報分野（AI，HRI，HCI など）で国際的に活躍中の 7 人の研究者に協力を依頼し，非母語（英語）を用いた電話会議でどの程度，会議内容の把握や言いたいことの発言ができているか（できていると感じているか）を質問した．これら 7 人の研究者は，国際会議の委員などをこれまで 10 回以上務めた経験があり，英語のレベルはきわめて高い．海外での滞在期間は，平均で約 2 年であった．

アンケートでは，彼らが英語を用いた対面会議でどの程度会議の内容を把握できていると感じているか，またどの程度言いたいことを発言できていると感じているかを 5 段階の Likert 尺度で答えてもらい，これらの値を電話会議の場合の値と比較した．**表 1** と**表 2** は，それらの結果をまとめたものである．表 1 と 2 により，電話会議では対面会議と比べて，理解度，そして言いたいことを発話できる度合いがともに大幅に減少していることが分かる．

表 1 非母語者が会議内容を把握できた度合い
Table 1 NNS perceived ability to follow conversation.

理解度(%)	0-20	20-40	40-60	60-80	80-100
対面会議	0	0	1	2	4
電話会議	0	0	4	2	1

表 2 非母語者が言いたいことを発話できた度合い
Table 2 NNS perceived ability to make an utterance.

伝達度(%)	0-20	20-40	40-60	60-80	80-100
対面会議	0	0	3	2	2
電話会議	3	2	0	1	1

今回の調査は小規模であるため，この結果が電話会議に参加するような非母語者に広くあてはまるということはできないが，日本を代表する研究者の 7 割が英語の電話会議で言いたいことの半分も発言できないと感じていることは興味深い事実である．

3. 研究目的と仮説

予備調査の結果をふまえ，本研究では，対面会議で議論できる程度の英語スキルを持った非母語者にターゲットを絞り，彼らの電話会議の参加支援方法を検討する．これらの非母語者が電話会議の参加時にかかえる問題は，英語スキルの欠如というより，複数の高度な情報処理を並列に処理しなければならないことに起因する [1], [20]．雑音の混入などによって失われた会話の断片情報を文脈から想像して補完することは，母語者にとってはさほど難しいことではないが，非母語者が音声情報のみを手がかりにしてこの処理を行うことは負荷が高く，会議内容を考えたり後続の発話内容を聞き取ったりすることが困難になってしまう [20]．このようにリソース不足な状態が続くと，非母語者は会話についていくことができず，置き去りにされてしまいがちである．

そこで我々は，非母語者に時間というリソースを付与することによって，彼らの負荷を軽減できないかと考えた．無音区間を会話中に挿入することによって，非母語者の不足しがちなリソースを補うことができれば，過負荷な状態から脱却し，タイムリーな情報処理が可能になるかもしれない．

本論文では，母語者間の話者交代時に微小な無音区間を挿入し，これによって非母語者の電話会議での活動を支援できるか否かを調べる．そのために，我々はまず，録音済みの会議音声に対して無音区間を挿入し，非母語者の理解度が向上するか否かを調べた．次に，リアルタイムの電話会議において，母語者間の通信に人工的に音声遅延を付加することによって非母語者が母語者よりも先に発話内容を聞き取れる仕組みを作り，これが非母語者の理解度や発話頻度に及ぼす影響を調べた．

本研究の貢献内容は以下の 2 点である．第 1 に，母語者

© 2013 Information Processing Society of Japan

間の話者交替時に 0.2 秒や 0.4 秒といった微小な無音区間を挿入することによって非母語者の会話に対する理解度を有意に向上させることができることが分かった．第 2 に，通信遅延および通信遅延のずれから発生する無音区間が母語者と非母語者のコミュニケーションにどのような影響を与えるかが明らかになった．これまで通信遅延がコミュニケーションに及ぼす影響について様々な研究がなされてきたが [19]，母語者と非母語者のコミュニケーションに及ぼす影響についてはほとんど報告がない [25]．

本論文の構成は以下のとおりである．まず関連研究を紹介し，本研究の位置づけを確認する．次に，録音済みの会議音声に無音区間を挿入したことが非母語者の理解度にどの程度影響を及ぼすかを調べた実験を紹介する．さらに，人工的な遅延を付加することによって無音区間を生成する仕組みを説明し，これをリアルタイムの音声会議に導入することによって非母語者の理解度や発話がどのような影響を受けるかを調べた実験を紹介する．最後に，これらの実験結果をふまえ，非母語者の電話会議参加を支援する方法について考察を述べる．

4. 関連研究

多人数が参加する電話会議は，非母語者に限らず母語者にとっても難しいことが指摘されている．たとえば，雑音混入や音質低下によって会議内容の把握が困難になりやすい点，誰が参加しているか分かりにくい点，誰が発言しているか分かりにくい点，被験者の理解状況を確認しにくい点などが電話会議の一般的な問題点としてあげられている [28]．

これまでに，これらの問題を改善する様々な手法が提案されてきた．残響除去や空間音響のように電話会議における音環境の改善に取り組むものから，電話会議をよりリッチなメディアに拡張することによって電話会議に足りない機能を補完しようとする試みまで，その取り組みは多岐にわたる．音環境の改善に関する研究では，電話会議システムに空間音響を組み込むことによって，話者と雑音の聞き分けを可能にし，これによって会話に対する理解度を向上させる方法が提案されている [9], [26]．電話会議の機能を拡張する研究としては，たとえば音声会議にビデオ映像を付加することによって，参加者の会話に対する理解度を向上させる方法が提案されている．特に多人数が参加する遠隔会議では，参加者のビデオ映像を見せることによって，円滑な話者交代 [24] や積極的な参加を促進することができる [7] とされている．また，Yankelovich らは，テキストチャットのチャネルを音声会議と併用することによって，参加者同士が会議の進行を妨げることなく議長に発話リクエストを出したり別の参加者と個別相談ができたりするようにした [27]．Junuzovic らは，なんらかの事情で会議を部分的に聞き逃した参加者が，少し前の会話内容を早送りで聞くことができるようにし，これによって参加者が会議の進行を妨げることなくキャッチアップできる機能を提案した [8]．

これらの提案の中には，空間音響 [5] やビデオ映像の付加のように，非母語者にとっても価値があると思われるものが存在する．たとえば Veinott らは，音声対話にビデオ映像を付加することによって，母語者が非母語者の理解度を推定しながら話せるため，相互理解を促すことができることを示した [23]．その一方で，母語者に有効とされた機能が必ずしも非母語者にも有効であるとは限らない．たとえば，電話会議に参加しながらテキストメッセージをタイピングしたり聞き逃した部分を早送りで聞き直したりすることは，非母語者にさらなる負担を強いることになる．非母語者が電話会議に参加するだけですでに過負荷な状態であることを考えると，非母語者にさらなる負担を強いるような機能は彼らの役に立つとは考えにくい．むしろ，非母語者の負担を軽減する機能が望ましい．

5. 本論文における研究

本論文は，電話会議に参加する非母語者の負担を軽減する方法を探ることを目的とする．本研究で対象とするユーザ（非母語で対面会議に参加できるくらいの外国語能力を十分に有するユーザ）がかかえる問題は，複数の高度な情報処理を並列に行わなければならない過負荷な状態に陥っていることである．これに対する本研究の提案は，付加的なリソース（たとえば，処理能力や時間）を提供することによって非母語者を支援することである．そこで研究の初期段階として，非母語者に付加的な処理時間を提供することを試みる．そのためには，会話の速度を遅くする方法と，会話の途中に無音区間を挿入する方法が考えられる [29]．電話会議では通信遅延のために，短い無音声状態を経験することがよくあることを考えると，後者の方法の方がユーザは違和感を感じにくく慣れやすいと思われる．そこで本研究では，無音区間を挿入する方法を採用することとした．

本研究では，この方法の妥当性を検証するために 2 つの実験を実施した．2 つの実験の違いは，無音区間がインタラクティブな状況に用いられたか否かということである．

6. 実験 1：微小な無音区間が非母語者の理解度に及ぼす影響

最初の実験は，3 種類の異なる無音区間（0 秒，0.2 秒，0.4 秒）が，英語の音声会話に対する「理解度」と「理解に際して感じる負担」に与える影響を計測することを目的とした．これらの微小な無音区間が，母語者と非母語者に同様の効果を与えるかどうかを確認するために，母語者と非母語者の比較も行った．

© 2013 Information Processing Society of Japan

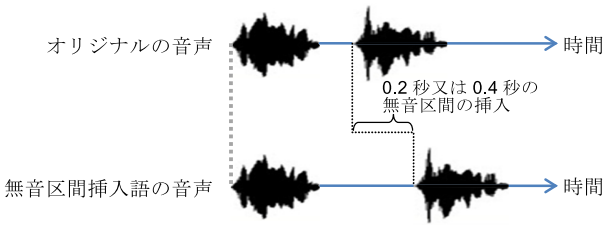

図 1　微小無音区間の挿入
Fig. 1　Insertion of short gaps.

6.1　実験方法

6.1.1　被験者

被験者は英語母語者 10 人と，英語非母語者 10 人であった．母語者はいずれも英語を母国語とする国で生まれ育った．非母語者は全員日本人であった．日本人被験者の英語の習熟度は様々であったが，全員英語を 10 年以上学んでいた．ほとんどの日本人被験者は英語には自信があり，8 人が英語に堪能であると自己評価していた．彼らが英語圏で生活した平均年数は 0.8 年であった．

6.1.2　実験設計

各被験者は，内容が違う 3 つの会議音声のヒアリングタスクを課せられた．会議音声に挿入する無音区間の長さは 3 種類（0 秒，0.2 秒，0.4 秒）とし，会議の内容と無音区間の長さの組合せ，および提示の順番はカウンタバランスをとった．

6.1.3　実験材料

録音済みの会議音声として，Linguistic Data Consortium (LDC)[*1]が公開している会議コーパスを利用した．この音声は，4 人の英語母語者が，9.11 の事件が彼らの実生活にどのように影響したか，また新技術によっていかにテロリズムを防げるかということに関して討論した様子を記録したものである．この音声データは，会議参加者に取り付けられた 2 つのマイク（ノイズキャンセル機能付きの接話型ブームマイクロフォンと，全方位型のラペルマイク）と，360 度全方位をカバーするように設置した数台の卓上マイクによって記録された．それぞれのマイクの音声は別々のファイルに記録されているため（図 2 の上から 2 番目と 3 番目）1 人 1 人の音声を個別に加工することができた．そこで，各参加者の発話のタイミングを手作業でずらした（微小無音区間を挿入した）のち，すべての発話を再び合成することによって，発話タイミングを人工的にずらした会議音声を生成した（**図 1** と**図 2** 参照）．

図 1 では，2 人の話者による発話の音声波形が時間軸上に記されている．上がオリジナルの音声で，下が微小の無音区間を挿入した後の音声である．図 1 では，もともと 2 つの発話の間に無音区間が存在し，無音区間挿入によってその区間が長くなる．一方で，母語者同士の会話では，しばしば発話がオーバラップする．そのような場合は，オーバラップ長は短くなる．図 2 のように 2 人の話者の発話に 0.2 秒以下（あるいは 0.4 秒以下）のオーバラップがある場合，そのオーバラップは解消される．

[*1] http://www.ldc.upenn.edu/

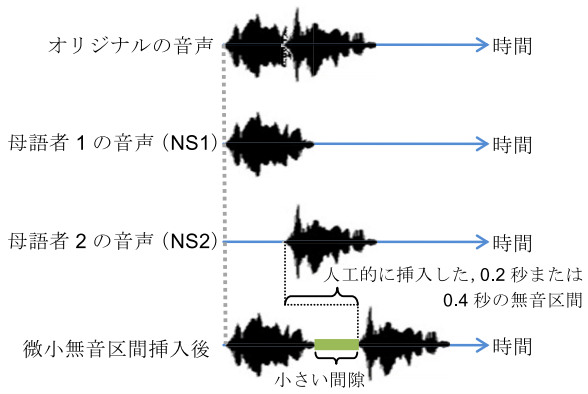

図 2　発話のオーバラップの解消
Fig. 2　Resolving speech overlaps.

ヒアリングタスクの材料として，まず会議音声からランダムに 3 つのフラグメント（会話断片）を切り出した．各フラグメントは約 45 秒間で，10，11 回の話者交代があり，そのうち 4，5 回の発話オーバラップがあった．これらのフラグメントを加工することによって，フラグメントごとに 3 種類の会議音声（オリジナルの音声と，0.2 秒および 0.4 秒の無音区間が各発話の間に挿入された音声）を用意した．0.2 秒と 0.4 秒の無音区間を挿入したフラグメントは，それぞれ 2 秒程度と 4 秒程度，オリジナルの音声よりも長くなった．

6.1.4　タスク

タスクは，異なる長さの無音区間が挿入された 3 種類のフラグメントを聞き，理解度について回答することである．被験者は，各会議音声を聞いた後，書き起こされた紙を渡され，聞き取ることができたすべての文や単語にマークをつけた．このとき，英語母語者は英語で書き起こされた紙を渡され，日本人は，日本語に訳された紙を渡された．

6.1.5　実験手順

被験者にはまず，ヒアリングタスクの中で話されている話題や，実験手順に関する説明を行った．ただし，実験結果に影響を与えないために，3 つの音声の違いについては説明しなかった．

実験説明に続いて，被験者に実験に慣れてもらうために，練習用のヒアリングタスクを実施した．練習用タスクのフラグメントは，本実験で使用したものと同じ会議音声から抜き出したもので，無音区間は追加しなかった．

次に，被験者は 3 つのフラグメント（加工なしのフラグメント，0.2 秒の無音区間を挿入したフラグメント，0.4 秒の無音区間を挿入したフラグメント）のヒアリングタスクを行った．ヒアリングの間，被験者はノートをとってもよいこととしたが，フラグメントの再生は 1 度のみとした．被験者は各フラグメントのヒアリングが終了するごとに，

© 2013 Information Processing Society of Japan

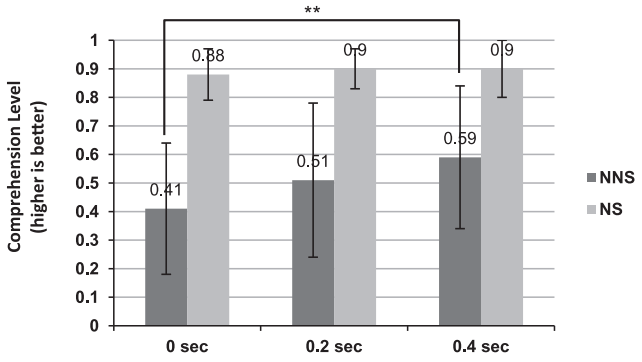

図 3 平均理解度の比較
Fig. 3 Mean comprehension scores by gap.

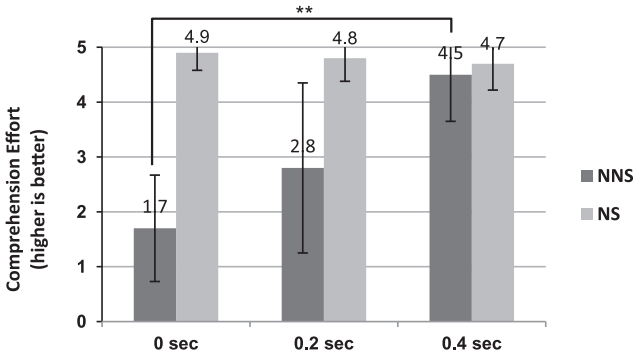

図 4 理解のための負担の平均値の比較
Fig. 4 Mean ratings of perception for comprehension effort.

理解度（書き起こし用紙へのチェック）と，理解に要した負担に関するアンケートに答えた．3 つすべてのヒアリングタスクが終了した後，最後に，3 つのトライアルの違いについて気が付いたことをインタビュした．

6.2 結果

微小な無音区間が非母語者の会話理解に与える影響を調べるために，客観的な評価指標と主観的な評価指標を用いた．客観的な評価指標では，各被験者が聞き取ることができた単語数を，そのフラグメントに含まれる全単語数で割った値（理解度）を計算し，主観的な評価指標では，被験者の理解の負担度合いをアンケートによって計測した．これらの評価値を 3 条件間で比較するにあたり，前者（客観的な評価指標）には反復測定分散分析を適用し，後者（主観的な評価指標）にはノンパラメトリックな Friedman 検定を適用した．

図 3 に見られるように，英語母語者（NS）の理解度は，すべての条件にわたって高かった．これに対して非母語者（NNS）の理解度は，無音区間が長くなるに従って向上した．無音区間長を反復因子として対応のある一元分散分析を行った結果，無音区間長の主効果に有意性が認められた（$F[2, 18] = 5.91$，$p < .05$）．次いで Bonferonni の多重比較を行った結果，非母語者の 0 秒と 0.4 秒の間に有意な差が認められた（$p < .01$）．母語者については，このような有意性は認められなかった．

次に，5 段階の Likert 尺度を用いたアンケートによって，被験者の理解の負担度合いを測定した．その結果，非母語者は，無音区間が長くなるに従って，負担が減ると感じることが分かった（**図 4**）．Friedman テストの結果，無音区間長の効果は有意であった（$\chi^2[2] = 10.47$，$p < .01$）．多重比較（Bonferroni の修正を施した Wilcoxon 符号順位検定）の結果，無音区間長が 0 秒と 0.4 秒の間に有意な差が見られ（$p < .01$），0.2 秒と 0.4 秒の間に有意傾向（$p = .06$）が見られた．理解度と同様に，母語者には有意な差は観測されなかった．

実験後のインタビュの中に，トライアル間の違いに関する興味深いコメントが見られた．多くの英語母語者は，特に無音区間が 0.4 秒のときに会話が少し不自然であると違和感を口にしたが，理解度や負担に影響はなかった．一方，多くの日本人は，無音区間を挿入したトライアルでは会話がゆっくりで聞き取りやすかったと答えた．実験後，日本人被験者に対して，実は会話がゆっくりになったのではなく，無音区間が発話の間に挿入されていたのだと説明すると，驚いた様子であった．

さらに，一部の日本人被験者は，発話のオーバラップがあるときは，会話が速く感じられて聞き取りにくいという感想を述べた．これは図 2 に示すように，手作業で挿入された無音区間によってフラグメント中の大半のオーバラップが解消されたことによって，日本人被験者の会話を聞き取るための負担が減少したためと考えられる．

以上の結果から，会話に微小な無音区間を挿入し，オーバラップを解消することによって，非母語者は少ない負担で会話を理解できるようになることが分かった．

7. 実験 2：微小な無音区間が実時間対話に及ぼす影響

実験 1 により，録音済みの音声対話に無音区間を挿入することが非母語者に対して有効であることが分かった．しかし，通常の実時間対話に同様の無音区間を挿入するということは，母語者に対して発話のたびに，一定の待ち時間を入れてから発話してもらうことになるため，現実的ではない．そこで，通信回線を介した実時間対話において，擬似的に無音区間を挿入する手法を考えた．

図 5 と**図 6** は人工的な遅延を挿入することによって，母語者間の発話と発話の間に無音区間を挿入する方法を示している．話者の直接の音声を黒い色で，ネットワークを介して伝わった音声を灰色で示している．ただし，これらの例はネットワークによる遅延がない場合を示している．図 5 は，母語話者同士の回線にのみ人工的に遅延を挿入し，母語者と非母語者の間はリアルタイムに音声を伝達することによって，非母語者が母語者 NS1 の音声を他の母語者よりも早く聞くことができる様子を示している．

© 2013 Information Processing Society of Japan

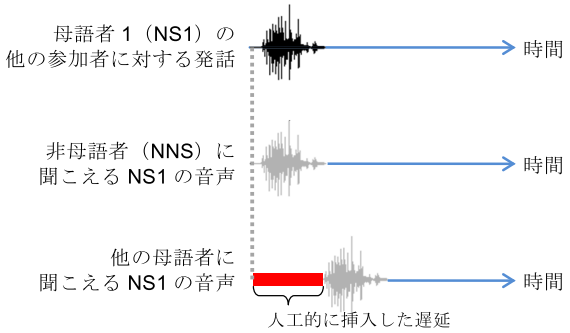

図 5　発話のオーバラップの解消
Fig. 5　Providing NNS with additional processing time by inserting artificial delay among NSs.

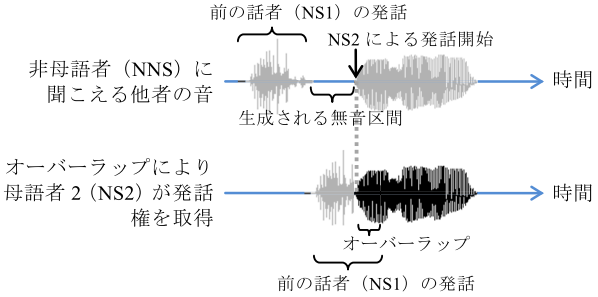

図 6　母語者の発話オーバラップの解消
Fig. 6　Resolving conversation overlaps between NSs.

母語話者間の通信に挿入された遅延によって，非母語者の処理時間を稼げるだけでなく，発話のオーバラップも減少させることができ（図 6），これによって非母語者は各話者の発話をより完全に聞けるようになることが期待できる．

こうして生成された微小な無音区間は，非母語者にとって有効であると期待されるが，母語者に対する影響にも注意しなければならない．実際，従来の研究によって，ネットワーク遅延がコミュニケーションに悪影響を与えることが指摘されている [6], [10], [22]．たとえば，Krauss らは，0.3 秒の音声遅延があると，コミュニケーションのプロセスに障害が生じ，0.9 秒の遅延になると，2 者の対話に致命的な影響を与えることを明らかにしている [10]．同様に，Tang らは，0.57 秒の遅延によって，発話交代の調整が難しくなることを見いだしている [22]．音声遅延に関する過去の研究をまとめると，遅延が 0.2 秒以下であれば人はほとんど気がつかず，遅延が 0.2 秒から 0.4 秒で問題が生じ始め，0.45 秒よりも長くなるとコミュニケーションや会話調整のプロセスに大きな障害が生じ始めることになる．

2 つ目の実験では，0.2 秒と 0.4 秒の人工的な遅延を，母語者の通信間に挿入することによる影響を調べる．先述の理由により，この程度の遅延であれば，母語者に悪影響を与えることなく，非母語者の会話活動を向上させることができるのではないかと考えた．具体的には，本実験では，以下の点について明らかにすることを目的とする．

- 提案する遅延挿入手法が期待どおりに機能するか．すなわち，リアルタイムの電話会議において，母語者の通信間にのみ人工的に遅延を挿入することによって，母語者の発話間に無音区間を生成し，オーバラップを減少させることができるか．
- 提案する遅延挿入手法が，母語者のコミュニケーションにどのような影響を与えるか．挿入する遅延が長くなるに従って，話者交代が難しくなったり，苛立ちや不満感が増大したりするか．
- 提案する遅延挿入手法によって，非母語者の会話活動が向上するか．すなわち，非母語者の理解度や自発的な発話数が向上するか．
- 提案する遅延挿入手法が，電話会議の合意形成にどのような影響を与えるか．また，母語者と非母語者はどの程度，会議での合意事項に納得しているか．

7.1　実験方法

7.1.1　被験者

本実験には，被験者 70 人が有償で参加した．5 人で 1 グループとし，合計 14 グループを構成した．各グループは 5 人中 4 人（合計 56 人）が英語を母語とし，1 人（合計 14 人）が英語の非母語者であった．非母語者は全員日本人であったが，TOEIC の点数 860 点以上，海外滞在期間は平均して 1.6 年であり，日常会話であれば十分な英会話能力を持っていた．

7.1.2　実験設計

5 人一組のグループ（4 人の母語者と 1 人の非母語者）に対して，合意形成タスクを 3 回課した．各トライアルでは 0 秒，0.2 秒，0.4 秒のいずれかの遅延時間が挿入される条件が適用された．1 グループ 5 人とした理由は，ほぼ全員が問題解決のためのディスカッションに参加できる程度に少ない人数であり，その一方，自然に発話の順番が決まってしまうほど少人数ではない人数にするためである．ディスカッションの題材と，遅延時間の長さの組合せは，グループ間でカウンタバランスをとった．同様に実験の実施順序についても，順序効果を考慮してカウンタバランスをとった．

7.1.3　タスク

ディスカッションの題材は，グループトレーニングに頻繁に用いられる「生き残り問題（場所は砂漠，北極，月とした）[11]」を採用した．このタスクは，被験者らのグループを乗せた飛行機がいずれかの場所に不時着した状況を想定し，グループメンバが生き延びるために，機内で発見されたアイテムの内どれから先に機外に持ち出すべきか優先順位を付けるというものである．まず各被験者は個人で順位付けを行い，その後音声会議によってグループ討論を行って合意形成つまりグループとしての順位付けを行う．オリジナルのタスクでは 15 アイテム用意されているが，本タスクでは，簡単のために日常的に使われない道具（たとえば，「大きな絶縁性の布」など）は排除し，残ったものの中

© 2013 Information Processing Society of Japan

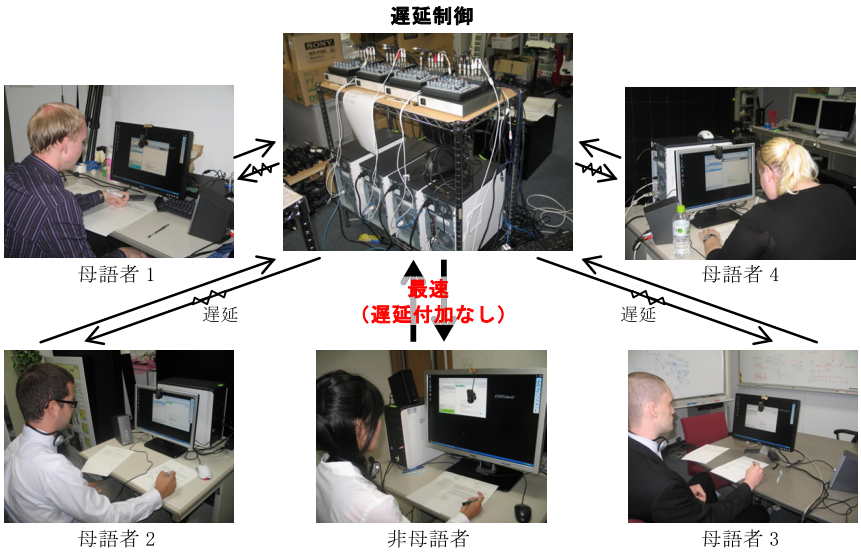

図 7　遅延付加した Skype を利用して 5 人の被験者（母語者 4 人と非母語者 1 人）が意思決定タスクを実施した

Fig. 7　Five participants (4 NSs and 1 NNS) working on a decision making task over Skype with delay control.

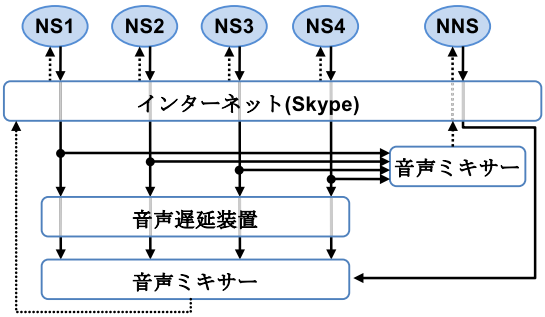

図 8　システムの概略図

Fig. 8　Wire diagram of the system.

からランダムに 6 アイテムを選んで被験者に提示した．

7.1.4　実験装置

音声対話用ソフトウェアには Skype を用いた．各被験者の端末には Skype のインタフェースが表示された．各端末間では，ネットワークの遅延と Skype の遅延によって，約 0.2 秒の通信遅延が存在していた．人工的な遅延は遅延付加装置（**図 7**，**図 8**）によって加えられた．もともと存在する 0.2 秒の遅延時間を考慮すると，母語者の発話は他の母語者のところに，0.2 秒，0.4 秒，0.6 秒の遅延で届き，非母語者のところには，いずれの遅延条件においても 0.2 秒の遅延で届くこととなった．

図 8 は，本研究で用いられたシステムの概略図であり，どのようにして母語者にのみ遅延が追加されたかということを示している．

実験中の発話は Skype の音声録音用のソフトウェアである Tapur を利用して録音した．Tapur はすべての端末にインストールしたため，各端末における音声の状況をそのまま記録することができた．

7.1.5　実験手順

手順（1）：まず被験者は同意書にサインをした後，5 つの部屋に分かれた．そして，簡単な音声テストをした後，Skype を介して自己紹介をした．

手順（2）：異なる遅延時間に対応して，以下の手順を 3 回繰り返した．被験者はまず個別に 5 分間かけて 6 つのアイテムの順位付けを行い，それを書き留めた．次に 15 分間グループディスカッションを行い，グループによる順位付けを行った．ただしディスカッションは，議長などをおかずに自由に行うこととした．これは，提案する遅延挿入手法が，非母語者の自発的な発話に与える影響（すなわち，非母語者が発話を促されない状態で，どの程度自発的に発話できるか）を調べたかったからである．グループによる順位付けが決まった後，グループディスカッションに対する納得度合いを調べるために，被験者は 2 回目の個別順位付けを行った．被験者はさらに，実験中の会話活動に関するアンケートに回答した．

手順（3）：3 つのトライアルを終了した後，被験者は最後のアンケートに回答し，それらのトライアルに関してインタビュを受けた．ただし，最初の実験と同様に，被験者は各トライアル間の違いについては知らされなかった．

7.2　結果

7.2.1　微小な無音区間の生成

まず，本研究で提案する通信遅延の挿入手法によって，実際に無音区間が生成され，非母語者に届く音声において発話のオーバラップを減少させることができたかどうかを確認することとした．そこでまず，母語者間の話者交代を，オーバラップがない場合とある場合の 2 種類に分類した．前者

© 2013 Information Processing Society of Japan

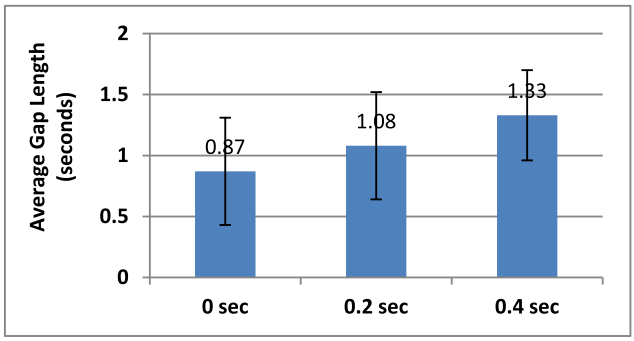

図 9　母語者同士の話者交替時の平均ギャップ長
Fig. 9　Mean length of gaps between native speakers' speech by delay scheme.

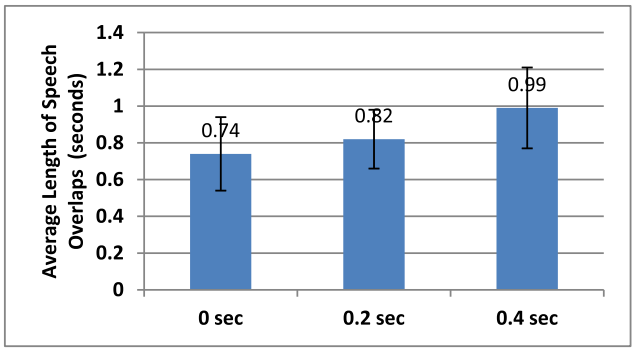

図 10　母語話者同士の発話の平均オーバラップ長
Fig. 10　Mean length of overlaps between native speakers' speech by delay.

の話者交代は，無音区間の平均長を算出するために利用し，後者はオーバラップの平均長を算出するために利用した．

図 9 は母語者間で話者が交代するときの無音区間長の平均値である．予想どおり，通信遅延を長くするに従って，無音区間も長くなっている．遅延時間を反復因子として対応のある一元分散分析を行った結果，遅延時間の主効果に有意性が認められた（$F[2, 26] = 24.09$，$p < .001$）．Bonferonni の多重比較の結果，遅延時間が 0 秒と 0.2 秒の間（$p < .01$），および 0 秒と 0.4 秒の間（$p < .001$）に有意差が認められた．

同様にして，母語話者同士の発話に発生したオーバラップの長さを計測した．我々は，遅延時間を増やすことによってオーバラップが短くなると予想したが（図 6 参照），予想に反し長くなってしまうという結果を得た（**図 10**）．先述と同様に対応のある一元分散分析を行った結果，遅延時間の主効果に有意性が認められた（$F[2, 26] = 9.70$，$p = .001$）．Bonferonni の多重比較の結果，遅延時間が 0 秒と 0.4 秒の間（$p < .05$），および 0.2 秒と 0.4 秒の間（$p < .05$）に有意差が認められた．

このような結果になった主な原因は，複数の母語者がほぼ同時に発話を開始した場合に，通信遅延のために音声が互いに届くまでに時間がかかってしまうため，両者ともそのことに気づかないまま発話を続けてしまうことに起因する（**表 3** の 1 行目，および後述の説明を参照）．

7.2.2　母語者に対する効果

総じて，母語話者は，遅延時間が長くなるほど発話権を獲得しにくくなることが分かった．以下に，詳細な分析結果を示す．

(1) **複数の母語者が同時に発話を開始する事例の増加：** 人工的な遅延が挿入された条件では，他の話者の発話が遅れて聞こえることになる．したがって，母語話者はときとして他の話者がすでに発話を開始していることに気づかずに話し始めてしまい，これによって発話のオーバラップが生じることとなった．このようなオーバラップの頻度を算出するために，まず，各トライアルにおいて複数の母語話者が発話権を獲得するために同時に発話を開始してしまった回数を計測し，次に，この値をそのトライアル間に発生した話者交代の回数で割った．分析の結果，表 3 に示すように，遅延時間が長くなるほどオーバラップ頻度も増加することが分かった．対応のある一元分散分析の結果，遅延時間の主効果に有意性が認められた（$F[2, 26] = 11.58$，$p < .001$）．Bonferonni の多重比較の結果，遅延時間が 0 秒と 0.2 秒の間（$p < .05$），0 秒と 0.4 秒の間（$p < .01$）に有意差が認められた．

(2) **NS 間の話者交代頻度の減少：** 上述のように複数の母語者が互いに気づかないまま発話を開始する状態では，話者交代（あるいは発話権の獲得）が難しくなることが予想される．そこで，話者交代の難しさを調べるために，各トライアルで話者交代が行われた回数を計測した．ただし，この計測では，新たな話者が発話を開始した場合でも，発言を完了できなかった場合は発話交代数から除外することとした．計測の結果，母語者間の話者交代数は，遅延時間が長くなるに従って少なくなることが分かった（表 3）．対応のある一元分散分析の結果，遅延時間の主効果に有意性が認められた（$F[2, 26] = 3.79$，$p < .05$）．以下の例で見られるように，発話交代が難しかった原因は，発話のオーバラップが増加したことと，遅延によって生じた長い沈黙（long pauses）にあったと思われる．

トランスクリプト 1（遅延時間 0.4 秒）

NS1:　When you say fishing kit what do you mean↑ exactly↓(0.8) uh:: well what do you have in mind when you say FISHING KIT↑(1.4)

NS2:　Pro [bably ro::d and]

NS1:　[Mosquito netting is] pretty thin it could rip (.) I mean you know (0.9) well it it it's pretty (0.2) I don't think it's uh good for uh::: for for catching fish.

この例で見られるように，他の被験者がタイミング

© 2013 Information Processing Society of Japan

表 3 遅延時間の増加にともなって，母語者の発話権獲得が困難になる
Table 3 Native speakers having difficulties in taking their turns as delays increase.

	0 sec	**0.2 sec**	**0.4 sec**
Percentage of conversations initiated simultaneously	22%	33%	41%
Frequency of speaker switches (# of speaker switches per min.)	7.8	7.3	6.6
Speaking effort (5pt Likert scale: higher is better)	4.2	4.2	4.0
Comprehension effort (5pt Likert scale: higher is better)	4.4	4.5	4.3
Frustration (5pt Likert scale: lower is better)	1.7	1.8	2.1

表 4 非母語者は遅延時間が 0.2 秒のときに，最もパフォーマンスが良い
Table 4 Non-native speakers performing best with 0.2 second delay trial.

	0 sec	0.2 sec	0.4 sec
Average rate of spontaneous speech	0.76	0.81	0.69
Speaking effort (5pt Likart scale: higher is better)	3.1	3.4	2.3
Comprehension effort (5pt Likart scale: higher is better)	4.1	4.1	3.5
Frustration (5pt Likart scale: lower is better)	2.9	2.7	3.1

良く発話権を獲得できなかったために，母語者が発話を継続した事例が多く見られた．このように話者交代数が少ないということは，話者間の相互作用性（interactivity）の低さを表しており，意見交換の場として好ましい状態とはいえない．

(3) **NS の不満の増加**：実験後のインタビュから，発話開始直後のオーバラップは，自分の話に無理矢理割り込まれたように感じられるため，不作法であるととらえてしまいがちであることが分かった．これによって，不満が増加したと考えられる．表 3 に，母語者の理解の負担の平均値，発話の負担の平均値，そして不満足度の平均値を示す．これらの結果から，いずれの値も，遅延時間が 0.4 秒の場合に最も悪いことが分かる．Friedman 検定の結果，不満足度に関して，遅延時間の主効果が有意であった（$\chi^2[2] = 7.23$，$p < .05$）．多重比較（Bonferroni の修正を施した Wilcoxon の符号付き順位検定）の結果，遅延時間 0 秒と 0.4 秒の間に有意差が認められた（$p < .05$）．

7.2.3 非母語者に対する効果

総じて，非母語者は遅延時間が 0.2 秒の場合を最も好み，成績も最も良かった．遅延時間 0.4 秒の場合は，非母語者にとっても，パフォーマンスと選好の点で最も悪かった．この傾向を示すデータを**表 4** に示す．

(1) **NNS の理解負担の増加**：1 つ目の実験結果に基づき，当初は，遅延時間（遅延挿入によって生成される無音区間）が長くなるほど，非母語者が会話を理解するための負担は減少すると予想していた．しかし実際には，遅延時間が長くなると（0.4 秒のとき），非母語者が会話を理解するための負担は上昇してしまうことが分かった（表 4）．Friedman 検定の結果，遅延時間の効果は有意傾向にあった（$\chi^2[2] = 5.52$，$p < .06$）．

(2) **NNS の自発的な発話の増加，そして減少**：非母語者が実験中に自発的に発話した割合は，各トライアルにおける非母語者の全発話数に対する自発的な発話回数（母語者の誘導によって発話した場合を除いた発話数）として計算した．その結果，遅延時間が 0.2 秒の場合，非母語者の自発的な発話の割合が最も高く，0.4 秒の場合に最も低いことが分かった（表 4）．対応のある一元分散分析の結果，遅延時間の効果は有意傾向にあった（$F[2, 26] = 3.23$，$p = .056$）．

(3) **発話に要する負担の減少，そして増加**：遅延時間が非母語者の自発的な発話に及ぼす影響はさほど大きくはなかったが，非母語者の感じる負担度合いには大きく影響していることが分かった．具体的には，非母語者は，遅延時間が 0.2 秒のときに比較的容易に発話できるが 0.4 秒のときには難しいと感じているようであった（表 4）．Friedman 検定から，遅延時間の要因に有意な効果があることが分かった（$\chi^2[2] = 6.59$，$p < .05$）．多重比較（Bonferroni の修正を施した Wilcoxon の符号付き順位検定）により，遅延時間が 0.2 秒と 0.4 秒

© 2013 Information Processing Society of Japan

表 5 グループによるアイテムの順位付けに対する母語者と非母語者の納得度合い
Table 5 NS and NNS's agreement level of group ranking.

	0 sec	0.2 sec	0.4 sec
Average correlation between NS and Group rankings	0.90	0.89	0.91
Average correlation between NNS and Group rankings	0.68	0.88	0.77

の間に有意差が認められた（$p < .05$）．実験後のインタビュにおいて，ある非母語者は，発話のしにくさについて，「（遅延時間 0.4 秒の条件では）みんな，他の人が話している途中で割り込んでいたので，他人の意見を聞きたくないのではないかと感じるときもありました．（…）それで自分の考えを述べにくくなってしまって，会話を理解することに集中することにしました」と述べた．

以上をまとめると，遅延挿入手法によって，母語者の発話間に無音区間を生成することには成功したが，発話開始時のオーバラップが増加してしまい（複数の母語者が次の発話権を取得しようとすることに起因する），結果的に非母語者に対しても悪影響を与えることとなった．この悪影響は遅延時間が 0.4 秒の場合に最も顕著であった．

7.2.4 合意形成に対する効果

最後に，本提案の遅延挿入手法が合意形成に与えた影響，すなわち被験者がグループとして決定したアイテムの順位にどの程度納得しているか，納得度合いが母語者と非母語者で異なるかどうか，そして納得度合いが各遅延時間に対して異なるかどうかを調べる．

本研究ではまず，各被験者の 2 回目の個別順位付けとグループによる順位付けの相関（Spearman's coefficient）を計算した．各相関係数は，被験者の納得度合い（順位付けの類似度）を表しており，−1（完全に逆）から 1（完全に一致）までの値を示す．母語者と非母語者の相関係数の平均値を比較するために Wilcoxon 検定を行った（**表 5**）．その結果，遅延を挿入しない条件では，非母語者は母語者に比べて納得度が有意に低い（$p < .05$）ことが分かった．この非母語者と母語者の納得度の違いは，0.2 秒と 0.4 秒の遅延挿入時のいずれの場合にも，有意ではなかった．

さらに，Friedman 検定によって，遅延時間と相関係数の関係を調べたが，遅延時間の要因に対する主効果は見られなかった．しかし表 5 からも見て取れるように，非母語者は遅延挿入時間が 0.2 秒のときに最もグループの順位付けに納得していたことがうかがえる．この結果は，前節における，非母語者にとって 0.2 秒の遅延挿入時間が最も好まれ，良いパフォーマンスをあげていたという結果と整合している．

8. 考察

本研究では，2 つの実験結果から 5 つの知見を得ることができた．

- あらかじめ録音された会議音声において，話者交代時のオーバラップを減少させ，微小の無音区間を挿入することによって，非母語者の会話に対する理解度（図 3）と負担（図 4）を有意に改善することができる．
- 我々が提案した遅延挿入手法（人工的な遅延を母語者間の会話に加えること）によって，リアルタイムの音声会議に対しても，母語者同士の発話の間に微小な無音区間を挿入することができた（図 9）．しかし，発話のオーバラップは意図したとおりには減少せず，むしろ増加した（図 10）．
- 母語者は，遅延量が増加するほど話者交代が困難になった（表 3）．具体的には，他者とほぼ同時に発話を開始した場合，遅延のせいでそのことに気づくことができず，しばらくの間オーバラップしたまま発話を続けてしまう様子が見られた．通常は，遅延なしの対話であれば，オーバラップと同時に話者が発話を止めるため，問題は起きにくいが，遅延挿入方式の場合は，遅延量が増加するほど，オーバラップの発生頻度が高くなり，オーバラップ長も長くなってしまった．
- 非母語者は，挿入した遅延量が 0.2 秒の場合に最も自発的に発話することができ，このとき発話の負担に対する自己評価も最小であった（表 4）．フラストレーションも同様の傾向を示し，挿入した遅延量が 0.2 秒の場合に最も低く，0.4 秒の場合に最も高かった．非母語者の会話理解にともなう負担は，遅延量の増加にともなって軽減されるであろうという我々の予想に反し，0 秒と 0.2 秒の場合は同程度で，0.4 秒の場合に増大した．
- 人工的な遅延が挿入されなかった場合（すなわち一般的な電話会議と同じ条件では），非母語者がグループの結論に対して納得した度合いは，母語者と比較して有意に低かった．この非母語者の納得度合いは，挿入した遅延量が 0.2 秒のときに最も高くなり，0.4 秒の場合に再び低下した．

以下では，これらの知見に関して考察した後，音声会議システムの設計指針や今後の研究課題について検討を行う．

© 2013 Information Processing Society of Japan

8.1 知見に関する考察

8.1.1 あらかじめ録音された会議音声に無音区間を挿入することが NS や NNS に与える影響

あらかじめ録音された会議音声に無音区間を挿入することが，非母語者に対して効果的に働いた理由として，以下の 2 つの要因が考えられる．まず，無音区間が入ることによって，音声情報を処理するために時間をかけられるようになり，理解に必要な負荷を下げることができたと考えられる．次に，発話間に無音区間を挿入することによって，発話のオーバラップが減少したことがあげられる．これによって，会話音声が全体的に明瞭になり，オーバラップによって聞き取りにくくなった部分を補うための負荷が軽減されたと考えられる．

このように無音区間の挿入によって，非母語者の理解度を向上させることはできたが，母語者に対しては同様の効果は見られなかった．母語者の理解度がつねに高く，負担に対する自己評価もつねに低かったことから，母語者は会議音声を理解することに関して，もともと負荷が小さかったと推測される．通常の会議音声であっても十分に理解できた母語者にとって，無音区間の挿入は天井効果により大きな効果を発揮できなかったと思われる．

8.1.2 リアルタイムの会議に遅延を挿入することが NS と NNS に与える影響

2 つめの実験により，遅延の挿入が母語者には逆効果であることが示された．この結果は通信遅延が遠隔対話者に悪影響を及ぼすという従来研究の知見と合致している．我々の実験では，遅延の増加にともなって発話開始時のオーバラップの回数が増加し，トランスクリプト 1 で見たような不自然に長い沈黙が観測された．こうした問題によって話者交代の回数が減少し，不満が増大したと考えられる．

非母語者にとっては，付加された遅延によって，有益な効果と弊害の両方が観測された．有益な点は，無音区間を生み出すことに成功し（図 9），これによって非母語者が余裕を持って発話を理解したり次に何を話すかを考えることができるようになったことである．実際，この無音区間によって，非母語者は母語者が話し始める前に発話を開始することができる機会が増え，自発的発話が増えることになった（表 4）．ただし，遅延が長くなると，発話のオーバラップが長くなってしまうケースが増え，有益な効果を相殺してしまう結果となった．

8.1.3 リアルタイム会議において無音区間が NNS の理解に要する負担を軽減できなかった理由

あらかじめ録音された会議音声に無音区間が挿入された場合，非母語者は比較的少ない負担で母語者の会話を聞き取ることができた．しかし，リアルタイムの音声会議においては，この程度の無音区間では理解に要する負担を軽減させるには不十分だったようである[*2]．むしろ，発話オーバラップの増加によって，会話を聞き取るためにさらなる注意と努力が必要になったと考えられる．

0.2 秒の遅延挿入時にもオーバラップ長が増加していたにもかかわらず（図 10），非母語者の理解度に対する影響がほとんど見られず，0.4 秒の遅延挿入時に影響が大きくなった要因について考察する．0.2 秒の遅延挿入時と 0.4 秒の遅延挿入時のオーバラップを比較すると，オーバラップの特徴に違いが見られることが分かった．挿入した遅延が 0.2 秒の場合，オーバラップ中に発話される単語数は平均して 1，2 語であり，遅延が 0.4 秒の場合には 2，3 単語であった．この違いは微小であるように見えるが，被験者のオーバラップに対する反応はまったく異なっていた．0.2 秒の場合には，被験者が同時に発話した場合，オーバラップに気づくタイミングが比較的早く，気づくとすぐに発話を止め，言い直しをすることが多かった．それに対して 0.4 秒の場合には，文の途中で割り込まれることが多く，そのような場合には，彼らは発話を継続したり，単に中断してしまって言い直さなかったりするケースが多かった．以下に，そうした例を引用する．

トランスクリプト 2（遅延時間 0.2 秒）：

NS1: Ok=
NS2: =Yeah (.)
NS1: So [anyway]
NS3: [I:::] I need I need the brandy

トランスクリプト 3（遅延時間 0.4 秒）：

NS1: my argument towards [the knife is tha:::t]
NS2: [the thing that kill you]
the fastest is dehydration (0.8) so:::
NS1: yeah (.)
NS3: water or dead

このようにオーバラップが発生した後の対処方法の違いによって，聞き手の負担に違いが生じたと考えられる．

8.2 設計指針

本研究で得られた知見と前述の考察によって，非母語者を支援する音声会議システムに対する指針が得られる．無音区間の挿入は非母語者にとって有効であるが，発話のオーバラップが生じてしまうとむしろ問題点の方が大きくなってしまう．したがって，オーバラップを増加させずに無音区間を挿入することができれば，音声対話における有効性が増すと考えられる．

そのような方法として，たとえば push-to-talk（PTT）方式が考えられる．PTT は通常，音声通信に参加する人数が多い場合に，ハウリングや背景の雑音を低減させるためによく利用されるが，注目すべきは，PTT を用いることによって参加者自身がオーバラップを減少させるような会話方式をとっているという点である [18]．人工的な遅延

*2 非母語者の理解度が実際に向上したか否かについては，本実験からは分からない．

© 2013 Information Processing Society of Japan

と PTT 機能を組み合わせることによって，遅延時間を比較的長くしても，NS-NSS 音声通信におけるオーバラップを最小限にし，非母語者に対する有効性を向上させられると期待できる．

また，人工遅延を挿入する代わりに，母語者間だけ話速をわずかに遅くする方法も考えられる．非母語者は通常の速度で発話を聞くため，母語者よりも早く発話を聞き終わることになり，結果として次の発話との間に無音区間が挿入されることになる．なお，この方法では，遅延はほとんど存在しないため，複数の母語者が互いの発話に気づかずに話し続ける現象（すなわち発話開始時のオーバラップ）を回避することができると考えられる．

9. おわりに

母語が同じ話者同士の対話に対して，通信遅延がいかに問題であるかということを示している研究は多い．母語の異なる者同士の対話に対して，通信遅延が与える影響に関する研究はほとんど見られない．これに対して本研究では，やり方によっては，遅延が非母語者にとって有益になりうることを示した．また，母語者と非母語者の多人数対話に，遅延が与える影響に関する知見を提供することができた．具体的には，あらかじめ録音された多人数音声対話に対して遅延（無音区間）が挿入されると，非母語者はより正確にその内容を理解することができた．リアルタイムの音声会議においては，短い遅延（0.2 秒）によって，非母語者が会話で自主的に発話する回数を増やすこともできた．しかし，遅延が長くなると（0.4 秒），非母語者に対する有効性よりも，たとえば発話のオーバラップ長の増加やフラストレーションの増加など，弊害の方が上回ってしまった．

本研究からは，非母語者のリアルタイム会議での理解度が 0.2 秒の遅延挿入時に向上したか否かは分からない．しかし，0.2 秒の遅延挿入時に非母語者の自発的な発話数が増加し，合意形成内容に対する満足度も向上したことから，非母語者にとって有効であったと判断できるだろう．

本研究の提案内容が本当に実用的であるか否かは議論の分かれるところであろう．今後は，非母語者の会議参加を支援する他の方法を試すと同時に，他の言語による実験も行いたいと考えている．たとえば，日本人にとって，英語は文法構造が異なるために，英語による会話に参加することが難しい可能性がある [21]．被験者の母語と使用される言語の組合せによっては，異なる結果が得られるかもしれない．

謝辞 本研究は日本学術振興会科学研究費基盤研究（S）（24220002，平成 24 年度～28 年度）の補助を受け実施したものである．ここに深謝の意を表す．

参考文献

[1] Broadbent, D.: *Perception and communication*, Pergamon Press, London (1958).

[2] Clark, H.H. and Schaefer, E.F.: Collaborating on contributions to conversations, *Language and Cognitive Processes*, Vol.2, No.1, pp.19–41 (1987).

[3] Ding, X., Erickson, T., Kellogg, W.A., Levy, S., Christensen, J., Sussman, J., Wolf, T.V. and Bennett, W.E.: An Empirical Study of the Use of Visually Enhanced VoIP Audio Conferencing: The Case of IEAC, *Proc. CHI '07*, pp.1019–1028 (2007).

[4] DiMicco, J.M., Hollenbach, K. and Bender, W.: Using visualizations to review a group's interaction dynamics, *Proc. CHI '06*, pp.706–711 (2006).

[5] Ezzatian, P., Avivi, M. and Schneider, B.: Do non-native listeners benefit as much as native listeners from spatial cues that release speech from masking?, *Speech Communication*, Vol.52, No.11-12, pp.919–929 (2010).

[6] Gutwin, C.: The effects of network delay on group work in shared workspaces, *Proc. ECSCW '01*, pp.299–318 (2001).

[7] Inkepen, K., Hegde, R., Czerwinski, M. and Zhang, Z.: Exploring spatialized audio & video for distributed conversations, *Proc. CSCW '10*, pp.95–98 (2010).

[8] Junuzovic, S., Inkepen, K., Hegde, R., Zhang, Z., Tang, J. and Brooks, C.: What did I miss? In-Meeting Review using Multimodal Accelerated Instant Replay (AIR) Conferencing, *Proc CHI '11*, pp.513–522 (2011).

[9] Kilgore, R., Chignell, M. and Smith, P.: Spatialized Audioconferencing: What are the Benefits?, *Proc. CASCON '03*, pp.135–144 (2003).

[10] Krauss, R.M. and Bricker, P.D.: Effects of transmission delay and access delay on the efficiency of verbal communication, *Journal of the Acoustical Society of America*, Vol.41, No.2, pp.286–292 (1967).

[11] Lafferty, J.C., Eady, P.M. and Elmers, J.: The Desert Survival Problem, Plymouth, Michigan, Experimental Learning Methods (1974).

[12] Li, H.Z., Yum, Y., Yates, R., Aguilera, L., Mao, Y. and Zheng, Y.: Interruption and Involvement in Discourse: Can Intercultural Interlocutors be Trained?, *Journal of Intercultural Communication Research*, Vol.34, No.4, pp.233–254 (2005).

[13] Luisa, M., Lecumberri, G., Cooke, M. and Culter, A.: Non-native speech perception in adverse conditions: A Review, *Speech Communication*, Vol.52, No.11-12, pp.864–886 (2010).

[14] Macnamara, J. and Kushnir, S.L.: Linguistic independence of bilinguals: The input switch, *Journal of Verbal Learning and Verbal Behavior*, Vol.10, No.5, pp.480–487 (1971).

[15] Mayo, L., Florentine, M. and Buus, S.: Age of Second-language acquisition and perception of speech in noise, *Journal of Speech, Language, and Hearing Research*, Vol.40, No.3, pp.686–693 (1997).

[16] Nabelek, A.K. and Donahue, A.M.: Perception of consonants in reverberation by native and non-native listeners, *Journal of Acoustical Society of America*, Vol.75, No.2, pp.632–634 (1984).

[17] Rogers, C., Lister, J., Febo, D., Besing, J. and Abrams, H.: Effects of bilingualism, noise and reverberation on speech perception by listeners with normal hearing, *Applied Psycholinguistics*, Vol.27, No.3, pp.465–485 (2006).

[18] Sawhney, N. and Schmandt, C.: Nomadic radio: Speech and audio interaction for contextual messaging in no-

© 2013 Information Processing Society of Japan

madic environments, *TOCHI*, Vol.7, No.3, pp.353–383 (2000).
[19] Short, J., Williams, E. and Christie, B.: *The Social Psychology of Telecommunications*, John Wiley, London (1976).
[20] Takano, Y. and Noda, A.: A temporary decline of thinking ability during foreign language processing, *Journal of Cross-Cultural Psychology*, Vol.24, No.4, pp.445–462 (1993).
[21] Tanaka, H.: Turn Projection in Japanese Talk-in-Interaction, *Research on Language and Social Interaction*, Vol.33, No.1, pp.1–38 (2009).
[22] Tang, J.C. and Isaacs, E.A.: *Why do users like video? Studies on multimedia-supported collaboration*, Sun Microsystems Laboratories, Inc., Mountain View, CA (1992).
[23] Veinott, E., Olson, J., Olson, G. and Fu, X.: Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other, *Proc. CHI '99*, pp.302–309 (1999).
[24] Vertgaal, R., Van der Veer, G.C. and Vons, H.: Effects of Gaze on Multiparty Mediated Communication, *Proc. Graphic Interface '00*, pp.95–102, Morgan Kaufmann Publishers (2000).
[25] Yamashita, N., Echenique, A., Ishida, T. and Hautasaari, A.: Lost in Transmittance: How Transmission Lags Enhances and Deteriorates Multilingual Collaboration, *Proc. CSCW*, pp.923–934 (2013).
[26] Yankelovich, N., Kaplan, J., Provino, J., Wessler, M. and DiMicco, J.M.: Improving Audio Conferencing: Are Two Ears Better than One?, *Proc. CSCW '06*, pp.333–342 (2006).
[27] Yankelovich, N., McGinn, J., Wessler, M., Kaplan, J., Provino, J. and Fox, H.: Private Communications in Public Meetings, *Proc. CHI '05*, pp.1873–1876 (2005).
[28] Yankelovich, N., Walker, W., Roberts, P., Wessler, M., Kaplan, J. and Provino, J.: Meeting Central: Making Distributed Meetings More Effective, *Proc. CSCW '04*, pp.419–428 (2004).
[29] 話速変換音声聴取に音声伸長量とポーズ長が及ぼす影響，電子情報通信学会技術研究報告，Vol.104, No.695, pp.43–48, SP2004-161 (2004).

山下 直美 （正会員）

1999 年京都大学工学部情報工学科卒業．2001 年京都大学大学院情報学研究科数理工学専攻修士課程修了．同年日本電信電話（株）コミュニケーション科学基礎研究所入所．博士（情報学）．CSCW，HCI の研究に従事．

エチェニーケ アンディ

カリフォルニア大学アーバイン校博士課程在籍中．異文化コラボレーション支援の研究に興味を持つ．

葛岡 英明 （正会員）

1992 年東京大学大学院工学系研究科情報工学専攻博士課程修了．博士（工学）．同年筑波大学構造工学系講師．現在，筑波大学大学院システム情報工学研究科教授．CSCW，HRI，HCI の研究に従事．

石田 亨 （フェロー）

1976 年京都大学工学部情報工学科卒業，1978 年京都大学大学院修士課程修了．同年日本電信電話公社電気通信研究所入所．ミュンヘン工科大学，パリ第六大学，メリーランド大学客員教授等経験．工学博士．IEEE フェロー．現在，京都大学大学院情報学研究科社会情報学専攻教授，上海交通大学客員教授．自律エージェントとマルチエージェントシステム，セマンティック Web 技術に取り組む．デジタルシティ，異文化コラボレーションプロジェクトを推進．

ハウタサーリ アリ

京都大学大学院情報学研究科博士課程在籍中．異文化コラボレーション支援の研究に興味を持つ．

© 2013 Information Processing Society of Japan