



日本人高校生のL2英語発話の流暢性を高めるフィードバックのあり方：自己評価・ピア評価・指導者評価の有効性の検証

増見, 敦
石川, 慎一郎

(Citation)

神戸大学国際コミュニケーションセンター論集, 13:76-89

(Issue Date)

2016

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81009749>

(URL)

<https://hdl.handle.net/20.500.14094/81009749>



日本人高校生のL2 英語発話の流暢性を高める フィードバックのあり方 —自己評価・ピア評価・指導者評価の有効性の検証—

増見 敦¹石川 慎一郎²

1. はじめに

昨今、高校教育現場では学習指導要領の趣旨を踏まえ、受信と発信のバランスのとれた英語力の育成が重要視されている。平成25年の第2期教育振興基本計画(文部科学省, 2013)では、中学校卒業時及び高校卒業時における具体的な成果指標が設定され、それぞれヨーロッパ言語共通参照枠(Common European Framework of Reference for Languages(これ以降CEFR))(Council of Europe, 2001)が示すA1レベル50%, A2～B1レベル50%の目標が掲げられた(CEFRではABCの3段階をさらに2分割したA1, A2, B1, B2, C1, C2の合計6段階のレベルで言語力を示している)。しかしながら、「平成27年度英語教育改善のための英語力調査」(文部科学省, 2016)によれば、発信面(「話すこと」及び「書くこと」)について課題が大きいことが示された。特に、高校3年生の「話すこと」について、A2～B1レベルの到達度は11%にとどまることが明らかにされた。これは国が示す目標の「高校卒業段階A2～B1レベル50%」には程遠い現状であり、指導の改善が求められている。

「話すこと」(発話力)を考えた場合、いくつかの側面が考えられる。例えばSkehan(1998)は発話力の構成要素を、CAF(複雑性:Complexity, 正確性:Accuracy, 流暢性:Fluency)に分類している。「複雑性」は多様で複雑な文構造や語彙を処理する能力を、「正確性」は誤りのない言語の理解及び産出能力を、「流暢性」は主にアウトプットに関係し、速いスピードで言い換えなどを利用して言語産出ができる能力を指すとされている(門田他, 2016)。しかしながら、L2学習者にとってCAFのバランスを取ることは難しい。最終的にCAFのバランスのとれた発話者の養成を目指すことは言うまでもないが、特に指導の初期の段階では3つの中でどれを優先するかを考えた場合、まず流暢性に焦点を当てて指導をすることが求められるであろう。

では、高校段階の発話指導で特に指導の重点を置くべき流暢性について、実際の学習者の状況はどうだろうか。前述の英語力調査結果では、「与えられた課題について、(特に準備をすることなく)即興で(ある一定量の英文を)話す活動をしていた生徒は3割程度」と報告されている。流暢性の育成について、西

¹ 神戸大学附属中等教育学校 a.masumi@people.kobe-u.ac.jp

² 神戸大学大学教育推進機構国際コミュニケーションセンター／国際文化学研究科 iskwhshin@gmail.com

(2010)等の実践例があるものの、実際の現場で即興による発話の流暢性を高める実践の広がりはまだまだ十分とは言えない。「話すこと」自体ができない(あるいは苦手な)学習者を考えた場合、複雑性や正確性を求める前提として流暢性能力の育成を無視することはできないであろう。

流暢性を重視した指導が十分に普及していない背景にはいくつかの要因が考えられるが、とりわけ顕著なのが評価の難しさである。そもそも評価には、評定のための資料を得るという目的と、学習者に肯定的な示唆を与えて学習を促進させるという目的とがあるが、流暢性を高める指導実践を考えた場合、重要なのは後者のタイプの評価である。ここで問題になるのは、果たしてどのような評価方法が学習者の発話の流暢性を伸ばすかという点である。一般に、発話の評価については、誰が評価主体となるかにより、(1)発話者が自らの発表を振り返って評価する(自己評価)、(2)同じ環境で学ぶ他の学習者が評価する(ピア評価)、(3)指導者が評価する(指導者評価)の3つのタイプがありうる。

(1)の場合、学習者が内省によって自身の流暢性の不足を痛感し、そのことで流暢性が増す可能性があるが、一方で、ピアとの比較の観点がないために、流暢性の不足が実感できず、流暢性が高まらない可能性もある。(2)の場合、自分と同じ立場の友人に見られることが競争心を刺激して流暢性が増す可能性があるが、一方で、友人に対する羞恥心が阻害要因となって流暢性が減じる可能性もある。(3)の場合、指導者に見られることが励みとなって流暢性が増す可能性があるが、一方で、ミスを犯したくないという気持ちが阻害要因となって流暢性が減じる可能性もある。つまり、3つの評価には、発話の流暢性を向上させるという点に関して、それぞれ一長一短が予想されるわけだが、筆者が知る限り、これら3つの評価と実際の流暢性の変化の関係を明確に解明した先行研究はほとんど存在しない。

そこで本研究では、(1)自己評価、(2)ピア評価、(3)指導者評価の3つを取り上げ、それぞれの評価で学習者はどのようなフィードバックを得るのか、またいずれの場合に学習者の流暢性が最も大きく伸びるかを実証的に検証することとしたい。流暢性は一定時間内での発話量と定義する。なお、こうした調査では、評価の前後で発話量がどの程度変化したかを調べる人が多いが、1回だけの評価では偶然による差異の可能性が否定できない。そこで本研究では、評価行為を2回行うこととし、第1回発話→第1回評価→第2回発話→第2回評価→第3回発話という実験デザインを採用する。これにより、2回の評価行為を通して発話量が安定的に上昇するのはどの評価方法であるかを明らかにする。なお、こうした議論を行う場合、学習者の個人属性、とくに、習熟度が一定の影響を及ぼしている可能性が考えられる。そこで、習熟度の影響についてもあわせて検証したい。

2. 先行研究

英語授業における発話の指導、評価に関わる研究は数多い。ここでは、流暢性向上への取組を扱った研究、評価タイプの違いを扱った研究に分けて、それぞれの一端を概観する。

まず、前者については、教育実践を踏まえ、発話の流暢性を調査した研究として、たとえば西(2010)では、日本人高校生を対象とし、ワードカウンター(相手が話す英語の語数を数えるために作成された1枚の紙から成る記録用紙)を用いた1分間スピーチを授業内で継続的に実施した。その結果、初回に平均55WPM程度の発話語数が約1年後に93WPMに上昇したことを報告(語数カウントは他の学習者)し、「語数」という客観的評価により、流暢さを重視することで、流暢さの伸びに加え、学習者の英語に対する自己効力感の向上等、情意面での効果も報告している。また、三宅(2013)では、日本人大学生を対象に1分間英語発話録音活動を11週計21回に渡り継続的に実施した(その際、評価は行われていない)結果、平均発話語数につい

て有意な上昇が確認されたことを報告している。

次に、後者に関して、発話指導においては、適切な評価者の存在が重要であり、評価者タイプの違いについて扱った研究も多い。たとえば学習者相互評価と自己評価の全般的な内容について、藤原他(2007)では、それぞれの特徴を言及し、自己評価では、自身の成果を振り返ることは次の課題遂行にとって有用である一方、心理的防衛のため、評価の際に自身の弱点を避ける場合もありうるといった自己評価の限界を指摘している。一方、相互評価については、ピアへの評価方法の学びを促進すると同時に、他の学習者を評価すること自体が自己への振り返りにつながると示唆し、相互評価の有効性を指摘している。また発話の評価者に関する研究について、たとえば菅沼(2013)では、日本人高校生を対象に1分間英語スピーチを用いて自己評価と学習者相互評価の特徴について比較調査した。その結果、Language Use(言語使用)に関する評価に関して、自己評価の信頼性の高さを報告し、積極的にLanguage Useに関する自己評価を授業に取り入れることを推奨している。また深澤(2009)では、日本人高校生を対象に発話活動の中のスピーチに焦点を当てた調査研究を行った。スピーチ発表時の生徒相互評価の妥当性について、項目応答理論のラッシュ・モデル分析を利用した検証を行い、生徒相互評価と教員評価との相関が高いこと、生徒相互評価が一定の妥当性を有することを確認し、平均点が指導者評価に比べてやや高くなる傾向はあるものの、学習者相互評価の活用により、より妥当性の高い、多面的な発話評価につなげることができると示唆している。

3. リサーチデザイン

3.1 目的とRQ

本研究の目的は、日本人高校生による発話の流暢性を高める上で、自己評価・ピア評価・指導者評価の3条件のうち、いずれが最も効果的であるかを明らかにすることである。設定したリサーチ・クエスチョン(RQ)は以下の通りである。

RQ1 3種の評価条件の中で最も充実した評価フィードバックが得られるのはどれか？

RQ2 学習者全体において、3種の評価条件の中で最も安定的に流暢性を伸ばすのはどれか？

RQ3 上位群・下位群において、3種の評価条件の中で最も安定的に流暢性を伸ばすのはどれか？

3.2 協力者

調査協力者は、関西圏の国立大学附属中等教育学校の5年生(一般の高校2年生に相当)で、「英語表現Ⅱ」の授業を受講している者とした。なお、帰国生徒等、習熟度が特に高い者を除外した100名(男子49名、女子51名)である。協力者の英語力は一般的な高校生の平均分布に近い。一通り基本的な文法事項等を学び、かつ大学受験の直接的な準備学習の段階には入っていないため、高校生全体の発話力を見る際のサンプルとして適切であると判断した。

本研究では、学習者を4条件のグループ(無評価(統制群)、自己評価、ピア評価、指導者評価)に分割し、2016年4月下旬から5月にかけて、1週間に1回、3週間にわたって所定の発話活動を実施する(合計3回)。以下、各グループの人数、活動内容、英語力について示す。なお、英語力に関しては、調査と同時期(2016年4月)に実施したGTEC for STUDENTS(Benesse®)('読む','聞く','書く')の3技能の運用能力をスコア型の絶対評価で診断するテスト)のトータルスコア(810点満点)の平均値と標準偏差を示している。各群とも、上位群・下位群の人数は同数(ピア評価群は8名ずつ、その他は14名ずつ)である。

表 1 設定された 4 グループ

	無評価	自己評価	ピア評価	指導者評価
人数	28 (男 13/ 女 15)	28 (男 13/ 女 15)	16 (男 9/ 女 7)	28 (男 13/ 女 14)
活動内容	発話録音	発話録音→語数測定→自己評価する	発話録音→語数測定→ピア評価を読む	発話録音→語数測定→指導者評価を読む
英語力(全体)	494.3 (SD 110.8)	495.4 (SD 91.0)	492.8 (SD 91.7)	494.5 (SD 129.6)
(上位)	576.4 (SD 93.3)	573.1 (SD 47.4)	548.0 (SD 102.2)	594.4 (SD 99.8)
(下位)	412.1 (SD 47.6)	417.6 (SD 44.0)	437.5 (SD 24.9)	394.6 (SD 58.3)

本研究では、自己評価・ピア評価・指導者評価の 3 種の評価条件による学習者の発話量の変化を検証するわけであるが、仮にいずれかの評価によって学習者の発話量が増加したとしても、それだけでは、評価条件(自分で評価する／他者からもらった評価を読む)に起因する増加なのか、単に発話回数を重ねることで発話に慣れたための増加なのかが切り分けられない。そこで、無評価条件(発話行為のみ行う)を設定し、これを統制群としている。なお、ピア評価の人数が少ないのは、同じクラス内での相互評価とさせず、他クラス生徒(ピア)から評価を受けるようにしたためである。

4 グループは、それぞれ英語力が同等になるように編成されている。まず、各グループの英語力について、一元配置の分散分析を行った結果、4 群間に差は認められなかった ($F_{(3,96)} = 0.002$, $p = 0.99$)。次に、上位群のみ、下位群のみについて、それぞれ一元配置の分散分析を行った結果、やはり 4 群間の英語力に差は認められなかった (上位群 $F_{(3,46)} = 0.50$, $p = 0.68$, 下位群 $F_{(3,46)} = 1.47$, $p = 0.24$)。最後に、各グループ内の上位群・下位群の英語力について、t 検定を行ったところ、4 群すべてで有意な差が確認された(無評価 $t(26) = 5.87$, $p < .01$, 自己評価 $t(26) = 8.99$, $p < .01$, ピア評価 $t(14) = 2.97$, $p < .05$, 指導者評価 $t(26) = 6.47$, $p < .01$)。

3.3 調査の手順

実施に当たり、事前指導ではすべての協力者に対し、発話において発話量が重要であること、発話活動を(場合によっては自己評価を)毎回行うものの、ピア評価・指導者評価を含め、学校での評定には一切関与しないことを説明した。

次に作業の流れについて説明する。まず始めに、発話を録音するための IC レコーダーを全員に配布した。その後、自己評価グループ、ピア評価グループ、指導者評価グループにはそれぞれ専用の評価シート(名前欄、作業の指示、ワードカウンター、評価記入欄等が印刷された用紙で、群により内容は異なる)を渡した。

実験の冒頭に、指導者が口頭で各回のトピック(1回目は“Things I want to do this year.”, 2回目は “What I am going to do during Golden Week.”, 3回目は “How do you usually spend your free time?”)を提示した。いずれも学習者にとって身近な話題であり、特別な事前知識がなくても一定量の発話ができるようにトピックが選ばれている。その後各自に、話すべき内容について考える準備時間(30 秒間)を与えた。この間にメモを取ることは禁じていないが、筆者が観察したところ大半の学習者はメモを取ることはなく、発話の内容を考えてい

たと思われる。準備時間の終了後、指導者の合図で各自の IC レコーダーに向かって一斉に発話を開始した。1 分経過後に指導者の「止め」の合図で一斉に発話を終了した。センテンスの途中であっても指示があれば終了するようにあらかじめ指示をしている。

その後、無評価グループの作業は存在しないが、他の 3 グループは、自分の発話を自分で聞き直し、語数を測定する。この際、事前指導に基づき、先述の西(2010)によって開発されたワードカウンターを利用した。使用したワードカウンターを図 1 に示す。実際の作業の流れとしては、発話者は自分の発話を IC レコーダーで再生しながら自分の話す単語数に合わせて、指を 1,2,3...と移動させ語数を数え、全体の語数を計量する。この語数については、シートに記入するようにと指示されている。

(1) 自分の発話した語数を数えます。以下の、ワードカウンター えなさい。										
1	2	3	4	5	6	7	8	9	10	11
21	22	23	24	25	26	27	28	29	30	31
41	42	43	44	45	46	47	48	49	50	51
61	62	63	64	65	66	67	68	69	70	71
81	82	83	84	85	86	87	88	89	90	91
100	101	102	103	104	105	106	107	108	109	110

図 1 使用したワードカウンター(一部)

自己評価グループは、続けて、自分の発話について振り返りを行い、用紙に記入する。用紙には以下の指示が記載されている。

(3) もう一度改めて録音を聞きなさい。聞いた後、以下の指示に従って自己評価を記述しなさい。
 プラスポイント(強み)について:録音を聞いて、良かったと思うことを短く 3 つ挙げなさい(発音, リズム, テンポ, 内容, その他)
 改善ポイント(弱み)について:録音を聞いて、良くなかった点, 改善したい点を短く 3 つ挙げなさい。

図 2 評価シート(一部)

ピア評価グループについては、別のクラスの生徒に音声を聞かせ、上記と同等のシートにアドバイスを書かせる。なお、友人に対して否定的なコメントを書くことへの抵抗感を減らすため、上記の「改善ポイント」は「アドバイスポイント」と呼び変えている。指導者評価グループについては、教師が生徒の音声を聞き、上記と同等のシートにアドバイスを記入した。なお、記入欄の大きさは、3 グループとも同等である。3 グループとも、評価シートは、次の発話作業の前に指導者より発話者に返却された。

3.4 データ処理

以上により、評価コメント(無評価群は除く)と3回分の発話語数データが収集された。これらのデータセット

に対し、以下の処理を実施する。

RQ 1(評価フィードバックの差)では、3つの評価グループにおいてそれぞれ10人を無作為に抽出し、2回分の評価コメント(10人×2回×3群の合計60サンプル)を得る。その後、各評価グループにおけるコメントの平均文字数を算出し、量的な分析を行う。加えてそれぞれの評価グループより代表的なコメントを数例取り上げ、質的に解釈を行う。

RQ 2(全体の発話量変化)とRQ 3(上位群・下位群別の発話量変化)については、1～3回目の発話量平均値を求め、グラフを作成する。また、発話タスクの回数(1～3回)を説明変数、各回のタスクにおける平均発話量を目的変数として単回帰分析を実施する。合わせて、説明変数と目的変数の間の相関係数を求め、当該値を平方することで説明力(決定係数 R^2 値)を計算する。一般的には相関係数で0.7以上、説明力で49%以上の値が出た場合に当該の単回帰式は一定の説明力を持っていると判断するが、今回の研究では説明変数が3変数のみであることから、基準を厳しくし、相関係数で0.8以上、説明力で64%以上の値が出ていれば一定の線形的な変化傾向が生じていると判断する。その後、 R^2 が0.64以上のものに限って、回帰係数の大小を比較し、係数が高い方が発話の流暢性が向上していると判断する。

4. 結果と考察

4.1 RQ1 評価フィードバックの差

まず、各条件における評価フィードバック量(字数)の平均は以下の通りであった(表2)。

表2 各評価群のコメント平均文字数

	自己評価	ピア評価	指導者評価
平均	52.4	81.7	73.2
SD	14.6	26.4	17.1

一元配置分散分析を行ったところ、群間の差は有意で($F_{(2,57)}=11.3, p<.01$)、事後検定により、自己評価とピア評価、自己評価と指導者評価の間の差も有意であった(いずれも $p<.01$)。これにより、量的に最も充実した評価フィードバックが得られるのはピア評価および指導者評価であり、自己評価のフィードバックは量的に少なくなることが確認された。なお、標準偏差に示されるように、ピア評価の場合、フィードバック量は個人差が大きい。

次に、各条件におけるフィードバックの実例は以下の通りであった(表3)。

表3 評価群ごとのフィードバックの実例

評価群	評価例	コメント	
		1回目	2回目
自己評価	A	[プラスポイント] なるべく多くを話そうと努力した。	[プラスポイント] 文の形を意識できた。リズムが前よりもよくなった。
		[改善(弱み)ポイント]	[改善(弱み)ポイント]

		構文がバラバラになっている。文を考えるために何もしゃべられなかった時間があったし、リズムよくしゃべられなかった。	語彙が少ないので、上手く文章を表現できなかった。まだ文と文の間で詰まってしまう。
	B	[プラスポイント] ほんとに発音が悪すぎたというのはなかったと思う。なるべく自分なりに分かりやすい言葉で言えた。 [改善(弱み)ポイント] 即興で考えられなかった。内容が薄かった。語数が少ない。	[プラスポイント] 正しい英語は少しでも使えたと思う。また言いたいことが言えた。分からない言葉を分かる言葉に置き換えられるようになった [改善(弱み)ポイント] 文と文の間が長い。簡単な英語しか使えない。リズムも悪く内容も薄い。
ピア評価	C	[プラスポイント] 発音がとてもよくて英語慣れしてるのがとても分かりました。つなぎ言葉を時々使っているのがとても聞きやすかったです。目立って止まるがなくてとても良いテンポでした。 [アドバイスポイント] 僕のような英語が苦手な人が聞くにはもう少し適切な間をとってもらえるとありがたいです。	[プラスポイント] 発音素晴らしい。理由がしっかりしている。語彙が豊富。 [アドバイスポイント] 良すぎてコメントが思いつかないのですが、しいて書くなら、間があるといいかなと思います。
	D	[プラスポイント] 内容が分かりやすい。伝えようとしていてよかった。聞きやすい。 [アドバイスポイント] リズムに気を付けて、さらにすらすら話せるといいなと思います。	[プラスポイント] 発音すごくいい。聞き取りやすい速さ。内容がおもしろかった。 [アドバイスポイント] 「えー」っていうのを少なくした方がいい。
指導者評価	E	[プラスポイント]この 1 年間、勉強を頑張ろうという強いメッセージが伝わってきました。 [改善ポイント]多くの英文を音読して、スラスラ感に慣れましょう。授業で使った英文でいいです。家で毎日 5 分頑張ってみてください。きっと力がつきますよ！	[プラスポイント]発音がはっきりしてきて聞き取りやすくなりました！ [改善ポイント]毎日 5 分でいいので、授業で扱っている音声データを先生からもらってシャドーイングの練習をしてみましょう。さらにスラスラ感が出てくると思いますよ！
	F	[プラスポイント]言いたいことを頑張って英語で表現してみようという気持ちが	[プラスポイント]少しずつ語数が増えましたね！

		大きく伝わってきました！ 〔改善ポイント〕 言いたいことを短い文章でいいので伝えてみましょう。使う単語も難しく考えず、知っている易しい単語で大丈夫ですよ！	〔改善ポイント〕次は、気持ちを分かりやすく伝えるために、話す前に頭の中で話す内容を2つ位に絞ってみましょう！
--	--	---	--

自己評価の場合は、分量は少ないものの、「構文がバラバラ」「何もしゃべられなかった時間があった」「リズムよくしゃべられなかった」「語彙が少ない」「文と文の間で詰まってしまう」「内容が薄かった」「語数が少ない」「文と文の間の長さが長い」「簡単な英語しか使えない」「リズムも悪く内容も薄い」等、自分の発話を直接モニターすることで、とくに自分ができなかったことについて客観的に見つめ直している様子が見られる。

ピア評価の場合は、分量は多いものの、その大部分は相手を傷つけないようにという配慮に基づく文言であって(例:「僕のような英語が苦手な人が聞くには…」), プラスポイント, アドバイスポイントのいずれにおいても、具体的な内容が乏しい傾向が見られる。とくに、評価例Cに見られるように、評価者の側が学習者よりも英語力が低い場合には、相手に積極的なコメントを与えられず躊躇している様子もうかがえる。

一方、指導者評価の場合は、学習者に励みとなるコメントを与えているだけでなく、多くの英文の音読, シャドーイング, 話す内容を絞る, 短い文章で話す等、具体的な学習方法をアドバイスしている様子が見られる。

フィードバックの質の比較は困難だが、上記のサンプルに限って言えば、自己評価と指導者評価の内容が充実しており、ピア評価の内容はやや浅いといえることができる。

4.2 RQ2 全体の発話量の変化

4つの評価条件による3回の発話の平均値は下図の通りであった(図3～6)。

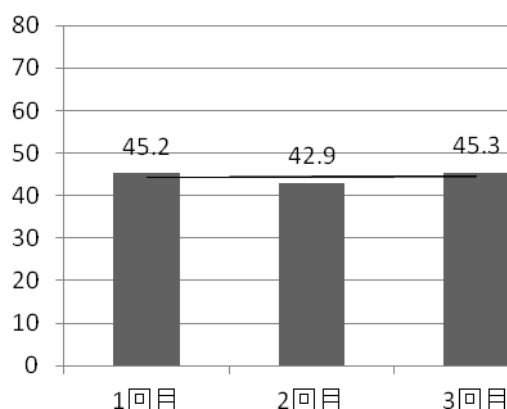


図3 無評価

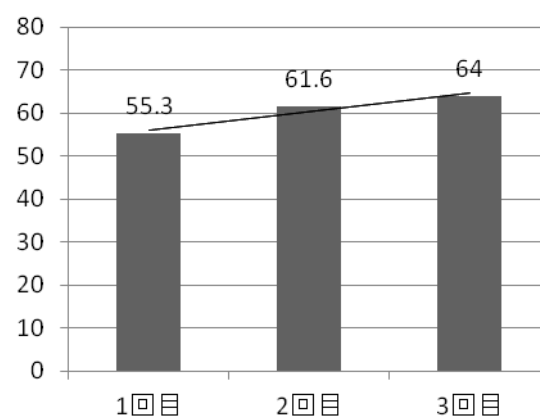


図4 自己評価

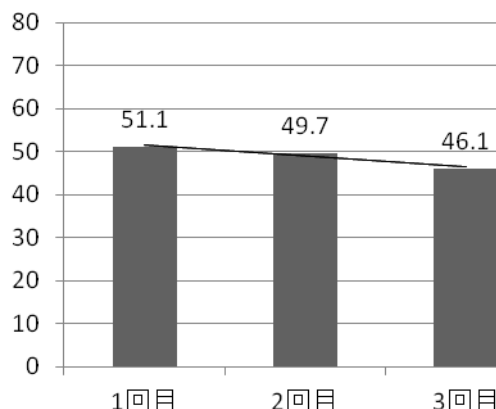


図5 ピア評価

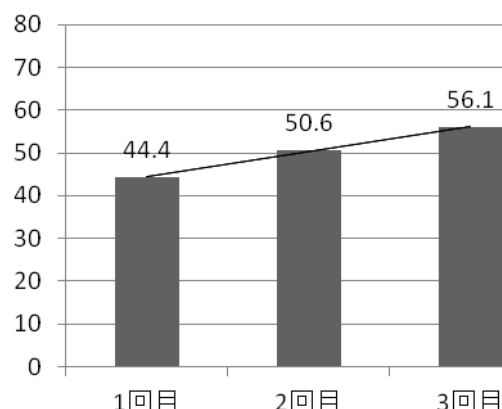


図6 指導者評価

回数(x)による発話量(y)の変化を線形的に示す回帰式はそれぞれ以下の通りである。

無評価 $y = 0.5x + 44.4$ ($R^2 = .001$)

自己評価 $y = 4.4x + 51.6$ ($R^2 = .94$)

ピア評価 $y = -2.5x + 54.0$ ($R^2 = .94$)

指導者評価 $y = 5.9x + 38.7$ ($R^2 = .99$)

決定係数は無評価を除くすべての群で64%を超えている。回帰式の係数は、指導者評価>自己評価>ピア評価であった。指導者評価と自己評価で正となり、ピア評価では負となっている。これより、(1)無評価(単に発話を反復するだけ)では発話量が増加しない、(2)指導者評価で最も多く発話量が増加し、自己評価がそれに次ぐ、(3)ピア評価では発話量が減少することが確認された。以下、各々について検討する。

まず、(1)については、従来、言われてきた「習うより慣れろ」というアプローチだけでは、発話の流暢性を高めることは難しいことが示された。また、統制群である無評価条件で有意な増減が確認できなかったことは、以下で見る3種の評価における発話量の変化が、評価行為そのものに起因することを示す。

次に(2)について、指導者評価で最も大きな流暢性の向上が認められた原因としてはいくつかの点が考えられる。1つは、今回の取り組みが評価に関係しないという説明があったにもかかわらず、協力者が評価評定を過度に意識し、「たくさん話せばいい評価がもらえるかもしれない」と考えた可能性である。ただし、そうであれば、学習者は初回から多く話していたはずで、回数を重ねるごとに発話量が順調に伸びていることをふまえると、やはり、「多くの英文を音読」、「スラスラ感に慣れ(る)」、「シャドーイングの練習」、「短い文章でいいので伝えてみ(る)」、「前に頭の中で話す内容を2つ位に絞ってみ(る)」といった具体的で詳細な指導者による評価フィードバックが一定の効果を挙げたと考えられるであろう。また、フィードバックの量においては他より少なかった自己評価で次に大きな流暢性の向上が認められた原因としては、分量は少なくとも、内省によって学習者が自身の発話の課題をより深く認識していたことが想定される。とくに、「文を考えるために何もしゃべられなかった時間があった」、「リズムよくしゃべられなかった」、「文と文の間で詰まってしまう」、「内容が薄かった」、「語数が少ない」、「文と文の間が長い」等はいずれも発話量に直結する内容で、こうした点の自覚化が発話量の向上をもたらしたのではないかと思われる。

一方、(3)について、量的に最も多くの評価フィードバックが得られたピア評価で発話量が減少しているのは興味深い事実である。この点については、自分の発話を友人に聞かれることで、「多く話そう」という気持ちよりも「英語の間違いをしてしまうと恥ずかしい」といった気持ちが強くなり、これが発話量を抑制させる方向で働いた可能性がある。合わせて、ピアは自分自身の聞き取り能力の制約もあり、「もう少し適切な間をとってもらえるとありがたい」、「しいて書くなら、間があるといいかなと思います」、「聞き取りやすい速さ」等、どちらかと言えば発話量を減らす方向での要望や助言を多く書く傾向があり、このことが発話量の抑制原因になっていた可能性も考えられる。

4.3 RQ3 上位群・下位群別の発話量の変化

4つの評価条件による3回の発話の平均値は下図の通りであった(図7～14)。

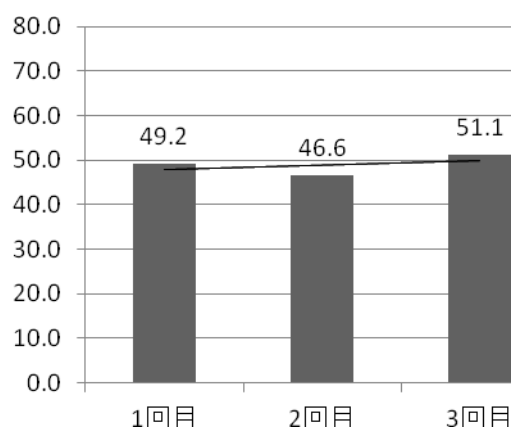


図7 無評価(上位群)

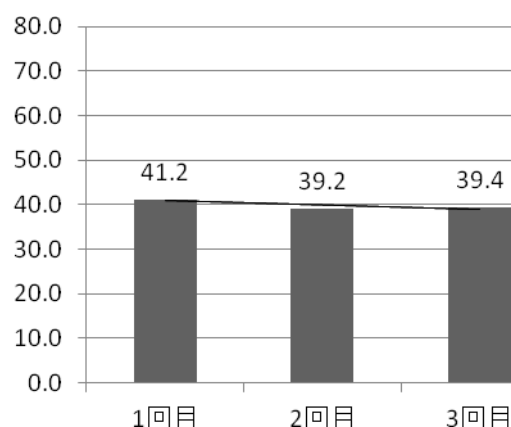


図8 無評価(下位群)

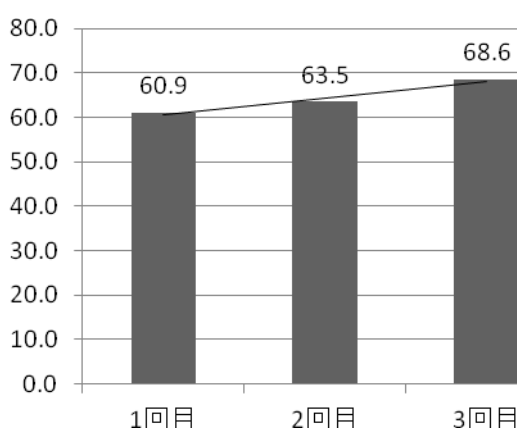


図9 自己評価(上位群)

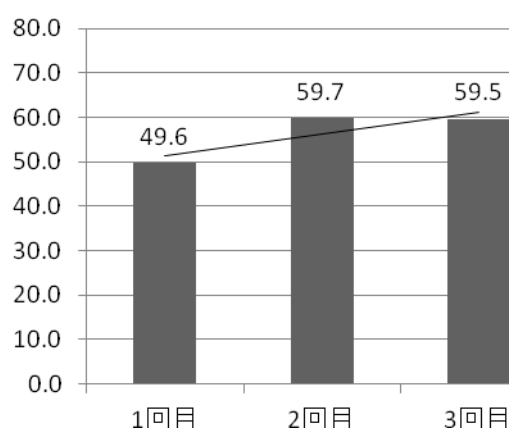


図10 自己評価(下位群)

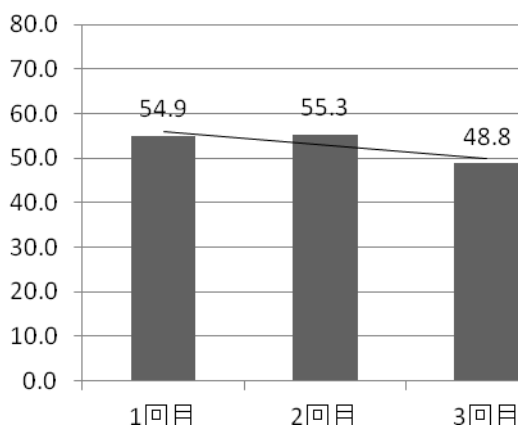


図 11 ピア評価(上位群)

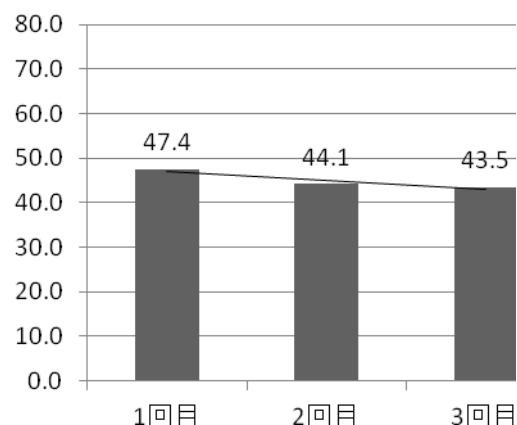


図 12 ピア評価(下位群)

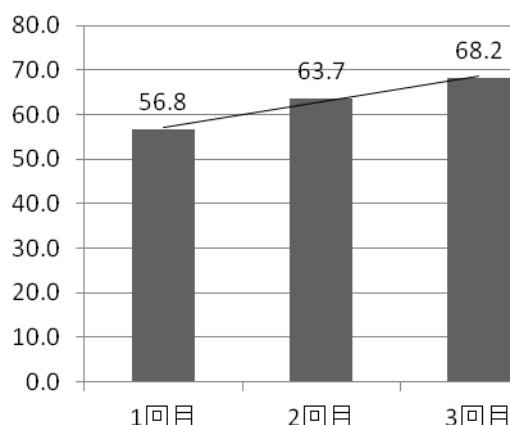


図 13 指導者評価(上位群)

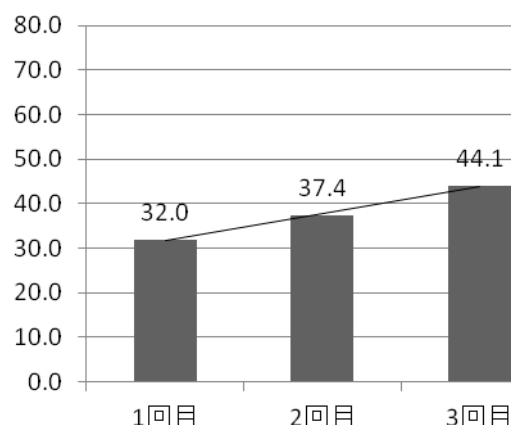


図 14 指導者評価(下位群)

回数(x)による発話量(y)の変化を線形的に示す回帰式はそれぞれ以下の通りである。

無評価	上位群	$y = 1.0x + 47.0$	$(R^2 = .18)$	下位群	$y = -0.9x + 41.7$	$(R^2 = .66)$
自己評価	上位群	$y = 3.8x + 56.7$	$(R^2 = .97)$	下位群	$y = 4.9x + 46.4$	$(R^2 = .73)$
ピア評価	上位群	$y = -3.1x + 59.1$	$(R^2 = .70)$	下位群	$y = -1.9x + 48.9$	$(R^2 = .87)$
指導者評価	上位群	$y = 5.7x + 51.5$	$(R^2 = .98)$	下位群	$y = 6.0x + 25.8$	$(R^2 = .99)$

決定係数は、無評価の上位群を除く全ての群で 64%を超えている。回帰式の係数は、上位群・下位群とも、指導者評価>自己評価>無評価>ピア評価の順で、ピア評価の上下位両群および無評価の下位群で係数は負になっている。これより、(1)上位群・下位群とも、指導者評価で最も多く発話量が増加し、自己評価がそれに次ぐ、(2)上位群・下位群とも、ピア評価では発話量が減少する、(3)下位群では無評価の場合に発話量が減少することが確認された。以下、各々について検討する。

まず、(1)については、すでに全体に対して確認された指導者評価および自己評価の効果が、上位群・下位群ともに有効であり、両群ともに最も効果があるのは指導者評価であることが示された。また、これらの評価

の効果は、特に下位群において顕著になる。これは、下位群では絶対的な発話量が少ないため、相対的に伸び幅が大きいためと考えられる。なお、自己評価の係数は、上位群では指導者評価に対してかなり開いているが(5.7と3.8で1.9ポイント差)、下位群では相対的に近接している(6.0と4.9で1.1ポイント差)。英語力の高い上位群は自分の発話について日頃から一定の内省的なモニターを行っているため、あえて重ねて自己評価を行っても変化の幅は小さいが、英語力の低い下位群は通常は眼前のL2の処理に追われ、内省的なモニターをあまり行っていないため、発話後に強制的に自己評価を行うことで、自身の課題をより強く意識できたのではないかと考えられる。

(2)については、すでにRQ 2で考察したことが上位群・下位群ともに同様にあてはまるものと考えられる。ただし、発話量の減少幅は上位群のほうが大きい。これは、上位群のほうが情意的なフィルターが高く、「文法や発音の間違いをすると恥ずかしい」という気持ちが大きいためと思われる。また、Yang *et al.* (2006)は、学習者間の英語力の差が大きい場合、相手に建設的なコメントを与えることができないと述べている。この点をふまえると、上位群のほうが、価値あるフィードバックをピアから得られていなかった可能性もあろう(RQ 1の評価例C参照)。

(3)については、英語力が一定の水準にない場合、発話の回数を重ねることによる「慣れ」が、流暢性を高めるというプラスの方向に寄与せず、むしろ、飽きて話さなくなるという否定的影響を及ぼす可能性を示している。

5. まとめ

以上、本研究では、自己評価・ピア評価・指導者評価という3種の評価が学習者の発話の流暢性を向上させるかどうかを検証してきたが、3つのRQについて以下の知見を得た。

まず、RQ 1(3種の評価条件の中で最も充実した評価フィードバックが得られるのはどれか?)については、量の点ではピア評価>指導者評価>自己評価となり、質の点では、自己評価≒指導者評価>ピア評価の順となることがわかった。

次に、RQ 2(学習者全体において、3種の評価条件の中で最も安定的に流暢性を伸ばすのはどれか?)については、指導者評価>自己評価>ピア評価の順となり、ピア評価の場合には発話量が減少することが確認された。

最後に、RQ 3(上位群・下位群において、3種の評価条件の中で最も安定的に流暢性を伸ばすのはどれか?)については、上位群・下位群いずれの場合も、指導者評価>自己評価>ピア評価の順となり、ピア評価の場合には発話量が減少することが確認された。

では、このような知見をどのように教育現場に活用していくことができるのであろうか。ここでは、4点に限って私見を述べたい。

1点目は、こうした評価を発話指導に組み込んでいくことの重要性についてである。今回の分析により、3週間程度の短期であっても、指導者評価や自己評価によって一定の流暢性の向上が得られることが示された。長期間にわたって取り組まないと効果が出ないのであれば、授業には組み込みにくい、短期の取り組みであっても一定の効果が出ることをふまえれば、こうした評価を積極的に指導に組み込んでいくべきだろう。

2点目は、こうした評価を可能にするハード面の整備の重要性についてである。今回の実践研究が可能になったのは、人数分の個人用ICレコーダーが使用できたためである。こうした機器の配備、あるいは、CALL

教室の活用等、学習者自身が日常的に自分の発話を録音して聞き直せる環境の整備が求められる。

3 点目は、指導法の組み合わせの重要性についてである。今回の分析により、習熟度を問わず、どの学習者に対しても、指導者評価が最も安定的に流暢性を伸ばすことが確認されたわけであるが、常にこうした丁寧な指導者評価を行っていくことは実際には容易ではない。今回の実践では、学習者 1 人につきフィードバックコメントの作成・執筆に約 2 分程度の時間がかかった。40 人規模のクラスで言うと、1 クラスにつき 80 分程度の時間が必要になる。1 人の指導者が多くのクラスを担当している現状では、これを毎行っていくことは現実的ではないだろう。その場合、指導者評価に別の評価を加えていくことになるが、この点については、ピア評価ではなく、自己評価が相対的に有効であろう。とくに、日頃、自分の L2 産出を振り返ることの少ない下位群の場合、まずは自己評価中心で指導を組むのが有効でありうる。

4 点目は、ピア活動の見直しの重要性についてである。今回の調査では、ピア評価によって発話量の減少が確認され、とくに上位群において減少幅が大きいことが分かった。このことは、評価に限らず、授業内で各種のピア活動を行わせる際に、その意味について指導者が注意深く検討する必要性を示している。対話的な学びを深める上でピア活動の意義が大きいことは言うまでもない。しかしながら、発話の流暢性を高めることを目的とした場合、ペアを組んで相手の発話語数を確認しコメントを与えるといった活動については、発話者の発話量を抑制する可能性も考えられるため注意を要したい。

今回の調査により、限られた時間の制約の中であっても、フィードバックを行う評価者を適切に選択することで、発話の流暢性を高めることが可能であることが示され、現場において、発話力の基盤となる流暢性を育てる指導を行うための実践のヒントが得られた。しかし、本研究には課題も残されている。今後は、調査協力者の情意面の調査(動機づけ、性格等)や、フィードバックコメントのより詳細な調査(形態素解析等)を組み合わせつつ、発話指導の実効性をあげる評価の在り方を引き続き検討していくこととしたい。

注

本研究は、JSPS 科学研究費補助金(奨励研究)16H00109 の助成を受けて第 1 著者が行っている研究プロジェクトの成果の一部である。教育実験の企画、データ処理、論文の作成には、神戸大学国際コミュニケーションセンター・附属学校英語教育連携推進会議「研究アドバイザー制度」に基づき、第 2 著者が協力した。

参考文献

- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, UK: Cambridge University Press.
- 藤原康宏・大西仁・加藤浩.(2007).「学習者間の相互評価に関する研究の動向と展望」.『メディア教育研究』, 4, 77-85.
- 深澤真.(2009).「スピーチにおける生徒相互評価の妥当性 ―項目応答理論を用いて―」.『STEP BULLETIN』, 21, 31-47.
- 石川慎一郎.(2016).「日本人学習者の L2 英語の発話量:母語話者及びアジア圏学習者との比較」.『日英言語文化研究』, 5, 15-26.
- 門田修平・泉恵美子(編著).(2016).『英語スピーキング指導ハンドブック』東京:大修館書店.
- 三宅ひろ子.(2013).「Online Audio Recording を利用した1分間英語スピーキング活動」.『全国英語教育学会 第 39 回北海道研究大会発表予稿集』, 190-191.

- 文部科学省.(2013).「第2期教育振興基本計画」. http://www.mext.go.jp/a_menu/keikaku/detail/1336379.htm
- 文部科学省.(2016).「平成27年度英語教育改善のための英語力調査について ―高校3年生の調査結果―」.(平成28年度「小・中・高等学校を通じた英語教育強化事業」に関する説明会配布資料3)
- 西巖弘.(2010).『即興で話す英語力を鍛える！ワードカウンターを活用した驚異のスピーキング活動22(目指せ！英語授業の達人)』.東京: 明治図書出版.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
- 菅沼洋子.(2013).「自己評価と他己評価を利用した自律的英語学習の探求」.『STEP BULLETIN』, 25, 168-185.
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15, 179-200.