異言語ビデオ会議におけるNSとNNSの協調学習の数量的表示と判定基準
-アジアン・ミス・コミュニケーション方略分析を通して-

*砂岡和子　**俞敬松　***高媛媛
*早稲田大学政治経済学術院　ksunaoka@waseda.jp
**北京大学軟件与微電子学院　yjs@ss.pku.edu.cn
***北京大学対外漢語教育学院　highland2000@163.com

概要　対面オンライン遠隔会議で、母語話者(NS)と非母語話者(NNS)の(L2)学習者が発揮する協調学習の特徴を、構築中の「アジアン・ミスコミュニケーション・コーパス」の分析例を通して具体的に報告する。日本語会議はポーズとフィラーの使用に関し、中国語会議は「発話単位理解度」と「コミュニケーション達成度」について、それぞれNSとNNSの協調的談話方略を定量的に分析した。得られた結果を元に、評価の可視化と半自動ラベリングの手法を提案し、外国語教育や異文化学習への応用を検討する。
キーワード　自然発話、協調学習、談話方略、ミス・コミュニケーション・コーパス

The Evaluation of the NS and NNS' Collaborative learning ability By the Multilingual videoconferencing
- Analysis of the Communication Strategies by Asian Miscommunications Corpus –

*Kazuko Sunaoka　**Yu Jingsong　***Gao Yuanyuan

Keywords: Spoken Language，Collaborative Learning，Communication Strategies，Miscommunications Corpus

1. 日中異文化コミュニケーション学習支援

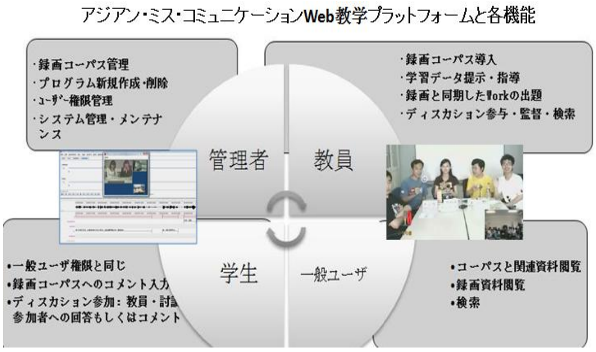
国際交流の拡大と情報通信技術の高度化に伴い、ICTを利用した異文化コラボレーションの機会が増えている。早稲田大学では過去8年間、早稲田・慶応・北京大・清華大・台湾師範大[1]・韓国高麗大の6校間でビデオ会議を開催してきた。アジア6大学の学生たちがそれぞれの教室からTV画面を通じて対面し、時事問題や身近な関心事をテーマに討論を行う。会議用語は原則的に隔週、日本語と中国語を交互に使用する。母語での討論は心理的緊張感を緩和し、非母語話者にとっても次回の発言に必要な情報を吸収できる。参加者の大半が大学入学後に学習を開始した第二外国語であるため、高く語学力は望めない。異文化コミュニケーションでは、多数の参加者が協調的に学び合う過程が相互理解を促し（三宅 2009)、学習支援も母語話者に近い正確さを目指す伝統的外国語教育とは観点が異なる。

異文化間コミュニケーションには言語スキル以外に、課題遂行能力や状況への柔軟な対応力、対人関係や感情処理など、異文化間ソーシャル・スキルが求められる（Yusuke Kondo他2008a)。アジア6大学TV会議の参加学生たちも、さまざまなソーシャル・スキルを動員して言語の壁を乗り越え交流を試みるが、日中間の情報格差を短時間で埋めることは容易ではない（砂岡他2008)。

異文化コミュニケーションは交流の支援法と到達度判定基準が明確ではなく、遠隔授業を授業科目として導入するさい、シラバスや教材作成に困難をきたす。「アジアン・ミスコミュニケーション・コーパス」は異文化交流障害の緩和支援を目的に、過去蓄積してきた日本語と中国語の学生会議録画約100本を資源とするアノテーション付き動画コーパスである。コーパスの閲覧・コメント・検索可能な教育・研究用Web教学プラットホームも構築中で[図 1]、個人情報保護と版

[1] 2008年10月から淡江大学に交代。

権処理を行い、インターネットで公開を目指す[2]。一過性の対面会議を振り返り、交流障害要因を可視化して相互にコメントの書き込み・閲覧が可能な音声画像データは、自立型対話実践の有力支援ツールとして期待できる(大塚他 2009)。



［図 1］

2. コミュニケーション障害の分析環境

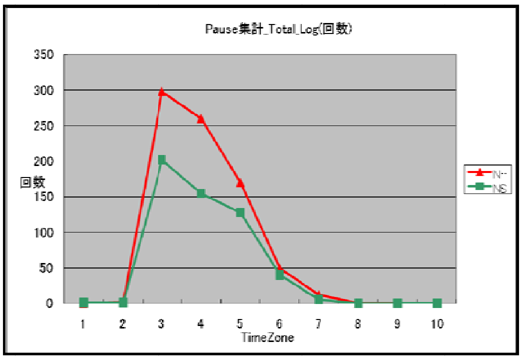
異文化コミュニケーションで発生する障害を、言語スキルとソーシャル・スキルに 2 分し、さらにそれぞれ下位項目を策定し、録画に評価コメントを注記してゆく。教員や母語話者による一方的な評価ではなく、討論の参加者が相互評価できるよう、ELAN3をカスタマイズし、動画・音声・テキストデータの検索閲覧と、多言語注記ができるアノテーション環境を構築した。テキストのみの転記に比べ、ELAN では動画シーンと同期させコミュニケーションの経緯を可視化できる。アジア学生会議の録画資料を用い、現在までに母語話者(NS)と非母語話者(NNS)のポーズやフィラー〔filler〕の出現傾向、ならびに協調学習に関して調査報告を行ってきた（砂岡他 2007,2008a)）。語彙や語法情報と異なり、しぐさや態度、沈黙や言いよどみなど、パラ言語的要素は非母語話者にとって把握が難しい。ポーズやフィラーは応答詞と共に、NS と NNS の協調的談話方略に大きな作用を及している。自然発話を特徴づけるこれら副言語的、非言語的要素の運用実態を、ELAN で直感的、かつ物理的に情報分析できる。以下、実例で、本 Web 教学プラットホームの利用法を示いてゆく。

[2] http://www.f.waseda.jp/ksunaoka/enkaku/index.html
[3] Language Archiving Technology（LAT）
http://www.lat-mpi.eu/tools/elan/

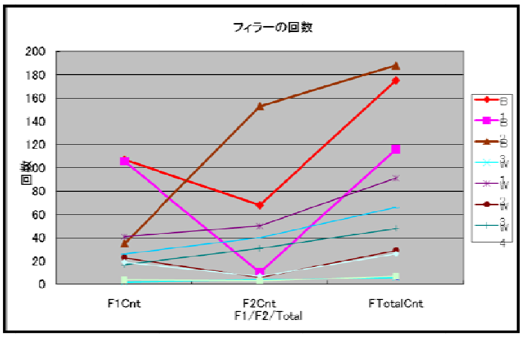
3. NS と NNS のポーズ、フィラーの使用特徴

10 分弱の日本語会議発言中に出現する NS8 名と、NNS3 名のポーズ回数を、話者ごとポーズ長別に集計し、対数表示した[図 2]。NNS は NS に比べポーズ使用回数が多いが、ポーズ長には双方大きな違いがない(横軸ポーズ尺度 3~5[200ms~800ms]に収束)。発話 10 秒当たりのポーズ出現回数のデータも同様の傾向を示しており、録画の聴覚的印象とも一致する。この結果を日本語学習に生かすと、NNS にもポーズは必要だが、一発話内に頻繁にポーズを挟まないよう注意すべきだ。

［図 2］



同様に NS8 名と NNS3 名のフィラー回数を話者ごとに集計した[図 3]。フィラーは属性を[F1；音声的フィラー][F2;語彙的フィラー]の 2 種類に分けた。NNS は NS に比べフィラー使用量が総じて多く、B3 を除いた NNS2 名は、音声的なフィラー[F1]の使用が語彙的フィラー[F2]より著しく多い。対して NS8 名は[F1]と[F2] のフィラーをバランスよく使用している。これから NNSL2 学習者は、音声的フィラーの乱用を控え、語彙的フィラー運用能力を高めるとよい。



［図 3］

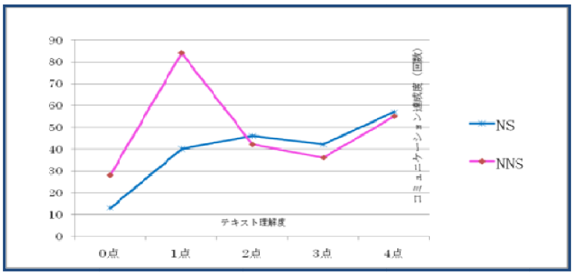
4. NS と NNS の協調的コミュニケーション方略

次に、NS と NNS がオンライン対話で発揮する第二外

国語習得(L2)の特徴を、協調的コミュニケーション方略に焦点を当て分析を行った。討論言語は中国語で、参加者は NNS 3 名・NS 5 名（全員中国語非専攻で初級レベル)。発話を最小発話単位「チャンク」に分け、全 443 チャンクの有音声区間テキストを得た。これを「テキスト内での理解度(A)」と、実際の対話中における「コミュニケーション達成度(B)」の 2 種のラベルを付け、両者の関係を定量的に観察した。採点は各 4 段階とし、中国人語学教員が手作業でラベル付けを行った。

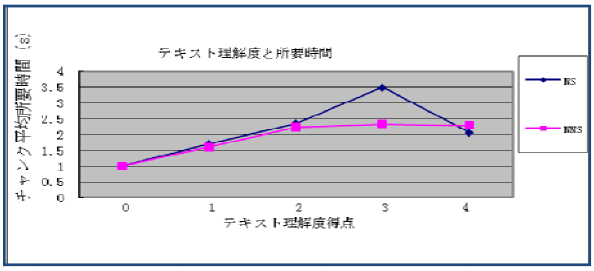
4. 1. NS 対 NNS 発話単位の情報差

［図 4］は NS 対 NNS のテキスト理解度（横軸）とコミュニケーション達成度（縦軸）の相関関係を表す。母語話者 NS のテキスト理解度とコミュニケーション達成度はほぼ正比例し、高得点(2-4 点)になるほど NNS を凌ぐ。対して NNS はテキスト理解度、コミュニケーション達成度がともに低い 1 点(大半が語断片)に発話数が集中し、2-4 点の高得点群と 0 点(大半が笑い声やフィラーなど非言語要素）が均等に分布する。有効なコミュニケーションチャンクをうまく産出できない NNS の様子が一目瞭然である。

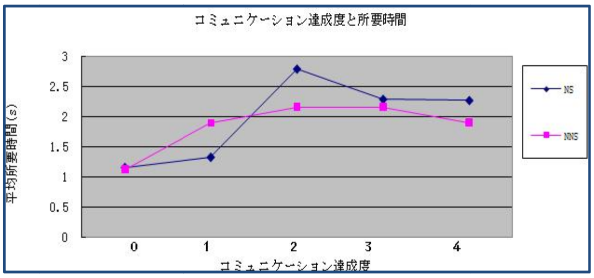


［図 5］

時間との相関でも、NS はテキスト理解度(3 点)、コミュニケーション達成度(2 点)を頂点とする発話分布に対し、NNS は低得点群で NS に比肩するが、大半が語断片やポーズで、情報伝達に貢献しない［図 4］［図 5］。母語話者 NS の得点ピークが満点(4 点)ではなく、コミュニケーション達成度の頂点が、テキスト理解度のピークより低得点に移動 (3→2 点) している。NS が母語を使って NNS と対話するさい、交流達成の任務を優先し、テキスト単位で見ると理解度が低いチャンクを多用すると解釈できる。テキスト規範性を犠牲にしても、コミュニケーション達成度を優先する協調的談話方略が可視化された。

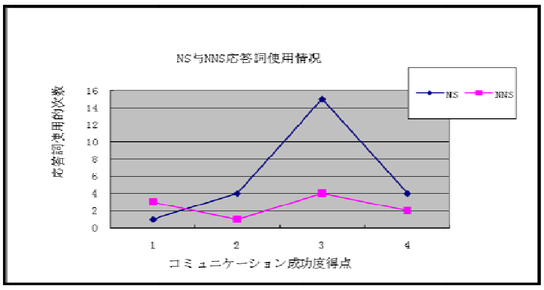


［図 4］



［図 5］

［図 6］で、NS の応答詞はコミュニケーション達成得点 3 に集中して出現する。NS は応答詞の挿入により、テキスト理解度がやや低めのチャンクを自然に繋ぎ、NNS との協調的談話を達成する。対して NNS は応答詞を使用せずにテキスト理解度の向上を図ろうとするが、発話総数や時間の浪費のわりに、コミュニケーション達成度が低い結果となっている。



［図 8］

5. NNS のコミュニケーション方略

以上の結果から、非母語話者 NNS が NS の母語を使ってコミュニケーション達成度を向上するには、母語話者と同じ満点を目指す必要はない。多くの場合、相手の NS の協調的談話方略により、双方の交流レベルが接近する。肝要なのは、ポーズや笑い声、フィラー、語断片など、発話中の非言語要素をより言語的質量の

高い内容に変換するスキルである。接続詞など応答詞の使用は、コミュニケーション達成度を上げるのに有効であり、難しい語彙を使うよりは、使い慣れた平易な語句を淀みなく発話し、平均2秒程度の短い発話を往還すれば、コミュニケーション達成度が向上する。

6. 評価の可視化と半自動ラベリング

アジア学生会議のように、多人数・異文化インタラクション活動では､階層的な教育コンテンツを共有し、その上に付加価値を創造し続ける自律的コミュニティとして存続することが望ましい。そのためには録画分析と各種アノテーション処理が簡便に行えることが条件となる。上記分析は多く手作業によるが（例：録画データ書き起こし、チャンクのラベル付け、フィラーへのアノテーション)､一部はプログラムによる半自動処理を行った（例：録画音声声無音区間の自動分割ツール（認識率70%-90%)、ELAN機能によるテキスト時間情報の抽出])。Matlabはまだ高価で使用難度も高いが、ELANはフリーソフトで扱いやすく、機能も十善である。ELANへの情報入力が正確であれば、データは自動的に吐き出され、それを定量分析し、かつ結果をグラフなどに可視化する作業はさほど難しくはない。Web教学プラットホームには各種アプリケーションの使用マニュアルと、研究事例や分析方法を提示し、外国語教育や異文化学習への利用普及を図ることができる。下記、チャンクのラベリングのうち「テキスト理解度」は4段階評価データをコーパス化すれば、コンピュータによる照応によってほぼ自動化が可能のはずである。「コミュニケーション達成度」に関しては、明示的文法規則を欠く中国語の自動判定は困難が予想される。

A　テキスト理解度

4点　テキスト単独で完全に理解できるフレイズ

3点　一つ以上完全な句を含む

2点　不完全な句

1点　単語だけ、または非動詞性語句

0点　語断片、語素など

B　コミュニケーション達成度

4点　完全にコミュニケーション達成

3点　やや達成（完全ではないが、ほぼ達成）

2点　前後の文脈がないと意味不明、或いは二義文

1点　前後文脈情報を借りても、理解困難

0点　その他、語断片、語素、ポーズなど

なお本分析では併せて使用語彙と句型の特色についても初歩的調査を行ったが、紙幅が尽きたため割愛した。詳細は(砂岡他2008b)を参照されたい。

謝辞　本研究は平成19-21年度文部科学省科学研究費補助金〔基盤（B）課題番号：19320087研究代表者砂岡和子〕の助成で進行中の成果の一部である。共同研究者の俞敬松（北京大学）のほか、高媛媛、王瑩（北京大学)、保坂敏子(日本大学)、河内彩香（早稲田大学)、山口真紀(東京工業大学)の協力を得、一部データ分析は(株)アイアール・アルトに委託した。

参考文献

[1] 砂岡和子，俞敬松，アジアン・ミスコミュニケーションコーパスの構築，人工知能学会第53回言語・音声理解と対話処理研究会予稿集2008a

[2] 砂岡和子、Yu Jingsong，日中ビデオ会議にみる共同学習-アジアン・ミス・コミュニケーションコーパスの開発-電子情報通信学会,思考と言語研究会，2008b

[3]Yusuke Kondo 1, Aya Kitagawa 2, and Michiko Nakano,Second language speech: subjective evaluation and objective measures, LangSpSciResLabWSNo.2，2008

[4] 三宅なほみ, 多人数インタラクションを活用した学習とその支援，人工知能学会誌Vol.24, 2009

[5] 大塚裕子、森本郁代、水上悦雄、富田英司、山内保典、柏岡秀紀「科学技術コミュニケーションにおける対話のデザイン」人工知能学会Vol.24, 2009

[6] 国立国語研究所・情報通信研究機構『日本語話し言葉コーパス（CSJ)』2004