英語能力テストに関する研究

面接方法が発話に与える影響

―ロールプレイを用いた個別面接方式とペア面接方式の比較―

栃木県／栃木県立鹿沼東高等学校　教諭　川島 智幸
申請時：栃木県／栃木県立小山高等学校　教諭

概要 本研究では，ペア面接方式の有効性を検証するため，英検準2級の2次面接と全国模擬試験により測定した英語能力の個人差が，ロールプレイにおける発話にどのような影響を及ぼすかを調べた。さらに生徒同士のペア面接方式と，教師との個別面接方式の2つの方式でロールプレイを行い，面接方式の違いが発話に及ぼす影響を分析した。実験では，公立高校2年生8人が行った4種類のロールプレイでの発話について，発話量，複雑さ，正確さ，流暢さ，発話の機能への影響の有無を調べた。その結果，面接方法の違いが流暢さへ及ぼす影響のみが有意となり，ペア面接において英語能力の個人差が，発話量や正確さ，複雑さ，流暢さ，発話の機能に影響しないことを確認した。また面接方式の違いが，発話量や正確さ，複雑さに影響しないことも明らかになった。これらの結果は，試験的ではあるもののペア面接の有効性が支持されたことを意味する。

1　はじめに

「英語が使える日本人」の育成は，今や国家的課題である。そのためには，実際に英語を使って何ができるかを測ることが不可欠となる。しかし，スピーキングテストを実施し発話を直接観察しようとすると，膨大な時間と労力を要する。こうした状況の下注目すべきは，イギリスのケンブリッジ大学海外試験評議会（UCLES）が英語検定試験で用いるペア面接方式である。試験時間の短縮だけでなく，対等な力関係の下でより自然なコミュニケーションを引き出せるとの期待が寄せられている。しかし同時に，能力の違う者を組ませると，受験者が十分に力を発揮できないのではという懸念もある。

UCLES は，1980年代からペア面接の試行を始め，1991年の Certificate of Advanced English（CAE）を皮切りに，日本でケンブリッジ英検の名称で知られる検定試験を順次ペア面接方式に切り替えた。現在は一番上級の Certificate of Proficiency in English（CPE）を除く4レベル，易しい順に Key English Test（KET），Preliminary English Test（PET），First Certificate in English（FCE），Certificate of Advanced English（CAE）の各テストで，ペア面接方式を実施している（Saville and Hargreaves, 1999; Langham, 2001）。しかし残念ながら UCLES は，予想されるペア面接への懸念を払拭するのに十分な研究結果を公表していない。それどころか UCLES でテスト開発に携わる Saville and Hargreaves（1999）は，ペアを組む受験生の能力や年齢，性別，国籍などが評価に影響する可能性を示唆した。

これまで個人面接において発話に影響しうる様々な要因が研究されてきたが，ペア面接自体がまだ一般的でないため，能力差が発話に及ぼす影響に言及した研究はごくわずかしかない。しかも英語学習者を対象とした研究は乏しく，能力差が発話に与える影響はいまだ明らかになっていない。そこで本研究では，ペア面接方式の有効性を検証するため，英検準2級の2次面接と全国模擬試験により測定した英語能力の個人差が，ロールプレイにおける発話にどのような影響を及ぼすかを分析した。さらに生徒同士のペア面接方式と，教師との個別面接方式の2つの方式でロールプレイを行い，面接方式の違いが発話に及ぼす影響を分析した。

以下では，まず面接方式ごとに主要なインタビュ

ーテストを分類し，簡単にそれぞれの面接方法の長短をまとめる。その後，テストとしてのロールプレイの功罪を論じ，様々な受験者要因と発話との関係に関する先行研究を概観する。

2 先行研究

2.1 面接方式

面接方式には，個別面接方式，ペア面接方式，テープ面接方式の3つがある。個別面接方式は，面接官1人に受験者1人で行う。この方式をとる主な面接に，ACTFL（全米外国語教育協会）が37言語で実施する Oral Proficiency Interview（OPI），Educational Testing Service（ETS）が，Test of English for International Communication（TOEIC）の筆記試験で730点以上の得点者を対象に実施する Language Proficiency Interview（LPI），UCLES が実施し，イギリスやオーストラリアなどで広く英語力の証明として利用される The International English Language Testing System（IELTS）の Speaking Test がある。また国内で開発されたものには，ACTFL と株式会社アルクが OPI を日本人向けに改良した Standard Speaking Test（SST）や，全国外国語教育振興協会の Standard Interview Test of English（SITE）などがある。

2番目のペア面接方式は，受験生2人を同時に面接する方法で，前述の通り UCLES が実施するケンブリッジ英検では，一番上級の CPE を除く4レベルのテストでこのペア面接方式が取り入れられている。厳密に言うと UCLES が実施するペア面接方式は，面接官と別に，もう1人評価官が同席し，面接官と同時に評価に当たる方法をとっている。また受験の手引きによれば，受験者が奇数の場合，最後の面接は受験者3人1組で行うと明記してある（UCLES, 2004）。ペア面接方式はケンブリッジ英検以外では，ハンガリーで2005年より高校卒業試験に導入されることになっている（Csepes, 2001）。本研究では，UCLES が実施している複数の評価者が面接に立ち会うやり方ではなく，単に1人の面接官が2人の受験者を面接する方法をペア面接方式と呼ぶことにする。

3番目のテープ面接方式は，ETS が Test of English as a Foreign Language（TOEFL）の受験者に希望で実施している Test of Spoken English（TSE）で使われ，受験者はテープから流れる問いに答え，その答えが自動的にテープに録音されるやり方である。

次に，個別面接方式とペア面接方式についてこれまで言われた指摘を整理してみよう。OPI をはじめとする個別面接方式では，常に主導権を握る面接官とそうでない受験者の間に不均衡な力関係が存在する。このため，面接官が面接を統制しやすく，面接の均等性を保ちやすい利点がある（Kitao and Kitao, 1996）。しかし同時に，現実の言語使用場面を広くカバーすることは難しく，どうしても面接官，受験者間の初対面のフォーマルなやり取りに終始してしまうとも言われている（Raffaldini, 1988 cited in 馬場, 1997, p.128）。例えば大人数の教室でよく見られる，教師の問いかけ（initiation）に生徒が答え（response），次に教師がそれに対してコメント（follow up）するようなやり取りが個別面接では多く見られる（Egyud and Glover, 2001; Hughes, 1989）。そして最悪の場合，McCarthy（1991, pp.122-3）が指摘するように，受験者は面接官からの問いかけに答えることしかやらなくなる。

これに対しペア面接は，受験者のストレスを軽減し（Ikeda, 1998），個別面接に比べ双方向の釣り合いのとれた発話を生む（May, 2002）とされる。また受験者同士や，面接官と受験者など多くの形態でのやり取りが可能となる（Saville and Hargreaves, 1999）。

しかしながらペア面接の問題点として，受験者の発話がペアを組む相手の発話に影響される可能性が指摘されている。例えば自己主張の強い無神経な受験者だった場合，面接を牛耳り，もう1人の受験者が正当に力を発揮する機会を奪ってしまうかもしれない（Hughes, 1989, p.105）。また Foot（1999）は，ペア面接により受験者のストレスを軽減することができても，リラックスしたからといって必ずしも良い成績につながるわけではないと，ペア面接の有効性に疑問を呈した。

さらに本稿の研究課題でもある受験者の能力との関連で，ペア面接がどのような受験者に有効かについてはまだ検証の余地がある。Foot（1999）は，ペア面接は英語力の高い受験者に適するが，英語力の低い者には適当でないと述べた。一方 Egyud and Glover（2001）は，ハンガリーの職業高校で行った実験から，ペア面接は能力の高い受験者同様能力の低い者にも適していると結論づけた。

2.2　ロールプレイ

受験者から引き出される発話が，実際に様々な場面で交わされている会話とかけ離れるというインタビューテストの欠点を埋めるため，ロールプレイが面接のタスクに利用されている。ACTFL OPI をはじめ，TOEIC LPI，TOEFL TSE，SST，SITE などのインタビューテストで，ロールプレイを取り入れている。ロールプレイは発話を現実のものに近づけ（Shohamy, 1988, p.172, cited in Salaberry, 2000），受験者がより対等な立場で会話に参加することを可能にする（Kormos, 1999）。更に Evans, Gaudin and Raveau（1987）は，ロールプレイを通して受験者の反応の即時性と，やり取りを維持する能力を測れる点を長所として挙げた。

また学習用タスクとして，ロールプレイは，予測不可能な実際のコミュニケーションに対応する力を養う（Salies, 1995; Ladousee, 1987, cited in Sato, 2001），発話の適切さを学ぶ（Salies, 1995），自信をつける（Salies, 1995; Maxwell, 1997），即興でのやり取りや話題の転換といった実際の会話の特徴に慣れる（Di Pietro, 1982, cited in Sato, 2001; Kormos, 1999），語彙を増やす（Shoemaker, 1991, cited in Sato, 2001）などの利点が言われている。

ロールプレイと個別面接で得られる発話を比較した，興味深い研究がある。Lynch（1999）は，日本人32人に個別面接とロールプレイを続けて実施し，発話量と発話の機能を比べた。そしてロールプレイは，発話量と発話機能の種類の豊富さで個別面接を上回るということがわかった。いかに個別面接の発話が機能面で限られているかの一例として Lynch（1999）は，合わせて5時間にも及ぶ個別面接の中で，質問をした受験者はわずか4人しかいなかったという。これ以外に，忠告や命令という発話の機能も，ロールプレイでのみ見られた発話機能だと言う。更に Lynch（1999）は，個別面接とロールプレイをそれぞれ7項目で評価し，その評定を比較し，個別面接とロールプレイの評定の間には高い相関（相関係数0.79）があることを確認した。以上のことから，ロールプレイは従来のインタビューテストにはない多くの長所を持つ，妥当性の高いタスクだと言えるだろう。

しかし最後に，ロールプレイテストの短所も忘れてはならない。馬場（1997, p.119）は，外向的な受験者が高い評価を受ける可能性があることと，与えられた役割に対する親密度がパフォーマンスに反映されやすい点を指摘している。またロールプレイは他のスピーキングタスクと比較すると，発話の複雑さや正確さが劣るという研究もある。Halleck（1995）は，107人の中国人英語学習者に，ロールプレイと描写・叙述の課題を与え，課題ごとに得られた発話の T-unit の平均語数や，誤りのない T-unit の平均語数を比較した。その結果，描写・叙述の課題の方がロールプレイよりいずれの平均語数をも有意に上回った。Halleck（1995）は，ロールプレイは常に相手の言うことに正しく反応しなくてはならないプレッシャーが伴い，十分に実力を出せなかったのではと，その理由を推察した。

2.3　受験者要因と発話との関係

2.3.1　親密さ

親密さの影響については研究結果が分かれている。アラブ人を対象に調査した Porter（1991, cited in O'Sullivan, 2002）は，親密さは発話に影響しないと結論づけた。一方 O'Sullivan（2002）は，12組の日本人大学生の発話を調べ，ケンブリッジ英検 FCE の評価基準及び発話の文法的複雑さにおいて，顔なじみ同士のペアの方が有意に高い評価を得たことを突き止めた。

2.3.2　性別

同様に面接官あるいはペアを組む相手の性別による影響についても，様々な結果が出ている。性別の影響はないとする O'Loughlin（2002）や，男性の面接官の方が成績は良いとする Porter（1991, cited in O'Sullivan, 2000），反対に女性の面接官の方が成績は良いという O'Sullivan（2000）の研究報告もある。このように相反する結果が出た理由の一つに，Porter がアラブ人を，O'Sullivan が日本人を対象に研究を行ったことから，被験者の文化的背景が関係している可能性が指摘されている。

2.3.3　性格

ペア面接における性格の影響について，消極的な学生同士を組ませると量的にも質的にも向上しないことが報告されている（Alvarado, 1992, cited in Spelman, 1992）。更に Spelman（1992）は，積極性に富む学生と欠ける学生を，自分の見ている絵を言葉で説明する情報の送り手と，説明を聞いて絵を描く受け手のいずれかの役にあてがい，役割と性格

の組み合わせが発話にどのような影響を与えるかを調べた。その結果，役割にかかわらず積極的な学生と組んだペアは，情報のやり取りが順調に進んだが，消極的な者同士のペアは，タスクを途中でやめてしまった。この研究から，受験者の性格がペア面接に重要な影響を及ぼすことがうかがえる。

2.3.4 英語能力の個人差

Foot（1999）は，ペア面接で英語力の違う者同士を組ませると，力のある者が力のない者に遠慮し，全力を出せないと述べた。彼はテニスを例にとり，自分より下手な相手とテニスをするよりも，上手な相手とテニスをした方が上手にテニスができることから，能力差のある者を組ませた時のペア面接の有効性に疑問を投げかけた。

この Foot の批判に対して UCLES は，CPE テスト受験者からペア面接方式と個別面接方式で得られた，発話に関する量的並びに質的分析結果を公表し反論した。Taylor（2000）によれば，ペア面接では個別面接の60％程度に減ると考えられている受験者の発話量が，実際は75％前後維持されていた。また全部で30のコミュニケーション機能のうち，個別面接では平均14機能を確認したが，ペア面接では平均26もの機能を引き出すことができたという。後に Taylor（2001）は，発話の言語機能を大きく，例えば意見を述べる informational 機能，説得などの interactional 機能，更には話し合いを終えるなどの managing interaction 機能の3つに分類し，その割合を個別面接とペア面接で比較した。そして，個別面接では平均80％以上を占めていた informational 機能が，ペア面接では平均55％まで下がり，その分他の2つの機能が増えたことを明らかにした。

しかし，残念ながら Taylor（2000, 2001）が紹介したデータは，不均整な発話という，個別面接について言われてきた批判への反論としては説得力があるが，英語力の違いがマイナスの影響を及ぼすのではないかという批判には十分に答えていない。更に Taylor（2000, 2001）が分析したのが，5レベルのケンブリッジ英検の中で，いまだペア面接方式を取り入れていない最高レベルの CPE テスト受験者の発話である点が，一般化を難しくし説得力を欠く原因となっている。

UCLES の専門家に代わって Foot（1999）の批判に反論したのは，ハンガリーの職業高校でペア面接を試験的に行った Egyud and Glover（2001）である。彼らは，Foot がペア面接をテニスに例えたことについて，テニスは勝ち負けを競うゲームであるが，ペア面接の目的は相手を打ち負かすことではなく，相手と協力してタスクを達成することだと述べた。また，能力差が原因で十分に力を出せないのではとの批判には，仮にそのようなことがあるとすれば，その最たる弊害が生じているのは個別面接ではないかと反論した。

さて次に，この他の先行研究を見ることにしよう。Iwashita（1996）は，オーストラリアで日本語を学ぶ大学生20人を日本語能力により2つのグループに分け，ペアの組ませ方と得られる発話を調べた。日本語能力の評定は，初めに行った面接結果，日本での滞在経験，授業担当者のコメントに基づき行われた。分析は2人の評価官による文法，流暢さ，発音など6項目の4段階尺度評定と，C-unit と Turn の数を用いた。C-unit は，後述する T-unit に動詞を伴わない独立した句も含めた分析単位で，Turn とは，1人の話し手が話し始めてから，他者が話す権利を受け継ぐまでのすべての発話（Sacks, Schegloff and Jefferson, 1974）を意味する。分析の結果，日本語能力の個人差は発話量には若干の影響を与えるが，面接の評価自体にはほとんど影響しないことを突き止めた。発話量について Iwashita（1996）は，一般化は難しいとしながらも，被験者は上位者と組んだ時に，より多く話す傾向があると述べた。その後 Iwashita（2001）は，念押しや確認などの Confirmation Check や，相手の発言への明瞭化要求 Clarification Request，自己修正 Modified Output の使用頻度や割合を調査項目に，日本語能力の個人差が発話に及ぼす影響の質的分析を試みた。しかし，先行研究同様，日本語能力の個人差による有意差は確認できなかった。

また，英語学習者を対象とする数少ない研究を行った Yule and Macdonald（1990）は，英語能力と与えられる役割が発話にどのような影響を及ぼすかを分析した。アメリカの大学で学ぶ外国人大学院生40人を TOEFL のスコアをもとに2つのグループに分け，地図の配達経路を口頭で説明するタスクを行った。その際経路を説明する情報の送り手側と，相手の説明を聞いて地図上で配達経路を確認する情報の受け手側との間に，英語能力の差が生じるようペアを組ませた。その結果，英語能力の高い者が情報の

受け手に就いた場合に，情報の送り手に就いた時に比べてやり取りが2倍以上長く続くことがわかった。これは英語能力の低い者が受け手になると，あいづち程度の反応になってしまうが，英語力の高い者が受け手に就くと，たくさんの内容確認のための発話がなされることを意味している。しかしながら，この実験に参加した上位者20人の TOEFL 平均点数が625点，下位者20人の平均が562点であることから，この研究結果が日本の高校生にも当てはまるかどうかは疑わしい。

英語学習者を対象としたもう一つの Csepes（2001）による研究では，ハンガリーで英語を学ぶ高校3年生120人が実験に参加した。Csepes（2001）は，全部で38の空所からなるクローズテストと教師の質問紙による評価，生徒の自己評価の3つの資料をもとに，被験者の英語能力を3つに分けた。そして英語能力が中レベルと判断された30人のコア被験者は，自分より上位，同等，下位の英語能力を有する3人の被験者とペアを組み，それぞれ1回ずつ異なるタスクを行った。その後テープに録音されたコア被験者の90回のペア面接を，会話遂行技術，文法，語彙，発音の4項目について2人の評価官が8段階尺度で評価した。そして評価の合計点を比較した結果，コア被験者の発話評価に，ペアを組む相手の英語能力の違いによる有意差は見られなかった。このことから Csepes（2001）は，ペア面接ではペアの相手が自分より英語能力が優れている場合も，逆に劣っている場合も，発話の評価は影響を受けないと結論づけた。

3　実験

3.1　目的

初級～中級レベルの英語学習者を対象にしたペア面接の有効性についての研究は非常に少ない。このため，ハンガリーの高校生を対象に調べた Egyud and Glover（2001）と Csepes（2001）の研究は大変貴重である。しかし残念ながら Egyud and Glover（2001）は，発話を書き起こした transcript の分析のみで，発話の量的分析まで踏み込んでいない。また Csepes（2001）は，被験者を分ける時に用いた英語能力を，クローズテストや質問紙を用いて測定し，実際にスピーキング能力を測っていない。

そこで本研究では，英検準2級の2次面接により測定したスピーキング能力に，全国模擬試験の成績を加え，グループ分けのもととなる英語能力とした。また受験者がより対等な立場で会話に参加することを可能にし（Kormos, 1999），発話量と発話機能の豊富さで個別面接を上回る（Lynch, 1999）と言われるロールプレイをタスクとして採用した。さらに，発話量や発話の複雑さ，正確さ，流暢さ，発話の機能を示す指標に注目し，発話の量的分析を通してペア面接方式の有効性を検証しようと試みた。

3.2　仮説

本研究では，次の2つの作業仮説を立てた。

仮説1　ペア面接において，英語能力の個人差は，発話量，正確さ，複雑さ，流暢さ，発話の機能に影響しない。

仮説2　面接方式の違いは，発話量や正確さ，複雑さ，流暢さ，発話の機能に影響しない。

3.3　被験者

栃木県内の公立高校に学ぶ2年生の女子生徒8人。これら8人の生徒は，同級生39人の中から，英検準2級の模擬面接結果と，全国規模で実施された2回の模擬試験の英語全国偏差値に基づいて選んだ。親密さの影響を指摘した O'Sullivan（2000）の研究を受けて，被験者は全員1つのクラスから選んだ。

実験の直前にクラス生徒全員に，英検準2級の過去の試験問題を使い2次面接を行った。面接は同一の面接官が3日間に分けて実施し，すべてテープに録音した。その結果は後日，英検面接委員有資格者2名が，実際の2次面接の評価基準に基づき採点した。採点が食い違った場合は，協議の上決定した。2人の採点者間の評価の一致度は，相関係数0.76であった。39人の成績概況（点数）は，33点満点中最高30点，最低12点，平均22.5点，標準偏差4.1であった。39人の中には，英検2級取得者が1人，準2級取得者が25人含まれていた。

もう一つの模擬試験の成績は，実験の11か月前（1回目）と3か月前（2回目）に，大手模擬試験業者がそれぞれ約44万5000人の高校1年生と約41万8000人高校2年生を対象に実施したものである。39人の成績概況（全国偏差値）は，次の通りである。1回目（最高72.2，最低39.5，平均54.6，標準偏差6.7）。2回目（最高75.1，最低45.3，平均55.3，標準偏差6.8）。

被験者8名は，クラス生徒の中から，英検準2級模

擬面接の評点と模擬試験の英語全国偏差値の平均が，そろって上位または下位の者の中から4名ずつを選んだ。また39人の中には数人の男子生徒が含まれていたが，性別の違いによる発話への影響を考え，被験者は全員女子の中から選んだ。更に，ロールプレイでは外向的な受験者が高い評価を受ける可能性（馬場，1997，p.119）に配慮し，クラス担任の意見を参考に，極度に外向的あるいは内向的と思われる生徒は被験者から除外した。

最終的に選ばれた被験者の英検準2級2次面接成績は，上位者平均が25.3（26，26，25，24），標準偏差0.96，下位者平均が17.5（14，16，19，21），標準偏差3.11だった。2グループ間には5％水準で有意差があることが確認された。また被験者の全国模擬試験英語の平均偏差値は，上位者平均が59.8（64.6，59.7，59.0，56.0），標準偏差が3.56，下位者平均が46.6（43.4，44.8，47.7，50.3），標準偏差が3.08だった。こちらも2グループ間に5％水準で有意差があることが確認された。

3.4　実験に用いたロールプレイ

Evans et al.（1987）は，ロールプレイを場面，役割，登場人物の性格，話の展開の4項目について，2人に与えられた役割が同じか異なるか，自由度がどの程度あるかという観点から分類する方法を提案した。そしてテストに適したロールプレイとして，次のようなロールプレイを挙げた。

> あなたたちは，異なる語学学校で働く語学教師です。ある日の午後，久しぶりに2人で喫茶店にやって来ました。お互いの仕事や生徒について話をしてください。

このロールプレイでは，同じ場面と役割が2人に与えられ，登場人物の性格や話の展開，更には話の内容については全く決められていない。Evans et al.（1987）は，このロールプレイがテストに適している理由として，指示が簡潔に済み，受験者が自由にその後の話の展開を決められる点を挙げている。しかし，ロールプレイを高校の授業で実践してきたSato（2001）は，話の展開が決められていない自由度の高いロールプレイは，英語で考え，瞬間的に英語を話すことに慣れていない，日本人の高校生には非常に難しいと述べた。

このため本研究では，まず役割に対する親密度が影響しやすいとの指摘に対応するため，高校生が遭遇しやすい場面で，互いに気心の知れた友人という対等の役割を設定した。話の展開については，自由度が高すぎて受験者が困らないよう，話の大まかな展開を提示することにした。また被験者は，自分のロールカードのみを示され相手の出方を予想できないため，あらかじめ会話を始める側をロールカードに明記した。

〈ロールカードの例〉

第1週　役割 A

> あなたは友達からカラオケに誘われましたが，歌が上手ではないので行きたくありません。テスト勉強を理由に誘いを断りなさい。

役割 B

> あなたは久しぶりに友人何人かを誘って，カラオケへ行く計画を立てました。これから友達に電話をしてぜひ一緒に来てくれるよう説得しなさい。ロールプレイはあなたが電話をかけるところから始まります。

（2週目以降のロールプレイは資料1参照）

3.5　手順

週1回50分の授業時間内に，4週連続で実験を行った。毎回生徒同士によるペア面接と，教師と生徒との個別面接を実施した。被験者には，今回のロールプレイは話す力を測るためのテストで，個別面接とペア面接のどちらが，より正確に話す力を測れるかを調べるための実験であることを伝えた。また，与えられたロールカードに書いてあること以外に自由に付け加えて，90秒間できるだけたくさん話をするよう指示した。面接順並びにペア面接の組み合わせは，表1の通りである。

順序効果を考慮し，前半の2週間は初めにペア面接（表の網掛け部分），その直後に個別面接を，後半の2週間は順番を入れ替え，個別面接終了後にペア面接を行った。また第1週と第3週は上位者と下位者を，第2週と第4週は上位者同士，あるいは下位者同士を組み合わせた。被験者は，同じ役割をペア面接と個別面接で演じた。また発話の機能を分析する際に偏りが生じないように，毎回上位者・下位者それぞれ2名が役割 A を，残り2名が役割 B になるよう割り当てた。ロールプレイはすべてテープに録音した。

■ 表1：面接順並びにペア面接組み合わせ

	1週目		2週目		3週目		4週目	
	role A	role B	role A	role B	role A	role B	role A	role B
1組目	a1	c1	a1	a2	t	c3	t	c4
	a1	t	a1	t	a4	t	c2	t
	t	c1	t	a2	a4	c3	c2	c4
2組目	a2	c2	a3	a4	t	c2	t	c3
	a2	t	a3	t	a3	t	c1	t
	t	c2	t	a4	a3	c2	c1	c3
3組目	c3	a3	c1	c2	t	a2	t	a4
	c3	t	c1	t	c1	t	a2	t
	t	a3	t	c2	c1	a2	a2	a4
4組目	c4	a4	c3	c4	t	a1	t	a3
	c4	t	c3	t	c4	t	a1	t
	t	a4	t	c4	c4	a1	a1	a3

教師 t　　上位者 a1 a2 a3 a4　　下位者 c1 c2 c3 c4

ロールプレイは1分30秒間で，初めてロールプレイをする場合のみ，1分間の準備時間を与えた。Wigglesworth（1998）は，107人を対象にテープ面接方式で1分間の準備時間の効果を調べた。そして準備時間の有無は，3つのタスクにおいて流暢さや文法，発話中の節の数に影響しないことを発見した。本研究ではこの研究結果に基づき，準備時間を1分間とした。

すべての個人面接は，被験者にロールプレイの内容が漏れることを避け，また面接官要因を抑制するため，1人の教師が50分の授業時間内に行った。このためこれ以上被験者数を増やしたり，ロールプレイの時間を長くすることは不可能だった。実験に参加した被験者以外の生徒は，他の英語科教員の指導の下，実験に用いたのとは異なるロールプレイを使い，英会話練習を行った。

3.6　分析方法

録音した48のロールプレイをテープから書き起こし，5つの分析項目（発話量，複雑さ，正確さ，流暢さ，発話の機能）について対応する指標を調べた。まずペア面接において能力差が及ぼす影響を調べるため，4回行ったペア面接のうち最初の2回分（上位者，下位者それぞれ延べ8人分）のデータを分析した。ペア面接を個別面接後に実施した後半の2回分については，練習効果の影響を避け分析には用いなかった。次に面接方式が及ぼす影響を見るため，4回分のペア面接と個別面接（上位者，下位者それぞれ延べ16人分）について，得られた指標を比較した。分析には SPSS 12.0J for Windows を用い，対応のある2つの母集団を比較するため，検定には Wilcoxon の符号付順位検定を行った。

3.6.1　発話量

指標：すべての発話から，単語1語からなる間投詞や繰り返し，言い淀み，自己訂正した語の数を引いた数

この指標を定めるに当たっては，日本人中学生のスピーキングの特徴を調べ，多くの分析項目の中から語彙や T-unit の数，話した秒数の3指標がスピーキング能力と中程度の関連があると報告した小泉・栗嵜（2002）を参考にした。単語1語の間投詞とは，意味上重要でない oh, ah, um, uhm, so を指す。また例えば，Yes, but I, I have, I was catch, catch a cold. の場合，繰り返しなどを除いた最終的な発話語数は，次のように6語と数えた。Yes, but (I 繰り返し), (I have 自己訂正), I (was catch 自己訂正), catch a cold.

3.6.2　複雑さ

指標： T-unit 1つ当たりの平均語数

T-unit とは1つの主節を含み，その主節に付加される従属節（Hunt, 1970, cited in 平野，1990）と定義される。例えば次のような重文（I have an uncle and his name is Roger.）の場合，T-unit は2つで，複文の場合（I have an uncle, whose name is Roger.）は，T-unit は1つと数える（平野，1990）。

言語能力が高ければ高いほど，より多くの情報を1つに組み入れ，より長い T-unit を書くことができるとされる。このため T-unit は，これまで書き言葉のみならず，話し言葉の統語的複雑度を示す，客観性と信頼性の高い尺度（平野，1990）として多くの研究に用いられてきた。T-unit を談話分析に用いた先行研究には，Long and Porter（1985），Spelman（1992），Young and Milanovic（1992），Halleck（1995），Young（1995），Kenyon（1998），Iwashita（2001）などがある。

3.6.3　正確さ

指標 1：誤りの数

例えば，I'm very poor (at) sing(ing) a song. のように続けて誤りが起こった場合は，併せて1つと数えた。

指標 2：全体の T-unit に占める誤りのある T-unit の

割合（誤りのある T-unit 数÷全体の T-unit 数）

3.6.4 流暢さ

指標： 1分間に話す平均語数（すべての発話語数÷話した秒数×60）

ここで示す語数は実際に話した語数ではなく，計算上割り出した語数である。

3.6.5 発話の機能

指標：ペア面接で得られた発話を，会話の舵取り，情報提供，会話維持，会話中断の4つの機能に分類し比較した。分類の際は，T-unit のみならず T-unit に満たない断片的な発話も分類の対象とした。会話中断以外の発話機能の定義は Spelman（1992）による。

会話の舵取りとは，会話の流れを先導し，相手の会話への参加を規制する機能。具体的には，命令文や疑問文，上がり調子の聞き返しなど，相手からの返答を求める発話がこれに属する。情報提供とは，相手に情報を提供する機能。相手の質問や命令に対する受動的な返答や，自発的な情報発信を指す。会話維持とは，新しい情報を含まずに，単に会話の流れを維持する機能。相手が直前に言ったことの繰り返しや言い換え，uhm や yes などのあいづちがこれに属する。会話中断は，10秒以上沈黙した状態を会話中断と定義した。

4 実験結果

4.1 ペア面接において能力差が及ぼす影響

4.1.1 発話量

上位者は上位者と組んだ時，下位者と組んだ時より発話語数が多かった。一方下位者は，上位者と組んだ時の発話量が，下位者同士でペアを組んだ時をわずかに上回った（表2）。上位者・下位者とも，ペアの組み方による発話量に有意差はなかった。また必ずしも常に上位者が，下位者より多く話すわけではないこともわかった。

4.1.2 複雑さ

T-unit に含まれる語数は，上位者と下位者が組んだ方が，わずかながら上位者あるいは下位者同士が組んだ時より多かった（表3）。しかしその差は有意ではなかった。いずれの組み合わせ方をした場合も，上位者の発話が，下位者の発話に T-unit に含まれる語数で勝っていた。

4.1.3 正確さ

上位者同士，下位者同士のペアの方が，上位者と下位者が組んだ時より，多くの誤りを犯す傾向が見られた（表4）。また同様に，能力差のないペアの発話の方が能力差のあるペアの発話より，誤りのある T-unit が全体に占める割合も高かった（表5）。

■ 表2：発話量（平均発話語数）への影響

グループ	平均発話語数（標準偏差）	増減	平均発話語数（標準偏差）	差
上位者 N=8	能力差あり 22.8 (8.7)	＜	能力差なし 27.0 (5.3)	4.2 有意差なし
下位者 N=8	能力差あり 26.6 (10.8)	＞	能力差なし 25.6 (11.8)	1.0 有意差なし

■ 表3：複雑さ（T-unit 当たりの平均語数）への影響

グループ	T-unit 当たりの平均語数（標準偏差）	増減	T-unit 当たりの平均語数（標準偏差）	差
上位者 N=8	能力差あり 5.9 (1.6)	＞	能力差なし 5.4 (0.8)	0.5 有意差なし
下位者 N=8	能力差あり 5.5 (1.5)	＞	能力差なし 5.2 (1.1)	0.3 有意差なし

■ 表4：正確さ（誤りの数）への影響

グループ	誤りの数（標準偏差）	増減	誤りの数（標準偏差）	差
上位者 N=8	能力差あり 0.6 (1.1)	＜	能力差なし 1.6 (0.9)	1.0 有意差なし
下位者 N=8	能力差あり 1.5 (0.8)	＜	能力差なし 1.9 (1.6)	0.4 有意差なし

■ 表5：正確さ（誤りのあるT-unitの割合）への影響

グループ	誤りのあるT-unitの割合（%）（標準偏差）	増減	誤りのあるT-unitの割合（%）（標準偏差）	差
上位者 N=8	能力差あり 22.9 (0.4)	＜	能力差なし 34.2 (0.2)	11.3 有意差なし
下位者 N=8	能力差あり 33.3 (0.2)	＜	能力差なし 38.9 (0.3)	5.6 有意差なし

■ 表6：流暢さ（1分間の発話語数）への影響

グループ	1分間の発話語数（標準偏差）	増減	1分間の発話語数（標準偏差）	差
上位者 N=8	能力差あり 162.1 (65.9)	＞	能力差なし 128.2 (79.4)	33.9 有意差なし
下位者 N=8	能力差あり 152.0 (88.0)	＞	能力差なし 123.1 (58.2)	28.9 有意差なし

4.1.4　流暢さ

上位者と下位者が組んだ時の1分間の平均発話語数は，上位者同士，下位者同士が組んだ時よりも多かった（表6）。上位者・下位者とも，英語能力が異なる相手と組んだ時の方が，平均30語ほど1分間の発話語数が増えたことになる。検定の結果，上位者あるいは下位者同士の平均発話語数との間に有意差は確認できなかった。

4.1.5　発話の機能

ペアを組む相手によって発話の機能がどのように変化したかを示したのが表7である。上位者同士を組み合わせた場合に，より多くの自発的な情報発信が見られた他は，目立った違いはない。しかし能力と発話機能との関係で見ると，聞き返しや会話中断，質問，命令といった発話機能に偏りが見られた。まず聞き返しは，上位者の発話に8回見られるが，下位者の発話には1回も見られなかった。次に会話中断は，下位者よりも上位者の方が多かった。逆に質問や命令の回数では，下位者が上位者を上回った。

4.2　面接方法が及ぼす影響

4.2.1　発話量

上位者・下位者とも個別面接時の発話語数が，ペア面接時の発話語数をわずかに上回った（表8）。被験者ごとに分析すると，8人中7人について個別面接時の発話量がペア面接の発話量より多いことがわかった。いずれの差も有意ではなかった。

■ 表8：発話量（平均発話語数）への影響

グループ	平均発話語数（標準偏差）	増減	平均発話語数（標準偏差）	差
上位者 N=16	ペア面接 24.9 (7.3)	＜	個別面接 28.1 (9.1)	3.2 有意差なし
下位者 N=16	ペア面接 26.1 (11.0)	＜	個別面接 27.5 (9.1)	1.4 有意差なし

4.2.2　複雑さ

個別面接よりもペア面接において，T-unit当たりの平均語数が多い傾向が，上位者・下位者ともに見られた（表9）。

■ 表9：複雑さ（T-unit当たりの平均語数）への影響

グループ	T-unit当たりの平均語数（標準偏差）	増減	T-unit当たりの平均語数（標準偏差）	差
上位者 N=16	ペア面接 5.6 (1.2)	＞	個別面接 5.0 (1.1)	0.6 有意差なし
下位者 N=16	ペア面接 5.4 (1.3)	＞	個別面接 5.3 (0.7)	0.1 有意差なし

■ 表7：ペア面接において，スピーキング能力個人差が発話機能に及ぼす影響

グループ	組み合わせ	会話の舵取り			情報提供		会話維持		会話中断
		命令	質問	聞き返し	質問の応答	情報発信	繰り返し	あいづち	沈黙
上位者 N=8	能力差あり	3	5	4	11	13	1	1	4
	能力差なし	1	6	4	7	18	0	3	3
下位者 N=8	能力差あり	3	9	0	8	17	0	2	2
	能力差なし	5	8	0	11	15	0	0	1

4.2.3 正確さ

英語能力にかかわらず，誤りの数（表10），誤りのある T-unit が全体に占める割合（表11）とも，個別面接の方が高かった。つまり上位者・下位者とも，ペア面接よりも個別面接において，誤りを犯す傾向があった。しかし面接方式による正確さに，有意差は認められなかった。

■ 表10：正確さ（誤りの数）への影響

グループ	誤りの数（標準偏差）	増減	誤りの数（標準偏差）	差
上位者 N=16	ペア面接 1.1 (1.1)	＜	個別面接 1.6 (1.1)	0.5 有意差なし
下位者 N=16	ペア面接 1.7 (1.3)	＜	個別面接 2.1 (1.7)	0.4 有意差なし

■ 表11：正確さ（誤りのある T-unit の割合）への影響

グループ	誤りのある T-unit の割合（%）（標準偏差）	増減	誤りのある T-unit の割合（%）（標準偏差）	差
上位者 N=16	ペア面接 28.5 (0.3)	＜	個別面接 38.6 (0.3)	10.1 有意差なし
下位者 N=16	ペア面接 36.1 (0.3)	＜	個別面接 51.1 (0.3)	15.0 有意差なし

4.2.4 流暢さ

英語能力にかかわらず，個別面接における1分間の平均発話語数は，ペア面接の発話語数を有意に上回った（表12）。これは上位者・下位者とも，ペア面接では1分間に約140語話す流暢さが見られたが，個別面接では70語に減り，流暢さが半分に低下したことを示している。

■ 表12：流暢さ（1分間の発話語数）への影響

グループ	1分間の発話語数（標準偏差）	増減	1分間の発話語数（標準偏差）	差
上位者 N=16	ペア面接 145.1 (72.6)	＞	個別面接 68.4 (18.2)	76.7**
下位者 N=16	ペア面接 137.5 (73.6)	＞	個別面接 68.8 (16.6)	68.7**

** p<.001

4.2.5 発話の機能

ペア面接に比べ個別面接では，生徒が発した質問の数が減少した。特に下位者ほどこの傾向は顕著で，質問数はペア面接の17から個別面接の7に減った（表13）。代わりに個別面接では，上位者・下位者とも質問への応答が，ペア面接時に比べて3倍に増えた。さらに会話中断の数も，個別面接ではペア面接の3倍近く増加した。

5 考察並びに今後の課題

本研究では，発話量，複雑さ，正確さ，流暢さ，発話の機能の5つの指標について，ペア面接方式において能力差が及ぼす影響と，面接方式が及ぼす影響を調べた。検定の結果有意であったのは，面接方法の違いが流暢さへ及ぼす影響のみであった。このことから，能力差は，ペア面接において発話量や正確さ，複雑さ，流暢さ，発話の機能に影響しないと

■表13：面接方法が発話機能に及ぼす影響

面接方法	グループ	会話の舵取り			情報提供		会話維持		会話中断
		命令	質問	聞き返し	質問の応答	情報発信	繰り返し	あいづち	沈黙
ペア面接	上位者 N=16	4	11	8	18	31	1	4	7
	下位者 N=16	8	17	0	19	32	0	2	3
個別面接	上位者 N=16	4	9	8	55	34	2	3	19
	下位者 N=16	3	7	2	61	33	2	6	12

いう仮説1は支持されたと言える。また面接方式の違いは，発話量や正確さ，複雑さ，流暢さ，発話の機能に影響しないとした仮説2について，今回の実験結果は面接方式が発話量，正確さ，複雑さに影響しないことを示唆した。有意差の見られた流暢さと，明らかに質問や質問への応答，会話中断の数に違いが見られた発話機能については，面接方式の違いが影響したと言える。しかし，今回の研究はサンプル数が少ない上，以下に述べる統制の課題もあり，能力差や面接方式の違いが及ぼす影響について完全に実証できたとは言い難い。

統制上の課題の一つは，練習効果である。ペア面接を先行実施した最初の2週間は個別面接の，個別面接を先行実施した後半の2週間はペア面接の発話量が多くなった。このことから，今回の実験では練習効果が働いたと判断できる。面接でのパフォーマンスには様々な受験者要因が作用するため，実験で操作する要因以外をできるだけ等しくする必要がある。しかしこれを完全に実行することは非常に困難である。この実験では被験者の英語能力を，唯一他と異なる要因（独立変数）としたが，実験結果には能力差よりも性格や反応の速さなど，個人差の影響の方が強いという Iwashita（2001）の指摘が当てはまる点も見られた。被験者8人を個別に見ると，一番発話量が多かったのは下位者に分類された被験者であり，逆に全部で11回と，最も多く黙り込んでしまったのは上位者の1人であった。このことは英語能力以外の受験者個人の要因が，ペア面接に影響を及ぼす可能性があることを示唆している。受験者要因をいかに統制し，同時に順番効果，練習効果を実験から排除するか，この先の検討は今後の研究にゆだねる。

2番目の課題は，英語能力とは何かをどう規定するかである。本実験では，英語能力を英検準2級2次面接と全国模擬試験で測られる能力と定めた。しかし今回の英語能力の中には，直接音声を聞かせて測定したリスニング能力は入っていない。英語能力をどう定義するかにより，今回と異なる実験結果が生まれる可能性も否定できない。実験結果の一般化のためには，英語能力のより厳密な定義づけが必要となるだろう。

次に，タスクとしてのロールプレイの均等性が挙げられる。4つのロールプレイについて，それぞれ回ごとに個別面接での発話量を分析したところ，3番目のロールプレイでの発話量が，他の3回に比べて少なかった。これは一つには，2回目まで個別面接をペア面接後に行ったため，被験者が個別面接を最初に行うことに戸惑い，普段以上に緊張したことが予想される。しかしもう一つには，3番目のロールプレイが他よりも難しかったから発話量が減ったとも考えられる。CD を貸した側は状況を説明することが，CD を借りた側は相手の言わんとする状況を理解することが難しかったようだ。このことは教師の発話量や被験者の沈黙回数からも読み取れる。3番目のロールプレイでは，教師の発話量と被験者が沈黙した回数が，他のロールプレイに比べて多い。与えられた状況を英語で表現できず，あるいは相手の言っていることが理解できず被験者が寡黙になってしまったため，教師がその分多く話さざるを得なくなったと考えられる。いくつかのロールプレイを用いて，そこで得られた発話を比較する場合，ロールプレイの難易度が均等でなくてはならず，ロールプレイを作成する上で大きな課題である。

ところで個別面接において沈黙が多くなった理由を被験者に尋ねたところ，発話に対する教師の反応が速く，答えを考える時間がなかったからという答えが多かった。これは多分にロールプレイをタスクに採用したことが，影響したかもしれない。Lier（1989）は，ロールプレイをインタビューに取り入れた場合，受験者が面接官との間の歴然とした準備状況の差に苦しめられる可能性を指摘した。面接官はあらかじめ話の展開を予想し練習が可能であるため，ロールプレイを初めて演じる受験者は，非常に不利な立場に置かれることになる。このような準備状況の違いが，被験者を更に苦しめる原因になったとも考えられる。このことからも今後ロールプレイを初級～中級の英語学習者用面接に用いる際は，より厳密な難易度調整が重要となるだろう。タスクの複雑さが，発話の流暢さや正確さ，複雑さへ及ぼす影響はないと結論づけた Elder, Iwashita and McNamara（2002）の研究もあるが，ロールプレイの場合にも同じことが言えるか検証する必要性があると言える。

更に言えば，ロールプレイを用いて面接を行う場合，評価項目と評価基準も課題である。残念ながら，ロールプレイの評価に関する先行研究は，非常に少ない。Kitao and Kitao（1996）は，ロールプレイをテストに用いる場合，謝る，頼む，情報を求めるなど要求された機能を遂行できるかどうかで，受験者

の英語能力を評価できると述べた。しかし今回のようにペア面接方式でロールプレイを用いる場合，2人のやり取りが面接官の予想通り進むとは限らず，予期した発話機能を引き出せない可能性もある。従って個別面接・ペア面接いずれの場合も利用できる，ロールプレイの全体的評価項目並びに評価基準の開発が，今後求められる。

6 結論

本研究では発話の量的分析に基づき，ペア面接方式の有効性を検証した。その結果，ペア面接では英語能力の個人差が，発話量や正確さ，複雑さ，流暢さ，発話の機能に影響しないことを確認した。また面接方式の違いが，発話量や正確さ，複雑さに影響しないことも明らかになった。これらの結果は，少数の限られた被験者から得られたデータであるため試験的ではあるものの，ペア面接の有効性が支持されたことを意味する。

また，個別面接とペア面接で得られた発話の流暢さを示す指標に，統計上有意な差が見られた。これは面接方式の違いだけでなく，ロールプレイをタスクに用いたことが一因と考えられる。話の展開についての情報の有無が，面接官と受験者が必要とする反応時間の違いに結びついたのだろう。今回の実験結果を一般化するためには，受験者要因の統制と順番効果の排除，英語能力のより厳密な定義づけ，タスクとしてのロールプレイの均等性確保，更にはロールプレイの全体的評価項目と評価基準の開発などの課題が残されている。

最後に，時間短縮とともにロールプレイをペア面接で行うことの利点に，教育上の波及効果を忘れてはならない。Saville and Hargreaves（1999）が指摘するように，ペア面接には教室で学習者同士のやり取りを奨励する，プラスの波及効果が期待できる。また，英語を使う機会を与えられず興味や動機を失いかけている，上級の日本人英語学習者に最適だと，Sato（2001）が評したロールプレイを面接に用いることで，自分の英語能力を確認し，英語学習に弾みがつくと予想される。今後，こうしたペア面接方式の教育的波及効果をも視野に入れた研究が望まれる。

謝　辞

このような貴重な研究の機会を与えてくださった，(財) 日本英語検定協会と選考委員の皆様に感謝いたします。また独りよがりな私の実験計画について，多くの示唆に富む助言をくださった選考委員の和田稔先生と筑波大学大学院の小泉利恵さんに，厚くお礼申し上げます。最後に，実験に快く協力してくれた同僚や生徒の皆様に感謝申し上げるとともに，陰で支えてくれた家族に感謝したい。

参考文献（＊は引用文献）

＊馬場哲生（編著）．（1997）．『英語スピーキング論』．河源社．

＊Csepes, I.（2001）. *Is testing speaking in pairs disadvantageous for students? A quantitative study of partner effects on oral test scores.* Paper presented at the 11th IATEFL-Hungary Conference. Retrieved February 18, 2003, from http://www.novelty.hu/htm2/vol91/csepes.html

＊Egyud, G. and Glover, P.（2001）. Oral testing in pairs — a secondary school perspective. *ELT Journal*, 55（1）, 70-76.

＊Elder, C., Iwashita, N. and McNamara, T.（2002）. Estimating the difficulty of oral proficiency tasks: what does the test-taker have to offer? *Language Testing*, 19（4）, 347-368.

＊Evans, C., Gaudin, J-M. and Raveau, F.（1987）. Testing advanced communicative competence through role-play. *British Journal of Language Testing*, 25（2）, 104-108.

＊Foot, M.（1999）. Relaxing in pairs. *ELT Journal*, 53（1）, 36-41.

＊Halleck, G. B.（1995）. Assessing Oral Proficiency: A Comparison of Holistic and Objective Measures. *The Modern Language Journal*, 79（2）, 223-234.

＊平野絹枝（1990）．「言語能力の客観的指標の妥当性―日本人 EFL 大学生の場合」．『上越教育大学研究紀要』第9巻第2分冊．65-77．

＊Hughes, A.（1989）. Testing oral ability. A. Hughes（ed.）*Testing for Language Teachers*, 101-115, Cambridge University Press.

＊Ikeda, K.（1998）. The Paired Learner Interview: A Preliminary Investigation Applying Vygotskian Insights. *Language, Culture and Curriculum*, 11（1）, 71-96.

＊Iwashita, N.（1996）. The validity of the paired interview format in oral performance assessment. *Melbourne Papers in Language Testing*, 5（2）, 51-66.

＊Iwashita, N.(2001). The effect of learner proficiency on interactional moves and modified output in non-native-nonnative interaction in Japanese as a foreign language. *System*, 29, 267-287.

Jenkins, S. and Parra, I.(2003). Multiple Layers of Meaning in an Oral Proficiency Test: The Complementary Roles of Nonverbal, Paralinguistic, and Verbal Behaviors in Assessment Decision. *The Modern Language Journal*, 87, 90-107.

＊Kenyon, D. M.(1998). An Investigation of the Validity of Task Demands on Performance-Based Tests of Oral Proficiency. A.J. Kunnan (ed.), *Validation in Language Assessment*, 19-40, Lawrence Erlbaum Associates.

＊Kitao, S. and Kitao, K.(1996). Testing Communicative Competence. *The Internet TESL Journal*, 7(5), May. http://iteslj.org/Articles/Kitao-Testing.html

＊小泉利恵・栗嵜逸美(2002).「日本人中学生のモノローグにおけるスピーキングの特徴」. 関東甲信越英語教育学会『研究紀要』第16号. 17-28.

＊Kormos, J.(1999). Simulating conversations in oral-proficiency assessment: a conversation analysis of role plays and non-scripted interviews in language exams. *Language Testing*, 16(2), 163-188.

＊Langham, C.S.(2001). Putting English Language Proficiency Tests on the Curriculum.『東京家政学院筑波女子大学紀要』第5集. 153-159. http://www.kasei.ac.jp/library/kiyou/2001/11.LANGHAM.pdf

＊Lier, L.V.(1989). Reeling, Writhing, Drawling, Stretching, and Fainting in Coils: Oral Proficiency Interviews as Conversation. *TESOL Quarterly*, 23(3), 489-508.

＊Long, M. and Porter P.(1985). Group Work, Interlanguage Talk, and Second Language Acquisition. *TESOL Quarterly*, 19(2), 207-228.

＊Lynch, J.(1999). *The Role of Role Play*. Unpublished MA thesis. University of Surrey.

＊Maxwell, C.(1997). *Role Play and Foreign Language Learning*. Paper presented at the Annual Meeting of the Japan Association of Language Teachers. ERIC Document Reproduction Service No. ED 416 688.

＊May, L.(2002). *An Exploration of the Validity of the Paired Candidate Interaction*. Abstract of the paper presented at 24th Language Testing Research Colloquium. Retrieved March 5, 2003, from http://www.engl.polyu.edu.hk/ACLAR/ltrc_documents/ltrc_networks/may%20edit.doc

＊McCarthy, M.(1991). Spoken Language. M. McCarthy(ed.), *Discourse Analysis for Language Teachers*, 118-145. Cambridge University Press.

＊O'Loughlin, K.(2002). The impact of gender in oral proficiency testing. *Language Testing*, 19(2), 169-192.

＊O'Sullivan, B.(2000). Exploring gender and oral proficiency interview performance. *System*, 28, 373-386.

＊O'Sullivan, B.(2002). Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing*, 19(3), 277-295.

＊Sacks, H., Schegloff, E.A. and Jefferson, G.(1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, 50, 676-753.

＊Salaberry, R.(2000). Revising the revised format of the ACTFL Oral Proficiency Interview. *Language Testing*, 17(3), 289-310.

＊Salies, T.G.(1995). *Teaching Language Realistically: Role Play Is the Thing*. ERIC Document Reproduction Service No. ED 424 753.

＊Sato, R.(2001). *Role Play: Effective Role Play for Japanese High School Students*. ERIC Document Reproduction Service No. ED 454 709.

＊Saville, N. and Hargreaves, P.(1999). Assessing speaking in the revised FCE. *ELT Journal*, 53(1), 42-51.

＊Spelman, M.(1992). *The Importance of Discourse Style in Pairing Students for Interactive Communication Tasks*. ERIC Document Reproduction Service No. ED 355 601.

＊Taylor, L.(2000). Investigating the paired speaking test format. *Research Notes*, Issue 2, 14-15, UCLES. http://www.cambridge-efl.org/rs_notes/0002/rn2.pdf

＊Taylor, L.(2001). The paired speaking test format: recent studies. *Research Notes*, Issue 6, 15-17, UCLES. http://www.cambridge-efl.org/rs_notes/rs_nts6.pdf

＊UCLES(2004). *A Detailed Guide to KET*. http://www.cambridge-efl.org/support/dloads/pet/PET_HB_sampleS.pdf

＊Wigglesworth, G.(1998). The Effect of Planning Time on Second Language Test Discourse. A.J. Kunnan (ed.), *Validation in Language Assessment*, 91-110, Lawrence Erlbaum Associates.

＊Young, R.(1995). Conversational Styles in Language Proficiency Interviews. *Language Learning*, 45(1), 3-42.

＊Young, R. and Milanovic, M.(1992). Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition*, 14, 403-424.

＊Yule, G. and Macdonald D.(1990). Resolving Referential Conflicts in L2 Interaction: The Effect of Proficiency and Interactive Role. *Language Learning*, 40(4), 539-556.

資料1：2週目以降の実験に用いたロールプレイ

第2週　役割 A

あなたの家に，久しぶりに仲の良い友達が遊びに来ています。せっかくなので見せたいビデオがあります。もう少し長居するよう友達を説得しなさい。

役割 B

あなたは友達の家に遊びに来ています。遅くなってきたので，そろそろ帰ろうと思います。友達に失礼することを伝えなさい。ロールプレイは，あなたが帰りたいと言い出すところから始まります。

第3週　役割 A

あなたは今日友達に返すはずだった CD を家に忘れてきてしまいました。ロールプレイは，あなたが友達に会ったところから始まります。

役割 B

今日友達が，以前あなたが貸した CD を返すことになっています。そこであなたは，今日のうちにその CD を別の友達に貸す約束をしました。ロールプレイは，あなたが CD を貸した友達に会ったところから始まります。

第4週　役割 A

あなたは今度の日曜日に，友達と買い物に行く約束をしています。しかし風邪をひき，行けそうにありません。友達に電話をかけ事情を説明し，予定を来週末に延ばしてもらえるよう頼みなさい。ロールプレイは，あなたが電話をかけるところから始まります。

役割 B

あなたは今度の日曜日に，友達と一緒に買い物に行くことを楽しみにしています。そして来週末は，家族で出かけることになっています。ロールプレイは，一緒に買い物に行く友達から，電話がかかってくるところから始まります。

資料2：Transcript の実例

能力差のあるペアと能力差のないペアの例

Transcript 1

C2: This is @@@. May I speak to @@@ please?
A2: Speaking.（@@@ には名前が入る）
C2: I want to go to Karaoke with my friend, so let's go to Karaoke with my friends.
A2: Sorry, I can't go because I have, I have exam. So I must study.
C2: Oh, so when you, your exam is finished, shall we go Karaoke with me?
A2: Oh, okay. Ah, when I was finished my exam, I would go to Karaoke with you.
C2: Oh, so, when, when, when we go?
A2: My exam finish, will finish next week.
C2: So let's go to Karaoke on Sunday.
A2: Okay.

Transcript 2

C2: @@@, I want come back my house.
C1: Yes, but I, I want watch video. You see ...（... は沈黙を示す）
C2: But I must go home because it's six o'clock.
C1: ... But let's watch this video. It's interesting.
C2: Oh, but ... I, but I ..., I don't go home. I..., I ...

この2つの Transcript は，下位者（C2）が上位者とペアを組んだ時のロールプレイ（Transcript 1）と，下位者同士がペアを組んだ時のロールプレイ（Transcript 2）を示したものである。本研究では能力差の影響は検定の結果有意ではなかったが，Transcript 1のように下位者が上位者と組んだ時に会話が弾んだ例も少数ながら存在した。下位者同士が組んだ Transcript 2では，まず相手の沈黙が原因で会話が中断し，しまいには自分もしどろもどろの発話になってしまっている。

個別面接とペア面接の比較

Transcript 3

A1: Hello.
Teacher: Hello.
A1: We promised to go shopping next Sunday.
Teacher: Yes.
A1: But I catch a cold.
Teacher: Really?
A1: So I don't go.
Teacher: You cannot come? You cannot go?
A1: Yes. So please ...
Teacher: You want to change?
A1: Yes.
Teacher: Okay. When, when shall we go?
A1: Next ... next week.
Teacher: Next week? When, next week?
A1: ...
Teacher: Sunday, I'm going with my family.
A1: Oh ... next Saturday.

Teacher: Next Saturday. Okay, okay, I'm free next Saturday.
A1: ...
Teacher: So where shall we go next Saturday?
A1: Oh, I ...

Transcript 4

A1: Hello.
A3: Hello.
A1: We promised to go shopping next Sunday, but I catch a cold.
A3: Really?
A1: Yes, so I, I don't go.
A3: Oh, that's too bad.
A1: So, ... can you change this promise?
A3: Ah, okay, okay. When?
A1: Last week.
A3: Last week?
A1: Ah, next week.
A3: Ah, next week, I am going to go out with my family next weekend.
A1: Saturday or Sunday?
A3: On Sunday.
A1: So, we meet ...

この Transcript は，上位者（A1）が個別面接方式で実施したロールプレイ（Transcript 3）を，その直後に上位者同士のペア面接方式で行った際の発話（Transcript 4）を書き起こしたものである。個別面接では，途中で教師が確認のための質問をすると，それ以降被験者は一言二言答えるだけで済ましてしまう。対照的に Transcript 4からは，同じ被験者が一方的に応答するだけでなく，落ち着いてやり取りを交わしていることがわかる。ただ発した言葉自体は Transcript 3とほとんど変わっていないため，この場合被験者に見られる態度の変化が，面接方式の違いによるものか，練習効果によるものかわからない。

個別面接で被験者が終始主導権を維持した例

Transcript 5

A3: Hello.
Teacher: Hello.
A3: This is @@@.
Teacher: Hi.
A3: How are you?
Teacher: Fine, thank you.
A3: I have a plan to go to Karaoke.
Teacher: Oh, yes.
A3: Would you join us?
Teacher: When?
A3: Ah, next Sunday.
Teacher: Next Sunday? Oh, I have a test on Tuesday.
A3: Oh, really? Ah ... A lot of friends will come.
Teacher: Oh, it's a pity. I want to go, but I have to study.
A3: So um how about ... how about last month?
Teacher: Last month?
A3: Ah, next month.
Teacher: Next month, um, let me see. Maybe, but I'm not good at singing.
A3: Oh, don't worry. Me too.
Teacher: You, too?

これは上位者（A3）の個別面接である。この Transcript のように，個別面接で被験者が終始主導権を維持した例は非常に少ない。

沈黙により会話が中断された例

Transcript 6

A4: Hello, this is @@@.
C4: Hello, this is @@@.
A4: I, I plan to go to Karaoke.
C4: Oh, I ...
A4: Please come with me.
C4: Ah ... I must study for test. So I can't go.
A4: Um ...
C4: But ....
A4: Our other friends come to Karaoke.
C4: ...

これは，上位者（A4）と下位者（C4）が組んだペア面接である。途中で2人とも言葉につまり，会話が中断してしまった。このように一度中断すると，被験者だけではどうにもならないことが多かった。ペア面接では，このような場面での面接官の対応も大きな課題である。