

千葉大学CALL英語履修者によるシステム評価結果の予備的分析(2)

A Preliminary Analysis of Students' Evaluation of the CALL System at Chiba University, Part 2

土肥 充、竹蓋 幸生

1. はじめに

本論文は、千葉大学言語教育センター『言語文化論叢』第5号に採録された同名の報告(土肥2011)の続編である。前編には、2001年度から2009年度までの9年間の千葉大学の「CALL英語」履修者10,813名から回収されたすべてのデータを項目別にまとめたものを提示した。この39項目の設問は本論文の最後の補遺に示したが、前編でのデータの観察は教材についての印象(1~5)、受講後の印象(19, 20, 21, 23, 24)、それに学習時間(22)に関するもの、等についての大まかな分析であった。後編となる本報告では、残りの項目に対する回答を「コースウェア」と「ソフトウェア」、それに「語彙指導」に対するより詳細な分析的評価としてまとめ、さらに項目間のクロス集計を行うなどして「英語教育の総合システム」としての評価を加えた。

2. 研究の目的

本研究の目的は、土肥(2011)の続編として、千葉大学での11年間(2001~2011年)にわたる「CALL英語」の受講者(11,682名)による評価をいくつかの項目ごとに再編成し、また項目間のクロス集計を行い、今後の教材開発、評価法や授業の改善に資する形にまとめて報告することであった。

3. 研究の方法

3-1 データの収集

本研究では、千葉大学で開発された三ラウンド・システム(竹蓋1997; 竹蓋・水光2005)の構想に基づいて制作された教材を採用した「CALL英語」の授業のアンケート結果を分析するが、授業の方法、使用教材、アンケートの実施方法等は基本的に前回の研究(土肥2011)と同じである。ただし、前回の研究では2001年から2009年の9年を対象としたが、本研究で新たに加えた一部の分析には2010年度および2011年度前期のアンケートも対象に含めた。評価のためのデータは、39項目からなる質問への回答で、実質2~4か月の授業後に履修者全員から回収した英語学習の効果、CALLによる英語学習・教育システムの構造、等についての印象評価のデータを収集したものである。

データ収集の方法としては、自由筆記によるアンケートも含めたが、今回の報告で分析

の対象としたデータは精神計量学の一手法である「等現間隔法」(Equal-Appearing Intervals)を採用して質問紙方式で収集したものである。

3-2 データの分析

「等現間隔法」を採用して収集されたデータは、項目ごとに計算された平均値と、いくつかの項目をまとめた項目群への再編成により得られたデータの平均値を求める形での分析やいくつかの項目間のデータのクロス集計で観察された。

3-3 データの比較、解釈

等現間隔法では、1～5と区切られた評価の基準が、「5：非常によい、4：よい、3：どちらでもない、2：悪い、1：非常に悪い」と区別されることが多い。しかし、平均値を計算すると5と1の間のあらゆる数値が出てくるので境界を決めるのが難しい。そこで、本研究では平均値の形で得られた分析結果の比較を、以下に示した基準を独自に作成して、わかりやすい形に表現し直し、数値と併記して報告した。

履修者による評価（5～1）の項目毎の平均値（ \bar{x} ）の解釈基準

$5 \geq \bar{x} \geq 4$ 非常によい

$4 > \bar{x} \geq 3$ よい

$3 > \bar{x} \geq 2$ 悪い

$2 > \bar{x} \geq 1$ 非常に悪い

また、必要に応じて、統計学の一手法であるt検定の手法を採用して平均値間の有意差の有無の検定も行った。t検定を実施した際は、人数が多いことを考慮して有意水準を原則1%とし、有意差がある場合は**の記号をつけたが、5%水準で有意の場合も参考までに*の記号をつけ、有意差がない場合はn.s.と表記した。数値はいずれも丸めの誤差を含む。

収集されたアンケートデータやその分析、比較のために使われた統計的手法の妥当性の検証は以下の2種の方法で行われた。

- 1) 項目2の質問に対して異なる回答をした二履修者群の質問項目19, 20, 21, 23, 24への回答の平均値の差にはすべて有意差が認められたが、「3」と回答した履修者群全員を奇数番と偶数番の二履修者群に分けて上記5項目への回答のそれぞれの平均値の差をt検定により比較した結果、いずれにも有意差がなかった。
- 2) 本論文の表7および表8にまとめた27項目の設問に「1～3と回答したグループ」と「4または5と回答したグループ」が上記5項目の設問に回答した結果の平均値を比較した結果、すべてのケースで有意差のあることが確認された。

4. 結果

4-1 学習・指導の効果についての評価

まず、履修者が「学習の効果」についてはどのような印象を持ったかを観察した。全39件の質問項目のうち、5件(19, 20, 21, 23, 24)の項目を履修者がCALLによる「英語学習・指導の効果」を評価した項目と捉えて、その平均値をまとめたものを表1に示した。

表1 千葉大学の「CALL英語」による英語学習・指導の効果の評価

項目	設問内容	評価	回答の割合(%)		
			5~4	3	2~1
19	成就感(教材理解力の向上)	4.2	81	12	7
20	成就感(応用力の向上)	3.9	69	22	9
21	情意面(学習が楽しい、満足感)	4.0	72	18	10
23	継続学習意欲(もっとこの種の教材で学習したい)	4.1	75	17	9
24	総合システムとして(この授業を取ってよかった)	4.4	85	11	4
平均		4.1	76	16	8

(表1のデータは2001年から2010年までの10年間のものである)

上記、5質問項目への評価値の平均は4.1であり、この平均値を「 \bar{x} 」として、「データの比較、解釈」の項に示した換算表で見ると千葉大のCALLシステムがまず効果の面から「非常によい」と、履修生に高く評価されていることがわかる。この評価の基になっている「回答の割合」で見ても、「5」と「4」と回答した学生があわせて76%で、「2」と「1」が8%しかないという事実は、公表されている東京大学や京都大学の学生の教養の授業、外国語の授業に対する満足度の28%、30%(小林他2001:165-166;水光2002:13-14)と比べるまでもなく、大学の英語教育として千葉大学のCALL英語による指導が成功していることの証として受け入れてよいものであろう。

項目別に見ると、項目20の「成就感(応用力の向上)」が唯一3.9で、1ランク下の「よい」の評価であるが、これは実質2~4か月という短い期間の学習への評価であることを考えれば、長い期間の学習が要求される外国語学習の過程としては、むしろ当然の評価と言えよう。どちらかと言えば、Krashen(1981)が仮説として提唱する少し難しい教材(i+1)を与えられながら、半年足らずの学習で教材が聞けるようになったという項目19への4.2の評価が注目すべき結果であると考ええる。

さらに特筆すべきは、授業全体への評価(項目24)が4.4(「5」と「4」の回答を合わせて85%)と最高の評価であったことである。CALLというと、All or Nothingの考え方をする教員が少なくなく、良いも悪いも機器、ソフト、教材しだいと単純にとらえられることが多いのであるが、千葉大のCALLでは、きめ細かな対応、情熱、わかりやすい説明、学習環境の向上、等を目指した担当教員の努力がCALLシステムと組み合わせられ、教育の「総合システム」として非常に高い評価がされたと考ええる。

4-2 CALLシステムの分析的評価、その1

次に、本研究では、CALLシステムの分析的評価のために、設問をカテゴリー別に大分類、中分類、小分類等に再編成してそれぞれに評価の平均値を求め、その結果を以下にまとめて示した。分類の基準となったカテゴリーは、大枠で、「聴解力養成システム」、「語彙力養成システム」、「学習時間」の3枠に分けられるので、それぞれのデータを表2、表3、表4として分けて示した。この部分の分析に使用されたデータは、土肥（2011）で使用されたものと同じ、2001年度から2009年度までのものである。

表2を一瞥してまず明らかになることは、千葉大学のCALL英語で採用されている「聴解力養成システム」が、4.1という評価結果から推定できるように、全体として「非常によい」と、好意的に受け入れられていることである。なかでも、履修者全員の平均値で見ると学習作業の直後に得られる、いわゆるKR(Knowledge of Results) 情報と呼ばれるもの（フィードバック：即時型）の価値を高く評価していることが、4.7という高い評価値から推定できる。同じフィードバック情報でも各ユニットの学習後と、やや遅れて実施される遅延型のユニットテストの効果については、決して悪くはないが、即時型に比べればやや低い、4.0という評価であった。

表2 アンケート項目の分類別評価の平均値（聴解力養成システム）

全体平均	中分類	中分類平均	小分類	小分類平均	設問内容	項目	項目平均
4.1	素材	3.8			興味	1	3.8
	コースウェア	3.9	各種情報	3.9	指示は明確	7	4.0
					事前情報（写真）	6	3.9
					ヒント情報	8	4.0
					参考情報（辞書）	9	4.4
					解説記事	12	3.9
					Key Sentences	13	4.0
					Column	14	3.9
					Tips & Hints	15	3.6
					Coffee Break	16	3.9
					発展情報	17	3.4
	フィードバック	4.4	即時型	4.7	正解の提示	10	4.6
					正解の英文	11	4.7
			遅延型	4.0	ユニットテスト	18	4.0
	ソフトウェア	4.1	構成	4.2	画面構成デザイン	25	4.2
			ボタン	4.0	ボタン配置操作性	26	4.0
			文字	4.5	文字の大きさ	27	4.5
文字の配色					28	4.4	
写真			4.1	写真の大きさ	29	4.2	
				写真の枚数	30	4.0	
				写真の画質	31	4.1	
動画	3.9	動画の画質	32	3.9			
避負荷過大	3.9	音声文字タイミング	33	3.9			

ソフトウェア関連の評価では、意外とも言えるが、文字の大きさや配色が4.5と高く評価され、指示や説明の読みやすさが評価されていることがわかる。ソフトウェア全体としては全体平均と同じ4.1の評価で、「大変によい」との評価であり、とくに悪いと評価された項目もなく、使いやすさが評価されたものと考えられる。音声や文字の提示のタイミングのずれは、マルチメディア情報の同時提示による負荷過大を避けるために導入された独自の手法であるが、動画の画質とともに3.9と、ソフトウェア関連項目の中では低い評価しか得られなかった。これは、前者が個人差への対応が難しいこと、後者はメディアのサイズの制約のため、どちらも教材作成時点の技術水準を見れば致し方がないとも言えるが、反省材料として銘記したいと考える。

コースウェアについては、全体として3.9と、フィードバックやソフトウェアに対する評価よりやや低かったことは残念である。自由筆記では高く評価されている、定着度に影響のある「分散学習」や学習の負担を軽減する「学習タスク間の有機的関連」の効果等に関する設問がなかったことが影響しているかもしれない。この項目では辞書情報が4.4と比較的高く評価され、ヒント情報とKey Sentencesの提示がそれに次いで高く評価された。辞書情報の評価の高さは、文字だけでなく、音声による提示を含んだものであること、熟語の提示も含まれていること、それに提示される意味が素材中で使われているものに限られている実用性などが評価されたと考える。発展情報（応用表現）の評価がこのカテゴリの中では最低であったが、これは教材の理解に直接役に立つものではなく、応用力の向上のために含まれているのであるが、受講生全員には長い目で見た有用性にまでは気づけなかったのかもしれない。効果的な学習作業のなかで指示の明確さも不可欠な要件であるが、それも4.0と高く評価されていることは、他とは異なる面の多い教材の制作者にとって朗報である。

語彙力養成システム（表3）については、全体として3.9、つまり「よい」という評価であり、聴解力養成システムよりは少し低い評価であった。音声や用例のなかで学べることに對しては4.2と、高い評価を与えていながら、マルチメディア情報の同時提示を避けて、負荷過大を抑えるという温情的気配りが仇になったようである。授業に独創的な学習作業を導入する場合には学習者に、オリエンテーション等で、その内容や導入理由について十

表3 アンケート項目の分類別評価の平均値（語彙力養成システム）

全 体 平 均	中分類	中分類 平 均	小分類	小分類 平 均	設 問 内 容	項 目	項 目 平 均
3.9	コース ウェア	4.2	音 声	4.2	音声の使用	34	4.2
			用 例	4.2	用例の使用	35	4.2
	フィード バック	3.8	内 容	3.8	語彙テスト内容	38	3.8
			回 数	3.7	語彙テスト回数	37	3.7
	ソフト ウェア	3.6	避 負 荷 過 大	3.6	音声文字タイミング	36	3.6

分な説明が必要だということがこのデータから学べる。語彙テストの内容や回数も評価はあまり高くなかったが、履修者全員の目で見るとフィードバック情報の価値よりもテストは嫌という気持ちが勝ったのかもしれない。

表4 アンケート項目の分類と平均値（学習時間）

設問内容	項目	項目平均
聴解学習時間／週	22	2.7
語彙学習時間／週	39	1.6

学習時間（表4）については、聴解力養成に週平均2.7時間、語彙力養成に1.6時間かけたということは、自習も含めて、一週間に4時間も英語の学習にかけていることになる。このこと自体は一般的な傾向から見れば多いほうと言えるのではないだろうか。

聴解力養成システムの学習時間については、5時間、3時間、1時間と答えた履修者が「学習効果の項目（19, 20, 21, 23, 24）」にどう回答したかを観察した。土肥（2011）と一部重複するが、結果は表5にまとめて示した。学習時間5時間と3時間と回答した履修者がどちらも項目19（教材の理解度）で4.3と回答しており、有意差がないことが検証されたが、他はすべての比較で有意差が確認され、学習に時間をかけることの必要が明らかにされた。言い方を変えれば、教材の理解は週3時間でできるようになるが、応用力の向上を含む他の面では週5時間、つまり週末以外は毎日1時間の学習、が外国語の習得には必要らしいということが見えてくる。

表5 聴解学習時間が学習効果に与える影響の例（土肥2011に追加）

時間	5時間	3時間	1時間	5時間／1時間	5時間／3時間	3時間／1時間
人数	263	3947	1429	t		
項目19	4.3	4.3	4.1	2.9**	1.5n.s.	3.9**
項目20	4.1	3.9	3.8	5.2**	4.4**	3.4**
項目21	4.3	4.0	3.8	6.7**	5.4**	4.3**
項目23	4.5	4.1	3.9	8.5**	6.7**	7.0**
項目24	4.6	4.4	4.2	6.6**	5.3**	5.2**

**p < .01

語彙力養成のための学習時間については、聴解力養成にかけた時間に比べて短く、平均で週1.6時間であったが、表5と同様の分析結果を表6に示した。こちらでも1時間では不十分で、少なくとも2、3時間の学習が必要であると言える。

表6 語彙学習時間が学習効果に与える影響の例

時 間	3時間	2時間	1時間	3時間/1時間	2時間/1時間
人 数	1084	3613	5576	t	
項目19	4.3	4.3	4.2	2.5*	3.6**
項目20	3.9	3.9	3.8	3.4**	3.6**
項目21	4.1	4.0	3.9	6.5**	6.3**
項目23	4.1	4.2	4.0	3.1**	5.8**
項目24	4.4	4.4	4.3	4.1**	5.0**

*p < .05 **p < .01

4-3 CALLシステムの分析的評価、その2

本研究では、上記4-2の観察を一步進めて、各質問項目に「1, 2, または3と（否定的または中立的に）回答した学生群」と「4または5と（肯定的に）回答した学生群」のそれぞれが「教育の効果を示すと定義された項目（19, 20, 21, 23, 24）」のそれぞれにどのように回答したかを平均値で比較、観察した。この結果は表7および表8に示したが、さらにこの表の各項目に対応する「5個の平均値」の平均値を計算して各項目の評価を観察すると興味深い傾向が見えてくる。

表7 聴解指導関連の設問に否定的回答と肯定的回答をした履修者群の効果関連項目への評価

分類	項目	設問内容	1～3と回答した者の各項目平均値						4～5と回答した者の各項目平均値					
			項目19	項目20	項目21	項目23	項目24	平均	項目19	項目20	項目21	項目23	項目24	平均
指示	7	指示は明確	4.0	3.6	3.6	3.9	4.0	3.81	4.3	4.0	4.1	4.2	4.5	4.21
コース ウェア	6	事前情報（写真）	4.0	3.6	3.6	3.9	4.1	3.83	4.3	3.9	4.1	4.2	4.5	4.20
	8	ヒント情報	4.0	3.7	3.7	3.9	3.9	3.84	4.3	3.9	4.0	4.1	4.4	4.17
	9	参考情報（辞書）	4.0	3.7	3.6	3.8	4.1	3.82	4.3	3.9	4.0	4.1	4.4	4.14
	12	解説記事	4.1	3.7	3.6	3.8	4.1	3.84	4.3	4.0	4.1	4.2	4.5	4.21
	13	Key Sentences	4.1	3.6	3.6	3.7	4.0	3.79	4.3	3.9	4.1	4.1	4.4	4.16
	14	Column	4.1	3.6	3.6	3.6	4.0	3.77	4.3	3.9	4.1	4.2	4.4	4.19
	15	Tips & Hints	4.1	3.7	3.7	3.9	4.2	3.92	4.3	4.0	4.2	4.3	4.5	4.26
16	Coffee Break	3.9	3.6	3.5	3.7	4.0	3.75	4.1	3.9	4.1	4.3	4.5	4.18	
17	発展情報	4.1	3.7	3.7	3.9	4.2	3.91	4.3	4.0	4.2	4.3	4.5	4.28	
フィード バック	10	正解の提示	3.9	3.5	3.5	3.7	3.9	3.69	4.3	3.9	4.0	4.1	4.4	4.13
	11	正解の英文	3.8	3.5	3.5	3.7	3.9	3.67	4.3	3.9	4.0	4.1	4.4	4.12
	18	ユニットテスト	3.9	3.6	3.6	3.8	4.1	3.80	4.3	4.0	4.1	4.2	4.5	4.20
ソフト ウェア	25	画面構成デザイン	4.0	3.6	3.5	3.8	3.9	3.74	4.3	3.9	4.1	4.2	4.5	4.18
	26	ボタン配置操作性	4.1	3.7	3.7	3.9	4.1	3.88	4.3	3.9	4.0	4.2	4.4	4.18
	27	文字の大きさ	3.9	3.5	3.6	3.7	3.9	3.73	4.3	3.9	4.0	4.1	4.4	4.14
	28	文字の配色	4.0	3.6	3.6	3.8	4.0	3.78	4.3	3.9	4.0	4.1	4.4	4.15
	29	写真の大きさ	4.0	3.7	3.7	3.9	4.1	3.87	4.3	3.9	4.0	4.1	4.4	4.15
	30	写真の枚数	4.1	3.7	3.7	4.0	4.2	3.92	4.3	3.9	4.0	4.1	4.4	4.17
	31	写真の画質	4.1	3.7	3.7	3.9	4.1	3.89	4.3	3.9	4.0	4.2	4.4	4.17
	32	動画の画質	4.1	3.7	3.7	3.9	4.2	3.95	4.3	3.9	4.0	4.2	4.4	4.17
33	音声文字タイミング	4.0	3.7	3.7	3.9	4.1	3.88	4.3	4.0	4.1	4.2	4.5	4.20	

表8 語彙指導関連の設問に否定的回答と肯定的回答をした履修者群の効果関連項目への評価

分類	項目	設問内容	1～3と回答した者の各項目平均値						4～5と回答した者の各項目平均値					
			項目19	項目20	項目21	項目23	項目24	平均	項目19	項目20	項目21	項目23	項目24	平均
コース ウェア	34	音声の使用	3.9	3.4	3.4	3.7	4.0	3.68	4.3	4.0	4.1	4.2	4.5	4.20
	35	用例の使用	3.9	3.5	3.4	3.6	3.9	3.66	4.3	3.9	4.1	4.2	4.5	4.19
フィード バック	38	語彙テスト内容	4.0	3.6	3.6	3.9	4.1	3.84	4.3	4.0	4.1	4.2	4.5	4.24
	37	語彙テスト回数	4.1	3.7	3.6	3.8	4.1	3.85	4.3	4.0	4.2	4.3	4.6	4.26
ソフト ウェア	36	音声文字タイミング	4.1	3.7	3.7	3.9	4.2	3.92	4.3	4.0	4.1	4.2	4.5	4.23

まず、表7および表8の最右列にある「4、5」と回答したグループの「平均値の平均値」を項目毎にすべて比較する。これは、聴解力養成システムおよび語彙力養成システムの要素に関する各項目の質問に対して「非常によい」、または「よい」と評価したグループがその印象で見て学習・教育の効果はどれだけ大きかったと回答したかを示す指標である。これを表7で見て評価の高い順に4.2以上を並べてみると以下の通りであった。

- 4.28 項目17 「発展情報は役に立った」
- 4.26 項目15 「Tips & Hintsは役に立った」
- 4.21 項目7 「指示は明確であった」
- 4.21 項目12 「解説の記事(★)は理解の役に立った」
- 4.20 項目6 「写真、イラストは学習の助けや励みになる」
- 4.20 項目18 「ユニットテストは役に立った」
- 4.20 項目33 「音声、文字等提示のタイミングは適切だった」

同様に表8で、平均値の大きい順に4.2以上を並べてみると以下の通りであった。

- 4.26 項目37 「語彙テストの回数は適切だった」
- 4.24 項目38 「語彙テストの内容は適切だった」
- 4.23 項目36 「音声、文字等提示のタイミングは適切だった」
- 4.20 項目34 「音声を使用した語彙教材は記憶に残る」

上記2つのリストから推定できることは、相対的に見て、発展情報(応用表現)やTips & Hints(音声学的、言語学的情報および英語学習のための助言)、それにユニットテストのような「教材による学習作業の実践に対して必ずしも即効薬とは言えない情報」、情報のトップダウン処理を助ける写真・イラスト等の事前情報を軽視しない学習者、「語彙テスト」を嫌がらない学習者、そして、「能力に対する負荷過大を避ける必要」等が理解できる学習者に高い学習効果の向上が得られているということである。

続いて、表7と表8それぞれの左右二つのグループの「平均値の平均値の差」を観察した。これは、差の大きい項目が各項目で「高く評価したグループ」が「低く評価したグループ」より「学習効果の向上のために上手に活用した項目(システムの特徴)」であることを示すからである。それは、その値の大きい順に6位まで列挙すると以下の通りであった。

- 0.53 項目35 「用例を使用した語彙学習は役に立つ」
- 0.52 項目34 「音声を使用した語彙教材は記憶に残る」
- 0.45 項目11 「正解部分の英文表示は理解の役に立った」

0.45 項目25 「画面構成、デザインは適切であった」

0.43 項目16 「Coffee Break は役に立った」

0.43 項目10 「正解の提示は理解の役に立った」

上記リストから「語彙学習の方法」「言語記号（形式と内容）習得のためのKR情報」「視覚的環境」「異文化情報」（Coffee Break）を重視する学習者はそうでない学習者より高い学習効果が期待できると推定できる。

4-4 クロス集計による観察

土肥（2011）では、アンケートの各項目に対してされた一部の評価の間でクロス集計による観察を行い、「素材に興味」を持った受講生と「学習時間が長い」受講生は学習効果が高いと評価している事実が報告された。続編の本研究では、このような観察をさらに進めて、学習方法への興味の多寡（項目21「CALLでの学習は楽しかった」）、教材理解力向上の多寡（項目19「Step 1, 2, 3と進むにつれ聞けるようになった」）、それに応用力向上の多寡（項目20「聞き取りの力がついたと思う」）のそれぞれと継続学習意欲の関係を調査した。その結果、表9～表11に見られるように、いずれの場合も、興味や成就感の持てる学習が継続学習意欲につながることを示す結果が得られた。

表9 学習方法への興味の多寡と継続学習意欲の関係

項目21回答	人数	項目23平均	t
5	3804	4.6	38.9**
1	367	2.7	

**p < .01

表10 成就感（教材理解力向上）の多寡と継続学習意欲の関係

項目19回答	人数	項目23平均	t
5	5412	4.3	12.9**
1	247	3.4	

**p < .01

表11 成就感（応用力向上）の多寡と継続学習意欲の関係

項目20回答	人数	項目23平均	t
5	3065	4.4	17.9**
1	293	3.3	

**p < .01

外国語の習得には長い時間の必要なことは周知の事実である。とすれば、継続学習意欲をもたせることは授業・学習の成功にとって重要な条件のひとつである。にもかかわらず、2003年の国立教育政策研究所教育課程センターの調査で、中等教育の現場（中学校、高等学校）では学年が進むにつれて教科としての英語に対する好意的感情が低くなっていくという憂うべき実情が田地野（竹蓋・水光2005）によって報告されている。そのような中で、CALL英語という、興味を持てる教材が使用でき、高い成就感（教材理解力、応用力とも）の得られる学習・教育法が本学に存在することは誇りに思っよよいものと考えられる。

次に、結論に変わりはないのだが、後編の作成時に、土肥（2011）の表5に使われたデータに不備が発見されたため、難易度関連のデータの分析を以下に見られるような形で再度行うこととした。さらに本研究では、データの分析の前に、項目2と項目3の質問と回答の意味について、他の項目の質問（有効であった、役に立った）と異なること、それに以下の表12に見られるような、二つの質問項目で一つのスケールの半分ずつをカバーするものであるとの再確認も行った。

表12 項目2と項目3の質問に対する回答の内容と推定される受講生群

項目	回答	回答の内容	推定される受講生の類型
3	5	極端に難しい	習熟度が不足している
	4	たいへん難しい	
	3	難しい	
	2	少し難しい	
	1	どちらかと言えば難しい	
なし	0	中立（適切な難易度）	普通の学生
2	1	どちらかと言えば易しい	
	2	少し易しい	
	3	易しい	
	4	たいへん易しい	
	5	極端に易しい	習熟度が高い

次に、難易度の評価と学習効果の評価関連の素データを表13として以下に示す。表13のデータを一瞥してまず言えることは、難易度の極端に高い教材（項目3に4、5と回答した受講生）が難易度の極端に低い教材（項目2に4、5と回答した受講生）よりも効果の評価（項目19と20）が低いということである。習熟度が不足する受講生に難しすぎる教材が割り当てられ、学ぶものが多すぎるということは、学べるものがないに等しく、絶望感、嫌悪感まで引き起こしてしまうのであろう。他方、極端に易し過ぎる素材でも実力者にとっては復習にはなるので、極端に難しすぎる素材よりは効果が感じられるのかもしれない。この推定を裏付けるように、極端に易し過ぎる素材（回答4、5）が、学習の楽しさ（21）、継続学習意欲（23）、授業の総合的評価（24）の項目で、極端に難しすぎる素材（回答4、5）

よりもはるかに高く評価されている。これは、専門でない授業で「単位が楽に取れる」という感覚と、新しい学習には結びつかなくてもとにかく「授業が分かる」というプラスの感覚が合わさっての評価結果と推定される。

表13 難易度の評価によって分類した効果関連5項目の平均値

項目	回答	難易	人数	項目19	項目20	項目21	項目23	項目24	全平均
3	5	難しすぎ	297	3.7	3.3	3.5	3.7	4.0	3.7
	4		2162	4.1	3.7	3.8	3.9	4.3	4.0
	3		4187	4.2	3.9	3.9	4.0	4.3	4.1
	2		1959	4.3	3.9	4.0	4.2	4.4	4.2
	1	1938	4.4	4.1	4.1	4.3	4.6	4.3	
2	1		3031	4.2	3.9	4.0	4.2	4.5	4.2
	2		2856	4.2	3.8	3.9	4.0	4.3	4.0
	3		3336	4.2	3.9	3.9	4.0	4.3	4.1
	4		1096	4.3	4.0	4.0	4.2	4.3	4.2
	5	易しすぎ	224	4.3	3.9	4.0	4.3	4.5	4.2

しかし、ここで見過ごしてならないことは、「少し難しい素材（項目3への回答1）」に対する効果の評価（全平均4.3）には他のどの評価項目でも及ばないという事実である。その度合いが極端な場合、難しすぎる教材よりも易し過ぎる教材の方が効果が出るが、少しならば難しすぎる教材が最善で、その方が易し過ぎる教材よりも効果がでると考えているのである。何事も過ぎたるは及ばざるが如しなのであろう。困難なことではあるが、指導に当たって「少し難しい素材」を選定、採用することの重要性は明らかである。

上記推定の検証のため、項目2で「1」と回答した受講生群と項目3で「1」と回答した受講生群が効果関連5項目のそれぞれにどのように回答したか、さらに項目2で「2」と回答した受講生群と項目3で「2」と回答した受講生群が効果関連5項目のそれぞれにどのように回答したか（表13太字データ参照）を比較した。また、その平均値の有意差もt検定で検証した。その結果は表14として以下に示した。

表14 「少し難しい素材」と「少し易しい素材」を使用した場合の教育効果の比較

t検定対象の項目と回答（表13参照）	項目19	項目20	項目21	項目23	項目24
回答2の平均値の差（項目2と3）のt検定	6.9**	4.3**	2.4*	6.2**	3.9**
回答1の平均値の差（項目2と3）のt検定	5.5**	5.9**	4.0**	5.2**	3.8**

*p < .05 **p < .01

表13と表14から、項目2と項目3に「1」と回答した受講生群、「2」と回答した受講生群は、両群とも、少しならば難しいもののほうが易しいものよりも聴解力養成用教材の素材として高い学習効果が得られると評価したことがわかる。ところが、表13を見ると、項目2で「1または2」と回答した受講生は10,543名中5,887名（55.8%）で、項目3で「1

または2」と回答した受講生が3,897名（37.0%）であったことから、これまでの本学の学習者が使用してきた教材は、どちらかと言えば、少し易し過ぎた可能性があるということになる。学習者の習熟度も教材の難易度もその推定は容易ではないが、その適合度が明白に教育効果に影響があると学習者が表明している以上、今後はその最適な組み合わせの推定のための研究にも目を向けたいと考える。

他方、項目2と3で「4または5」と回答した学生が1,320名と2,459名で12.5%と23.3%にのぼる（表13）。これは適度な難易度の教材が存在しないか、または既存の教材の選定の誤りであるが、そのような学生に対してこれからも適切な難易度の教材を作成、または選定して提供し続けることの必要が見える。

4-5 各種属性別の傾向

本研究では、アンケートの項目別や項目間の分析だけでなく、各種属性を基準にデータを再分類して、以下のように効果関連5項目の平均値を求めて、その差の有意差の有無も調査した。いずれもアンケート内容以外の種々の要因が直接または間接的に影響している可能性があり、以下の数値だけを見て簡単に結論を出すことはできないが、今後の参考のために示しておく。

4-5-1 指導実践の前半5年間と後半5年間の指導効果の比較

2001～2005年度と2006～2010年度に分けて効果の差を調べ、表15に示した。

表15 指導の前半5年間と後半5年間の効果の比較

項目	2001～2005年度		2006～2010年度		t
19	4.3	>	4.2		2.7**
20	3.9	>	3.8		3.0**
21	4.0	>	3.9		2.7**
23	4.2	>	4.0		11.0**
24	4.4	>	4.3		8.6**

不等号は1%水準で有意の場合のみ

**p < .01

同一の指導理論に基づきながら、これまで10年以上にわたって教材開発の努力を継続して教材数が増え、授業方法も少しずつ工夫を重ねてきたにもかかわらず、少しずつではあるが5項目すべてで後半5年間のほうが効果が下がるという残念な結果となってしまった。教員、学生、教材、システム等の諸要因のうち、どの要因がもっとも強く影響したのかは不明であるが、教員が直接の原因であるとしても、その他の要因が悪化したのであるとしても、それを改善するのは教員の仕事であり、今後の向上に励みたい。しかしながら、授業の準備、評価、学生対応等の時間が他の仕事の影響、教員定数の削減等で少なくなっていることも原因の一部であるかもしれない。真の外国語教育の向上には教員定数の確保やTA(ティーチング・アシスタント)の増員等が望まれる。

4-5-2 クラスサイズ別

各クラスの履修者数を10名刻みで分類して、効果を比較したのが以下の表16である。

表16 クラスサイズ別の効果の比較

項目	10～19名	20～29名	30～39名	40～49名	50～59名	t
19	3.9	4.2	4.2	4.2	4.3	4.2**
20	3.4	3.8	3.9	3.9	3.9	5.8**
21	3.7	3.9	4.0	3.9	4.0	3.1**
23	3.7	4.0	4.1	4.1	4.1	4.5**
24	4.1	4.3	4.4	4.3	4.5	4.6**

tは「10～19名」と「50～59名」の差をt検定したもの

**p<.01

外国語授業は少人数で行うのが望ましいとされるが、一方では教員の人件費との兼ね合いでクラスサイズを大きくすることを求められることも多い。しかしながら、今回の調査では、わずかではあるが多人数クラスのほうが少人数クラスより効果が高いという結果になった。とくに少人数のほうが効果が下がるとは考えにくいことであるが、データのばらつき具合から見て多人数でも効果が下がらなかったと結論すべきかもしれない。

一方、50名以上のクラスで大人数にもかかわらず予想に反して高い効果が得られたのは、自律学習というCALLの特長をうまく活用した授業であるからと考えられる。しかし、CALL英語の授業は学生が自習するだけでなく、複数の教材やテスト、それにソフトウェア等を教員が同時進行で管理する必要があり、毎週学生の進捗度をチェックするために大量のテストをして、即座に手作業で採点をしてフィードバックを与えており、人数が多いと負担が大きいのも事実である。現状では教員の犠牲の上に成り立っているというべきであろう。今後は担当教員、TA等の増員（竹蓋2010）も必要である。

4-5-3 教材種類別

学習者が使用した教材は年度、学期、英語力等によって異なったが、主としてキャンパスライフを扱った4種類の教材を延べ5,112名が使用し、主として異文化・日常生活を扱った4種類の教材を延べ3,445名が使用し、ニュースおよび映画の素材を使用した2種類の教材を延べ472名が使用した。その結果得られた効果関連5項目の平均値の差は表17～表19に示した。

表17 二種の教材の効果の比較（「キャンパスライフ」
対「異文化・日常生活」）

項目	キャンパスライフ		異文化・日常生活	t
19	4.2		4.3	2.2*
20	3.9	>	3.8	3.4**
21	4.0		3.9	1.0n.s.
23	4.2	>	4.0	8.0**
24	4.4	>	4.3	6.6**

不等号は1%水準で有意の場合のみ *p < .05 **p < .01

表18 二種の教材の効果の差（「キャンパスライフ」
対「ニュース・映画」）

項目	キャンパスライフ		ニュース・映画	t
19	4.2		4.2	1.1n.s.
20	3.9		4.0	2.1*
21	4.0	<	4.2	4.0**
23	4.2		4.2	1.2n.s.
24	4.4	<	4.5	2.7**

不等号は1%水準で有意の場合のみ *p < .05 **p < .01

表19 二種の教材の効果の比較（「異文化・日常生活」
対「ニュース・映画」）

項目	異文化・日常生活		ニュース・映画	t
19	4.3		4.2	2.1*
20	3.8	<	4.0	3.6**
21	3.9	<	4.2	4.4**
23	4.0	<	4.2	4.7**
24	4.3	<	4.5	5.5**

不等号は1%水準で有意の場合のみ *p < .05 **p < .01

上記3つの表を見ると、いくつかの項目で有意差が見られ、表面的には「異文化・日常生活」<「キャンパスライフ」<「ニュース・映画」という緩やかな傾向があるようにも見られる。しかし、とくに「ニュース・映画」のカテゴリーはまだ2教材しか使用していないなど、各カテゴリーの教材数も少なく、現状ではこの見方を一般化すべきではないと考える。評価の結果に影響を与えたかもしれないもう一つの要因としては、学習者の英語力がある。本学では、これまで、「異文化・日常生活」の教材は比較的英語力が低い学生が使用し、「ニュース・映画」の教材は比較的英語力が高い学生が使用した。そして、外国語としての英語を国内で学ぶ大部分の学生にとってはメディアの英語が究極の実用レベルの英語としていちばん実感できる素材であったのかもしれない。このようなこともより強く評価結果に影響している可能性がある。

5. まとめ

千葉大学では2001年度から2011年度まで全学の学生を対象に、選択必修の形で「CALL英語」と呼ばれる授業を開講してきた。受講した学生は延べ11,682名にのぼる。本研究ではこの受講生群が質問紙法による調査に回答した結果をまとめて統計的に観察した。調査は等現間隔法を採用して39項目の質問に答えてもらったものである。本研究は今後の千葉大学における英語教育の教材開発、指導法改善に役立たせる目的でなされたものであるが、多くの有益な事実が明らかにされた。

最初に明らかにされた事実は、受講生が全体として「三ラウンド・システム」の理論に基づいたCALLシステムの構造を高く評価し、かつ、本学の「CALL英語」の効果が非常に高いと回答したことである。なかでも「この授業を取ってよかった」の項目には5点法で4.4という高い評価が見られた。これは、全受講生の85%が授業を肯定的に評価し、否定的に評価した者は4%に過ぎなかったことを示す。

次に明らかにされた事実は、受講生全員の意見としては、システムで提示される「参考情報（辞書）」や各学習作業ごとに提示される「正解」、それに「正解の英文」等が有効であったと考えていること、さらに、「文字の大きさ」や「文字の配色」が適切であったことであった。とくに、即時に提示されるKR情報の効果が高かったとしていることはCALLによる指導の特徴が好意的に受け入れられていることを示すと考えられる。

一方、特定の項目での質問に有効であった、適切であったと回答した受講生群の表明した教育効果に対する印象を高い順に見ていったところ、全員の意見とは異なり、「発展情報やTips & Hintsが有効であった」、「音声文字等の提示のタイミング（避負荷過大）がよかった」、それに「語彙テストの回数や内容を適切と思う」、などの項目があがってきた。興味深いことは、全員の回答が即時型や直接（教材直結）型KR情報の有効性を指摘しているのに対して、学習効果が上がったと回答している受講生群は遅延型や間接（応用）型の情報の有効性と語彙学習の有効性を指摘している点である。

次に、学習時間の多寡が学習の効果に大きくかかわっているということも学生自身の回答から明らかにされた。事実、外国語の学習には長期間におよぶ継続学習が求められるが、その実現には「学生自身に興味の持てる内容の素材」が重要であること、さらには「学習を通じて得られる成就感」や「教材の難易度と学習者の習熟度の適合性」が強い継続的学習意欲を生むことも学習者の意見から明らかにされた。

最後に、教師の間ではクラスサイズの大きさが英語教育の問題点として指摘されることが多いと田地野（竹蓋・水光2005）らが指摘しているが、本研究の結果でみる限り、三ラウンド・システムの理論に基づいたCALLでの英語指導の場合、必ずしも大人数クラスが少人数クラスよりも効果が低いとは言えないことが明らかにされた。

今回の研究の結果明らかにされたことの何点かは、すでに他研究者等によって明らかにされていることではある。しかしながら、それらのことが1万名を超える学習者自身の声で明らかにされたところに大きな意義があると我々は考える。

今回の受講生による評価から学べる反省材料も数点あった。その第一は、10年間の前半

5年間と後半5年間の指導効果を比較したところ、前半の方が後半より少しではあるが高かったという事実である。「初心忘るべからず」の警句をもう一度思い出して今後の指導に活かしたいと考える。しかしながら、一方で、指導の前半時期にはCALL担当教員が夜明けまでかかって教材の開発や改善、小テストの作成、採点等に尽力した日が続き、疲労が蓄積していった事実も見逃せない。教員の個人的な犠牲を強いる指導は長続きしないので、教員やTAの増員等の配慮も要求したいと考える。

二点目の問題点は個々の学生への最適教材の選定である。学習者の興味に合う内容のCALL教材は、高橋教授を中心とした本学教員や他大学教員の献身的な努力により充実されつつあるが、学習者の習熟度にあわせた「少し難しいレベルの教材の選定」に多少の問題があったことも今回の調査結果から見えてきた。この問題への対応策も今後考えていきたいと考える。

なお、今回の観察は、極めて大規模なものながら前期または後期の数か月間の学習に対する評価であった。今後は、年間または大学在学4年間等の長期にわたる継続的学習、それに生涯学習における問題点についても広く研究を発展させていくべきであろうと考える。そのことが真に実用コミュニケーション能力の養成を目指す英語教育に必要であると考えるからである。

謝辞

本研究の実施にあたり、土肥（2011）の研究と同様に、教材開発者、授業担当者の先生方、アンケート入力補助をされた職員の方々、授業とアンケートに真剣に取り組んだ受講生の皆さんをはじめ、全員のお名前を記すことはできないが、多数の方々のご協力をいただいた。「CALL英語」で使用した教材の一部は、

- ・文部科学省科学研究費補助金による特定領域研究「高等教育改革に資するマルチメディアの高度利用に関する研究」（領域代表者 坂元昂）の中の計画研究「外国語CALL教材の高度化の研究」（研究代表者 竹蓋幸生）
- ・メディア教育開発センターにおける「メディア教材開発事業」（制作担当：高津直己、山田恒夫 監修：竹蓋幸生）
- ・千葉大学教育改善推進費、「英語コミュニケーション能力養成のための初中級CALL教材New York Liveの開発」（研究代表者 高橋秀夫）
- ・文部科学省科学研究費補助金による基盤研究（B）、「英語レメディアルCALL教材の開発」（研究代表者 高橋秀夫）
- ・文部科学省科学研究費補助金による基盤研究（B）、「映画DVD対応CALL教材作成支援システムの開発」（研究代表者 土肥充）
- ・文部科学省科学研究費補助金による基盤研究（B）、「異文化理解を目指した英語聴解力養成用CALL教材の開発」（研究代表者 高橋秀夫）

を含め、複数の補助金をいただいて制作されたものである。また、本研究は

- ・文部科学省科学研究費補助金による基盤研究（C）、「CALLを活用する英語教育総

合システムの高度化に関する研究」(研究代表者 土肥充)
の補助金を受けて進行中の研究である。記して感謝の意を表したい。

主な参考文献

- 土肥充, 「千葉大学CALL英語履修者によるシステム評価結果の予備的分析」, 『言語文化論叢』, 第5号, 2011, pp.69-81.
- 小林雅之, 間淵泰尚, 大多和直樹, 「学部教育の問題点」, 『東京大学 現状と課題 (東大白書)』, 3, 東京大学総合教育センター, 2001, pp.164-173.
- Krashen, S.D. *Second Language Acquisition and Second Language Learning*, Oxford: Pergamon, 1981.
- 水光雅則, 「CALL教材 Listen to Me! を使った『授業』と『自習』と教師の役割」, 『京都大学総合人間学部マルチメディア教育運営委員会 MM News』, 5, 2002, pp.1-17.
- 竹蓋順子, 「CALLを活かす英語教育システム—教員が担う役割に関する考察」, 『e-Learning教育研究』, 第5巻, 2010, pp.10-19.
- 竹蓋幸生, 『英語教育の科学』, アルク, 東京, 1997.
- 竹蓋幸生, 水光雅則編著, 『これからの大学英語教育—CALLを活かした指導システムの構築』, 岩波書店, 東京, 2005.

補遺 アンケート項目（土肥2011より抜粋）

番号	設問内容
	聞き取り教材の内容について
1	内容、トピックに興味を持った
2	難易度は易しすぎた
3	難易度は難しすぎた
4	前半教材と比較し難しい（1年次のみ）
5	前半教材と比較し内容に興味を持った（1年次のみ）
6	写真、イラストは学習の助けや励みになる
7	指示は明確であった
8	ヒントは理解の役に立った
9	辞書は理解の役に立った
10	正解の提示は理解の役に立った
11	正解部分の英文表示は理解の役に立った
12	解説の記事（★）は理解の役に立った
13	Key Sentences は役に立った
14	Column は役に立った
15	Tips & Hintsは役に立った
16	Coffee Breakは役に立った
17	発展情報は役に立った
18	ユニットテストは役に立った
19	Step 1, 2, 3と進むにつれ聞けるようになった
20	聞き取りの力がついたと思う
21	CALLでの学習は楽しかった
22	聞き取り教材の1週間の平均学習時間は？
23	別の教材でも学習したい
24	この授業をとってよかった
	ソフトウェア（聞き取り教材）について
25	画面構成、デザインは適切であった
26	ボタンの配置、操作性は良かった
27	文字の大きさは適切だった
28	文字の配色は適切だった
29	写真の大きさは適切だった
30	写真の枚数は適切だった
31	写真の画質はよかった
32	動画の画質はよかった
33	音声、文字等提示のタイミングは適切だった
	語彙学習教材の内容について
34	音声を使用した語彙教材は記憶に残る
35	用例を使用した語彙学習は役に立つ
36	音声、文字等提示のタイミングは適切だった
37	語彙テストの回数は適切だった
38	語彙テストの内容は適切だった
39	語彙学習の1週間の平均学習時間は？

（設問22と39以外の回答は肯定から否定までを5～1とする5段階評価）