

C-testを教材として使えるか

An investigation into the use of the C-test as a teaching aid

神 本 忠 光 ・ 折 田 充

Abstract

The C-test, constructed by deleting the second half of every other word in a passage, is claimed to have a higher degree of ecological validity in that it is capable of presenting items in context in comparison to other vocabulary tests which are most likely to present items in isolation. However, it is not widely used mainly because its construct validity and face validity have been questioned by many studies (e.g., Grotjahn, 1986). Carroll (1987), while highly critical of the C-test as a measurement tool, suggested potential use of the test as a teaching aid. However, this idea has never been tested. This paper addresses the issue and compares vocabulary knowledge of same words between two types of presentation: in isolation and in context.

A C-test was developed by using one English passage with a result of 50 items. A group of 78 Japanese EFL students took a C-test and responded to a survey on the test. One week later, the same words were given to the same group, but in isolation where the participants were asked to translate test words into Japanese. Results indicated a mean score of 57.21% on the C-test and a mean score of 88.18% on the decontextualized vocabulary test. The difference between them was statistically significant (t -value

= 26.82, $df = 77$, $p < .001$), implying that some vocabulary knowledge, which appeared to be acquired when measured by a decontextualized test, may be judged to be yet unusable when measured by a C-test. The survey indicated that 95% of the participants judged that the C-test measures their vocabulary knowledge and about 85% felt that the C-test measures grammatical knowledge. Asked whether the C-test can work as a teaching aid, about 80% of the participants rated more than 3 on a five-point Likert scale. In summary the paper concludes that C-test can be used as a teaching aid.

1. はしがき

Meara (1980) が語彙研究の少なさを嘆いたのは一昔前のことで、ここ30年の間に盛んに語彙研究が行われ、数々の成果が収められている。語彙テストを例にとると、語彙サイズをはじめ、深さ、連想、コロケーション、アクセス速度などを測るテストが開発された。研究が進むにつれ細分化が進み、昔は測れなかった語彙の種々の側面を測定できるようになった。しかし、これらのテストを通底すべき本質的なものが欠けているような気がする。大所高所から今までの成果を見直し、評価し直す時期を迎えているのかもしれない。

Read and Chapelle (2001) は、今までに開発された主な語彙テストや測定法を三つの次元から整理し、語彙テストを評価する理論的枠組みを提案した。その分析の際、Read and Chapelleは現状は語彙力を個別に測る脱文脈テストが席巻しているとし、今後の語彙力測定は文脈の中で語彙力を測る方向を目指すべきだと主張している。脱文脈テストと言うのは、単語を文脈から切り離し単独に測定するやり方である。L2単語を羅列しL1へ訳させたり、L1訳語を提示しそれに相当するL2単語を書かせる形式がそれにあたる。この形式の語彙テストは、作成者にとってあまり時間も労力もかからず作成できる。また、手軽に実施できるし、採点も比較的容易である。学習者もこの形式の試験であれば、L1とL2語

のリストを見て独力で準備ができるし、コツコツと努力すれば学習した達成感を味わうことができる。この単語レベル中心の脱文脈語彙テストは、実施する教師にとっても受験する学習者にとっても、非常に都合の良いテストである。

ただ、語彙テストとしてこの脱文脈形式だけしか使われていないのであれば、語彙知識の本質を十分に測っていないことになる。語彙がどのように実際に使われているかを見てみると、単独で使われることはまず考えられない。話し言葉であれ書き言葉であれ常に他の語と一緒に文脈を形づくり、その中で個々の単語は本領を発揮する。従って、語彙力テストとして最も妥当な形式は有文脈での語彙テストと主張することができる (Carroll & Hall, 1985; Weir, 1990)。

ではどんな有文脈語彙テストがあるかと言うと、Read and Chapelle (2001)によると、多肢選択式クローズテストとC-testがそれに相当する。ここではC-testを取り上げる。C-testのテスト項目は、英文中で1語おきに単語の後半を機械的に削除して作る。簡単に誰でもできる。脱文脈語彙テストと異なり、幅広い語彙力を測定できる。また採点も簡単である。ドイツ語や英語をはじめ、フィンランド語、韓国語などいろいろな言語で開発され (Klein-Braley, 1985, Lee-Ellis, 2009, McBeath, 1989)、将来性があるテストのように見える。

その一方でC-testは議論のあるテストである。非常に肯定的な評価もある一方で、否定的な評価も少なくない。構成概念妥当性が明確ではないという主張 (Grotjahn, 1986; Lee-Ellis, 2009; Wright, 1990) がある。何を測っているのか明確でないというのは、測定道具としては壊滅的打撃である。他方、総合的な言語能力を測定するという主張 (Dörnyei & Katona, 1992; Klein-Braley & Raatz, 1984) もあれば、語彙と文法を中心に測っているという主張 (Eckes & Grotjahn, 2006) もある。また、ほぼ語彙力だけを測るという主張 (Singleton & Little, 1991) もあり、共通してこれを測っていると言えないテストである。さらに、C-testは表面的妥当性に欠けるという批難 (Babaii & Ansary, 2001; Jafarpur, 1995; Wright, 1990) もある。C-testはテストらしくないという主張である。受験者がまともに試験として取り組んでくれない危険性を孕んでいるとい

うのでは、信頼度が高いデータ収集を見込むことができない。以上のような理論的不安定さのためか、C-testは一部の支持者の間でしか使われていない。

一部の人々に使われるだけではもったいない特徴がC-testにはある。今までのどんな語彙テストにはない強みである。それは、語彙テスト項目の文脈化に長けているということである。短い文章で様々な語彙を抽出しその文脈の中で復元力を試すことは、脱文脈語彙テストにはとても真似できないことである。この特徴を何かに活かすことはできないものであろうか。

C-testに関する先行研究の中には、教材としての有用性に言及した発言がある。例えば、Carroll (1987) はC-test批評の際、厳しい評価を行いながらも、“informal teaching aid” (p. 105) として高い可能性を指摘している。Chapelle (1994) も “it (C-test) was the type of exercise that learners found normal in their classroom setting” (p. 180) と述べている。さらには、小山・清水・木村 (1997) は「C-testは授業中のexerciseとしては好意的に受け止められるかもしれない」(p. 81) と発言している。しかし、この可能性に関して、実証的に確かめられてはいない。

この研究では、脱文脈語彙テストと有文脈テストの違いを明らかにしたい。さらに、C-testの教材としての可能性を探ってみたい。研究課題をまとめると、以下のようなになる。

研究課題

1. 同じ語彙知識を脱文脈語彙テストと有文脈語彙テストで測ったらどうなるか
2. C-testは授業中の練習問題として使えるか

2. 実験

2.1. 実験参加者

日本人EFL大学生の非英語専攻（工学部と理学部）の3グループが実験に参

加した。その授業登録者数は合計96名であった。しかし、連続2週にわたるデータ収集過程で、病欠などのためデータが不完全となった学生が12名いた。その結果、分析対象の学習者数は78名である。

2. 2. C-test

2. 2. 1. 教材への転換

C-testは元来NRT目的で、約100テスト項目からなる。英文の内容に偏りがなないように、文章を4～6種類集め、各文章から20～25語のテスト項目を作る (Carroll, 1987)。しかしここでは、授業の指導教材の一部として使えるかどうかその可能性を探る目的なので、教室で使う教材の一部という仮定で単一の文章から作成する。また、この論文でのC-testテスト項目数は、授業中にあまり時間を取らずに実施しやすいように50項目とした。

2. 2. 2. 作成用文章

この研究の目的からすると、C-test作成用の文章には授業で実際に使った教材を使うのが望ましい。しかし、実験参加3グループ間に共通した教材はない。ここでは誰も読んだことがないと思われる英文、“Kiwi country” (1994, 『Mini-World』) を使った。可読性の指標 Flesch-Kincaid Grade level では9.5である。また英文の全語彙は基本語 (2語—“Kiwi” と “immigrant” —以外は word families の3レベルの範囲内) なので知らない単語はまずないと考えられる¹。

2. 2. 3. 作成方法

C-testの作成手順は、一語おきに単語の後半を削除するという原則 (Klein-Braley, 1985) に従った。ただ、解き方を説明する指示文に関して、通常のC-testの場合と異なる点が二点ある。一点目は、一般的指示文では単語がどのように削除されているかについては何も触れないのが通例である。しかし、この研究では「ほぼ半分が削除されている」と明記し、削除原理に気づくかどうかの要因が不当に得点に影響しないように配慮した。一般的指示文との二点目の違いは、一般的指示文では、ただ「解きなさい」だけで、どうやって解いたらよいか

等の説明はない。ここでは、教材としての可能性を探る目的があるので、C-testの例を解答とともに提示した。そしてテスト項目の文脈次第で、名詞を複数形にしたり又は動詞を変化させたりする必要があることにも注意を喚起した。(資料1を参照)

2.3. C-testに関するアンケート

学習者たちはC-testを受けている最中に、あるいは受け終わった後、どんな感想を持つだろうか。つまり、どんな能力が試されたと感じるだろうか。さらには、教材としての適切性に関してどう評価するだろうか。以上のような点に関する学習者の考えを調査するために、アンケート(資料2参照)を準備した。アンケートでの言葉遣いについて、学習者にわかりやすくする意図で専門的ニュアンスが強い「測定する」ではなく一般的な「関係している」などを使った。

2.4. 脱文脈語彙テスト

この脱文脈語彙テストは、単語単独での語彙力を測定することを目的とする。上記C-test作成文章からテスト項目50語すべてを抽出し、そのL1訳語を刺激語として与えL2語で答えるように求めた。C-testの50テスト項目中には、“and”と“of”がそれぞれ3回ずつ出現している。これらの単語を脱文脈語彙テストで重複して問うのは意味がないので、それぞれ1語として扱った。その結果、この脱文脈語彙テストの合計テスト項目数は46語となった。問題用紙では、C-testでの出現順序ではなく、ランダムに配列した。また、解答欄には、正答の第一番目のアルファベット文字を併記して提示した(資料3参照)。この処置で、C-testでのテスト項目と限りなく同じ単語の知識を測ることができる。

2.5. データ収集

データ収集は二週にまたがり実施した。第一週目には実験参加者3グループにC-testとアンケートを実施した。すべてのグループに対して、C-test受験の仕方

を文字および口頭の両方で説明し、開始した。開始15分後にほぼ解答し終えている様子だったが、全員が確実に解く努力をするように合計20分を与えた。C-testを回収直後、予告なしにアンケート（資料2）を配布し記入を依頼した。全員が書き終えたのを確認し回収した。その1週間後、同じ被験者グループを対象に脱文脈語彙テストを実施した。全員が解答できるように十分な時間を与えた。

2. 6. 採点

C-testと脱文脈語彙テストの採点は厳密に行った。特にC-testの採点に関しては、学習者の解答が英文の単語とまったく同じ場合のみ正解とし2点を与えた。例えば、動詞のつづりが解答の動詞と酷似する程度に復元されていても、屈折形まで一致していない場合は1点と採点した。正答に似ていない綴りや無記入は0点で採点した。

3. 結果

まず、C-testと脱文脈語彙テストの成績のまとめを報告する。両テスト間の得点差を見ることで、文脈の有無で同じ語彙の成績にどの程度の差が見られるかが明らかになる。次に、同じテスト項目に関して、両テスト間でどの程度難しさが違うかを明らかにする。最後に、C-testに関するアンケート結果のまとめを示す。

3. 1. 基礎統計量

表1 C-testと脱文脈語彙テストの基礎統計量 ($N = 78$)

	C-test	脱文脈語彙テスト
平均	57.21	88.18
標準偏差 (SD)	11.29	6.41
最小	29.00	65.22
最大	82.00	96.74

注：C-testと脱文脈語彙テストのテスト間の項目数は異なるので、比較しやすいように%で示している。

表1が示すように、有文脈テストであるC-testの平均点は57.21で、脱文脈語彙テストの平均点は88.18となった。その差は30.97で、統計的に有意であった (t 値 = 26.82, $df = 77$, $p < .001$)。このことから、文脈無しの単語テストでは得点に結びついている語彙知識の約1/3が、文脈内で測定されると得点に結びついていないことがわかる。C-testのSDは脱文脈テストのSDのほぼ2倍で、C-testの得点分布はかなり広い。脱文脈テストの得点分布は、最小65.22で最大96.74からわかるように、非常に幅が狭く全体的に極端に右寄りである。つまり、脱文脈テストでは大部分の参加者が非常に高い得点を取っている一方、C-testになると得点がかかなり大きく分散していることがわかる。まとめとして、同じ単語でも文脈の有無に拘わらず発揮される語彙力もあれば、単独ではその能力があるのに文脈中では発揮されない語彙力の存在が明らかになった。しかも文脈中で使える語彙力になっているかどうかは、学習者間でかなり差が見られる。

3. 2. 項目分析

文脈を伴うかどうかで、同じ語彙知識であっても、見え隠れすることが明らかになった。文脈内で試される方が単独で試されるより難しいことが明らかになったが、この差はテスト項目単位でどの程度異なるのであろうか。また、内容語と機能語という単語の区別によって、難易度 (Item Facility index: IF) の差に違いはみられるのだろうか。テスト項目毎に、脱文脈語彙テストでのIFからC-testでのIFを差し引けばわかる。ここではそのテスト項目単位の資料 (資料4) を基に、全体的な傾向を把握できるようにIFの差の頻度表を示す。

表2 難易度指数の差の頻度表

IF差	頻度	(%)	内容語	機能語
0.8 ~ 0.9	1	(2.17)	1	0
0.7 ~ 0.8	3	(6.52)	0	3
0.6 ~ 0.7	4	(8.70)	2	2
0.5 ~ 0.6	1	(2.17)	0	1
0.4 ~ 0.5	8	(17.40)	7	1
0.3 ~ 0.4	5	(10.90)	5	0
0.2 ~ 0.3	6	(13.00)	4	2
0.1 ~ 0.2	5	(10.90)	2	3
0.0 ~ 0.1	11	(23.90)	9	2
-0.1 ~ 0.0	1	(2.17)	1	0
-0.2 ~ -0.1	0	(0.00)	0	0
-0.3 ~ -0.2	1	(2.17)	1	0
合計	46	(100)	32	14

注：C-testには“and”と“of”がそれぞれ3回ずつ表れている。ここでは、同じ3語の難易度指数の平均を計算したうえで示した。

この表2から特徴をふたつ指摘することができる。一つ目は、IF差は幅広く分布しているものの、ほぼ8割近くのテスト項目がIF差0.0~0.5の範囲内を占めている²。その内訳をみると、IF差が0.0~0.1のテスト項目数が約24%あり、文脈の有無にかかわらずほぼ同等に発揮できる語彙知識があることがわかる。IF差0.1~0.5までの4区分で、それぞれ10.90%、13.00%、10.90%、17.40%となっているので、IF差がテスト項目単位でかなり幅があることがわかる。二つ目の特徴は、内容語は圧倒的にIF差0.0~0.5の範囲を占める一方、機能語はIF差0.0から0.8の範囲まで幅広く分布していることである。内容語のIF差が機能語のIF差ほどないのは、テスト項目となった内容語は比較的頻度が高く学習者が知らない単語はほぼなかったからであろう。一方、機能語のIF差の幅が大きい結果は、機能語は一般的に易しいという主張（Chapelle, 1994; Klein-Braley, 1985; Kokkota, 1988）とは異なる。機能語は一般的に短いし簡単に復元できそうに見えても、学習者の能力や文脈の関係で必ずしも易しい訳ではないことを示唆している。この件には議論でさらに触れる。

3.3. アンケート

学習者はC-testを受けながら、どんな英語の技能や能力が試されていると感じたのだろうか。学習者のそれらに関する判断を「まったく関係ない」を1とし「非常に関係ある」を5とした5件法で尋ねたが、その頻度を%で表3に示す。関係性の尺度の「4」以上を選んだ割合を中心に見ていく。

表を通覧して一番顕著なのは、「語彙力」である。「4」と「5」を選んだ割合の合計は約94%である。次に頻度が高いのは「文法」で、約85%を占めている。次に続くのは「読む」と「書く」で、それぞれ約77%と約63%を占めている。残り二つの「話す」と「聞く」を選んだ割合は両方とも10%以下で、関係性が弱い「1」や「2」の方に非常に集中している。まとめると、C-testは書きことばを中心に、語彙と文法に特に強く関係しているテストだと判断していることが明らかになった。これはEckes and Grotjahn (2006) や小山他 (1997) と同じ結論である。

表3 C-testと英語力の関係の頻度表 (%)

	1	2	3	4	5
文法	0.00	0.00	15.38	51.28	33.33
読む	0.00	3.85	19.23	37.18	39.74
書く	0.00	7.69	29.49	39.74	23.08
話す	17.95	52.56	20.51	7.69	1.28
聞く	39.74	42.31	12.82	3.85	1.28
語彙	0.00	0.00	6.41	42.31	51.28

次にC-testを授業の復習用に使うと仮定した場合、どの程度賛成するかと尋ねたアンケート項目への反応を見てみる(表4)。上記のC-testと英語力の関係の場合と同様に「4」と「5」の合計を基準として採用すると、約46%が積極的に賛意を表している。中間的な態度である「3」の約32%まで加えると、約8割に達する。教材の一部としてC-testを使うことに、賛同の意を得たと解釈してよからう。

表4 頻度表：C-testを授業復習用に採用 (%)

賛成度	1	2	3	4	5
%	1.28	20.51	32.05	35.90	10.26

3.4. 自由記述のまとめ

アンケートには、上記の質問以外に記述形式で答える質問項目が三つあった。

1. の「(アンケート用紙に)与えられた英語の技術や言語能力以外に(C-testが)関係している能力があるか」の質問に関して、ほとんどの学習者が無記入であった。アンケートに列挙されていた英語の能力以外に、C-testが関係しているものは取り立てて何もないという判断だと解釈できる。

3. 「C-test受験体験の感想」に関しては、「難しかった」と述べた学生が42名いた。この数値は、実験参加者の半分以上が共通して指摘した感想だということ注目になる。C-testをはじめて受け、見慣れない形式だったことも、この反応の中に含まれているだろう。また、「語彙力不足」に触れている学習者が13名いた。C-test作成に使った英文は基本的に易しい単語で書かれているので、知らないはずはない。これは、削除語の綴りが半分見えているにも拘わらず、自由に書けない自分の語彙知識への不満と解釈できる。

4. の「C-testの問題形式をさらに試験らしくする方法」として学習者の意見を尋ねたら、様々な回答を得た。頻度が多い順序に表5に示す。

表5 自由記述のまとめ

意見	頻度 (%)
1. 削除語数を減少	26 (30.95)
2. 選択肢で置換	12 (14.29)
3. 日本語訳を追加	9 (10.71)
4. このままで可	8 (9.52)
5. 削除文字数を明記	3 (3.57)
6. その他	13 (15.48)
7. 無記入	13 (15.48)

注：数人が複数の意見を書いていたので、頻度の合計がN=78より多い。

改善案として頻度が一番高かったのは、削除数を減らすという提案である。全体の約3割を占めている。具体的には、現行の1語おきの削除方法を、2語おきにする、あるいは1文おきにするというものがあつた。次に高い頻度を占めたのは、削除語の後半を記入させるのではなく、語全体を選択肢にするという意見で約15%あつた。日本語訳を追加するというのは、約1割を占めている。また、少数ではあるが、削除文字数を明記するという提案もあつた。どの提案もC-testの改善案を考える上で注目に値するものである。

特記すべき反応のなかに、「4. このままで可」という意見があつた。その頻度は8名で学習者全体のわずか10%程度だが、注目に値する。この学習者たちは、はじめて受けたC-testで削除語を上首尾に復元することができ、好意的な意見を述べているのだろうか。もしそうだとすれば、小山他(1997)が述べているように、C-testは好成績の学習者に気に入られる傾向があるのかも知れない。実験参加者全体78名の中でこの8名のC-test得点順位を降順に示すと、2, 16, 27, 39 (x2), 52.5 (x2), 76位であつた。実に幅広く分布している。従つて、このままのC-test形式で良いという意見とC-testの成績の間には、特に強い関係があるとは言えない。C-testの成績が良くない学習者の中にも、この形式で十分試験らしいと思つている学習者がいることがわかつた。

4. 議論

本研究の目的のひとつは、同じ語彙知識を文脈を伴うテストと伴わないテストで測つたらどうなるかを調べることであつた。その結果は文脈を伴うテストでの成績の方が著しく低くなつた。これはいったい何を意味しているのだろうか。

4.1. 文脈の有無で生じる語彙知識の差

本研究の学習者の平均的な語彙知識の習得状況を、表1のデータから図1のように表すことができる。C-testと脱文脈語彙テストで同じ語彙知識を測つたが、その11.82%はどちらのテストでも未習得であつた。残りの88.18%は脱文脈語彙

テストでは習得済だと示された。ところが、C-testでは88.18%の57.21%は習得済だが、30.98%（斜線部分）は未習得と二分された。

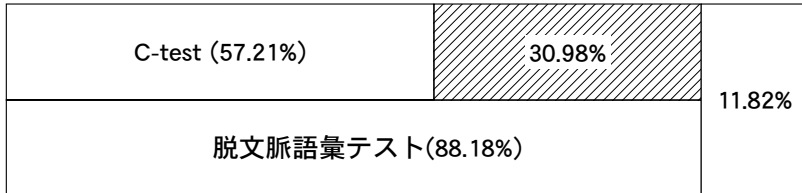


図1 2テストで見たテスト項目全体の習得・未習得状況

二つのテスト間で観察された約31%の差は何を意味しているのだろうか。脱文脈語彙テストで正解している語彙知識の中には、C-testで正解できる知識とできない知識の2種類があることを示している。その比率が平均的な学習者の場合、約57%と約31%ということになる。脱文脈語彙テストがもっぱら使われている現状から解釈すると、二つのテストで重複している約57%の部分は、有文脈で使える言語知識にも拘わらず、過小評価されている部分に相当する。さらには、脱文脈語彙テストでは、語彙知識があるように見えるが、有文脈テストでは使えない語彙力が約31%あることになる。

では、この語彙知識の差を生み出した二つの語彙テストの相違、文脈とは何かと改めて問う必要がある。脱文脈語彙テストには文脈はないので、C-testでの文脈を見てみよう。C-test内の文脈とは、テスト項目の(1)未削除部分と(2)前後の文脈が提供する意味・形態・統語的情報と言える。この情報を使える能力があるかどうかを、C-testは明らかにする機能を持つ。脱文脈語彙テスト中心の語彙知識測定では、この部分を測定できていないことを意味する。

C-testが要求する語彙知識を、脱文脈語彙テストの場合と比較してみる。表6に名詞と動詞の主なテスト項目をIF指数と共に示した。脱文脈語彙テストでは見出し語が正答となるところを、C-testでは複数形で書くことや、屈折形で答えることを求めている。表からわかるように、すべてのテスト項目で、C-test

でのIF指数の方が脱文脈語彙テストでのIF指数より低い。これは、脱文脈語彙テストでは正解できても、C-testでは正解できなかった率が高いことを意味する。例えば、“mountain”を見てみよう。脱文脈語彙テストでのIF指数は1.0なので、実験参加者全員が正解していることがわかる。しかし、C-testで求められている解答は複数形の“mountains”である。そのIF指数0.85というのは、ほとんどの学習者がこの文脈では複数形が必要だと気づいているが、それに気づかず“mountain”と解答した学習者が数人いることを示している。

表6 文脈の有無で求める語彙知識の違いの例³

	C-test		脱文脈語彙テスト		IF差
	key	IF	key	IF	
名詞	forests	0.49	forest	0.97	0.48
	mountains	0.85	mountain	1.00	0.15
	parts	0.39	part	0.78	0.39
動詞	beginning	0.69	begin	0.90	0.21
	swimming	0.96	swim	0.98	0.02
	worrying	0.67	worry	0.94	0.27

このC-testで明らかになった脱文脈語彙テストとのギャップを、学習者へフィードバックすることは大切である。両テスト間のIF差の大きさは、指導の留意点の目安となりうる。この種の指導を行うことで、学習者は文脈の中で要求される語彙知識に敏感になり、ひいては書いたり話したりする際に正確で適切な英語表現をできるようになる。

4.2. 教材としてのC-test

本研究での二つ目の目的は、C-testを教材として使う可能性について調べることであった。アンケートの結果、ほぼ8割の学習者が是認したことは先に示した。先行研究のC-testに関する類似した調査と比べると、この数値は非常に高い。例えば、Jafarpur (1995) は、C-testの妥当性や用途についてNNS学習者に質問をしているが、8項目すべてに対して否定的な評価である。また、小山他 (1997)

もEFL日本人大学生に対して類した調査を行っているが、本研究ほど高い評価には至っていない。

なぜ先行研究と本研究では大きな評価の差が生じたのであろうか。いくつかの要因が考えられる。例えば、C-testに使った英文や学習者の能力の違いなどが関係しているであろう。しかし、種々の要因の中で一番注目すべきはC-testの与え方の違いである。先行研究はどの研究でもC-testの一般的な実施要領に従い、受験者に対して削除原理の説明などは行っていないと考えられる。1語おきに単語の後半部分が削除されていることに気づくかどうかは、受験者次第である。これに対してこの研究では、削除原理を明らかにしC-testの例題も与え、さらには名詞の複数形や動詞の屈折形にも注意が必要なことを喚起した。復習教材としてC-testを使う時に学習者が心得ておくべき情報を、実験前に明示的に与えた。先行研究とのこの違いが、本研究の学習者たちにC-testに対してどう対処してよいかの情報となり、解いている最中に役に立ち、高い評価に至っていると思われる。

4.3. テスト項目作成

実際に教材の一部としてC-testを使う場合、テスト項目の削除原理を再検討する必要がある。学習者へのアンケートが示しているように、過半数の学習者が今の形式のC-testを難しいと感じている。なんらかの改善が必要なことを強く示唆している。

テスト項目作成に関して、考慮すべき視点を三点挙げる。一点目は、教科書の英文からテスト項目を抽出するとき、無作為か意図的かということである。無作為抽出では、テスト項目の選択に指導者の意図は介在しない。易しくするという目的なので、現行の1語おきという削除原則を2語おきにすることが考えられる。この抽出法の利点は、機械的にこの原則を当てはめていけばよいのでだれでも簡単に作成できる。その結果、いろいろな特徴を持つ語彙を選ぶことになり、学習者はいろいろな語彙知識を試されることになる。しかしその反面、無作為抽出法は指導上の焦点が明確でなくなる傾向になり、効率が悪いと主張可能であ

る。その点、意図的抽出法であれば、指導者の焦点を明確にでき、教育上の力点を操作することができる。しかし、削除語間の間隔が一律でなくなり、文脈で与える利点を失う危険性がある。また、類似した特徴を持つ削除語彙が集中して現れることになり、学習者に抽出語彙の意図を見抜かれ、本来の指導目的を完遂できない可能性がある。無作為抽出・意図的抽出のどちらにも一長一短があるので、C-testの使用目的を熟考し、また学習者の反応を見ながら柔軟に対処する必要があるだろう。

二点目の考慮点は、内容語と機能語に関係している。先行研究の中には、機能語の復元は易しいと報告しているものもある (Chapelle, 1994; Klein-Braley, 1985; Kokkota, 1988)。機能語は一般的に短く数は少なく、推測しやすくして易しい印象を与える。その結果、テスト項目として出題する価値はあまりないと解釈されかねない。しかし、本研究では必ずしも易しくないという結果を得た。C-testでの内容語と機能語のIF指数平均はそれぞれ0.59と0.53で機能語の方が難しく、統計的に有意 (t 値=2.78, df =154, p < .01) であった。テスト項目次第ではあるが、例えば、今回のC-testの中の機能語 “by” と “their” のIF指数は共に0.21でかなり難しいと言える。さらに面白いことに、同じ機能語であっても、文脈次第で難易度が変化する。今回のC-testにたまたま “and” と “of” が3回ずつ出てきた。その難易度を表7にテスト項目番号と共に示す。“and” のIF指数の範囲は0.36~0.68で、“of” のIF指数の範囲は0.54~0.79と、文脈次第で難易度がかなりの幅で変化している。以上のことから、機能語もテスト項目として出題する意義があることを指摘しておきたい。

表7 重複機能語の難易度

“and”		“of”	
Item No	IF	Item No	IF
9	0.68	20	0.79
38	0.40	43	0.54
44	0.36	47	0.67

三つ目の考慮点は、テスト項目の削除文字数に関してである。C-testの削減原理は、単語の後半部を削除するが、あまり厳密に適用すると、復元するのが極端に易くなる場合がある。例えば、今回のC-testでは、テスト項目番号10 “swim_____” や17 “moun_____” がそれらの単語に相当する。前者はswimの動詞全体が見えているので、屈折させればよいということが一目瞭然である。後者の方は、直前にある単語 “beaches” から “mountain(s)” を導き出すことはさほど難しいことではないだろう。つまり、単語の後半を機械的に削除するということは、見直して良いかもしれない。ではどのような削除方法をすればよいか。Kokkota (1988) によれば、正解となりうる語の数、項目難易度や項目の長さなどを挙げている。このことを参考にすると、先ほどの “swim_____” と “moun_____” は、削除原理よりさらに一文字ずつ削除し、“swi_____” と “mou_____” としても良いかもしれない。

C-testを教材として適したものに改善するために、テスト項目作成に関して考慮点を3点述べた。どの考慮点を採用するか判断は、実際の学習者を相手にして試していくうちに徐々に解決していこう。ただしどのような改善を行うにしろ、常に心しておくべきことがある。それは、文脈でテストする価値を失わないように最大限の努力をすべきだということである。そしてC-test採用の目的は、能力の測定ではなく、あくまで語彙指導上の一つの教材であることを銘記すべきである。

4. 4. 本研究の限界と今後の研究

この研究の問題点をふたつ指摘しておきたい。一つ目の問題は、C-testを教材の復習用としてその可能性を探るのが目的であるにも拘わらず、今回使用したC-testは実際の教材に基づいていない点である。できるだけ実験への参加者を多く確保するために、どのグループも見ることがない英文材料を使った。C-testを受けたのも初めてだと思うが、この初見の英文だったことも関係し、C-testを難

しいと感じた感想が多かったのかもしれない。実験参加者数は少なくても、もし実際に授業で使ったことがある英文を本研究のC-test材料にしていたら、内容や語彙など全般的に親しみを感じ、もっと教材としての妥当性を正確に反映した研究になったかもしれない。この点に関しては、今後の研究課題としたい。

二つ目の問題は、C-testのテスト項目を独立しているものとして扱っている点である。あるテスト項目がわかれば、その完成した知識を使って付近の別なテスト項目を解く手助けとなり得る。つまり、C-testは局所独立の原理を維持できないテストである。一般的なC-testのように複数の英文をもとに複数のC-testがあれば、各英文の合計点を一つの独立した得点として処理することができる。しかし、語彙指導を中心とした授業でのC-testは短く一つの英文に基づく可能性が高いので、そのような処理は不可能である。語彙指導の一環での教材ということを考慮すれば、局所独立の原理を維持しているかどうかの懸念は不要かも知れない。ただこの研究ではこのやむを得ない問題点が、採点や分析に含まれていることを指摘しておきたい。しかし、このC-testを教材として採用することは、相対的に資する方が大きいと確信している。

5. 結論

この研究では、有文脈語彙テストであるC-testはテストとしては様々な論議があるが、その特徴を活用すべく、教材として受け入れられるかどうかその可能性を探った。その結果、個別の単語テストでは測れない語彙知識に関して、学習者の意識を高める教材として十分使える可能性が明らかになった。しかし、このままのC-testの形式では難しい傾向があるので、削除原則を少し改変する必要があるかも知れない。その際には、文脈を与えすぎないように注意しなければならない。C-testの強みである文脈で削除語を復元するという学習者のタスクはテストであれ、教材であれ変わらない。言語能力を測定するテストとしてではなく、教材として使うことで、C-testの良さを潰すことなく、学習者も文脈での語彙の特徴に注意を注ぐことができる。教材としての新たな役割を得ることで、C-testは

自由にさまざまな応用が可能である。C-testは語学教師の指導資料の一つとして大きく貢献できる可能性を持っている。脱文脈語彙テストと棲み分けをすることで、学習者の語彙能力をいっそう伸ばすことが出来よう。

注

1. 分析には、Heatley, Nation, & Coxhead. (2002)のRANGE and FREQUENCY programsを使用した。
2. この頻度表にはIF差がマイナスのテスト項目が二つある。Europe (IF diff: -0.01)とZealander (IF diff: -0.26)である。C-test上の“Europe”は前半部分が見えているので、間違える可能性は後半部にしかない。一方、脱文脈語彙テストでは日本語の「ヨーロッパ」に対して語頭のEに続く単語大半が要求されている。この要求度の違いでIF差がマイナスになっていると考えられる。また、“Zealander”は“New Zealander”の一部であるが、C-testの冒頭に削除されない形で使われている。注意深い学習者はこれに気づき、C-testで正解者が多かったことを示している。
3. 「名詞の複数の形態素」に関して、表に掲載していないテスト項目がある。New Zealanderである。C-testの正解(New) ZealandersのIF指数は0.74で、脱文脈語彙テストでのNew ZealanderのIF指数は0.49である。本文で述べている傾向と矛盾しているように見えるが、これはこのテスト項目の無削除語がC-testの冒頭に出てきているので、学習者はそれを見逃さずに参考にした結果、脱文脈語彙テストのIFより高くなっている、と解釈できる。

参考文献

- 小山由紀江, 清水裕子, & 木村真治. (1997). 質問紙法による英語リーディングテストの表面的妥当性と波及効果の研究. *Language Laboratory*, 34, 75-92.
- Babaii, E., & Ansary, H. (2001). The C-test: a valid operationalization of reduced redundancy principle? *System*, 29, 209-219.
- Carroll, J. B. (1987). Book reviews: Klein-Braley, C. and Raatz, E. 1985: C-Tests in der Praxis. In *Fremdsprachen und Hochschule*, AKS-Rundbrief 13/14, Bochum:

- Arbeitskreis Sprachenzentrum [AKS]. *Language Testing*, 4, 99-106.
- Carroll, B. J., & Hall, P. J. (1985). *Make your own language tests: A practical guide to writing language performance tests*. Oxford: Pergamon.
- Chapelle, C. A. (1994). Are C-tests valid measures for L2 vocabulary research? *Second Language Research*, 10, 157-187.
- Dörnyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing*, 9, 187-206.
- Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23, 290-325.
- Grotjahn, R. (1986). On the methodological basis of introspective methods. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research* (pp. 54-81). Clevedon: Multilingual Matters.
- Heatley, A., Nation, P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. Retrieved February 7, 2006, from http://www.vuw.ac.nz/lals/staff/Paul_Nation
- Jafarpur, A. (1995). Is C-testing superior to cloze? *Language Testing*, 12, 194-216.
- Klein-Braley, C. (1985). A cloze-up on the C-Test: A study in the construct validation of authentic tests. *Language Testing*, 2, 76-104.
- Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test. *Language Testing*, 1, 134-146.
- Kokkota, V. (1988). Letter-deletion procedure: A flexible way of reducing text redundancy. *Language Testing*, 5, 115-119.
- Lee-Ellis, S. (2009). The development and validation of a Korean C-test using Rasch analysis. *Language Testing*, 26, 245-274.
- McBeath, N. (1989). C-Tests in English: Pushed beyond the original concept? *RELC Journal*, 20(2), 36-41.
- Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts*, 13, 221-246.

- Mini-World. (1994). Kiwi country. *Mini-World*, 36.
- Read, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing*, 18, 1-32.
- Singleton, D., & Little, D. (1991). The second language lexicon: Some evidence from university-level learners of French and German. *Second Language Research*, 7, 61-81.
- Weir, C. J. (1990). *Communicative language testing*. New York: Prentice Hall.
- Wright, F. (1990). What makes a test communicative? *Teanga: Journal of the Irish Association for Applied Linguistics*, 10, 53-62.

資料 1

次の英文は、途中から英単語が一語おきに後半のほぼ半分が削除されています。削除されている部分を書きなさい。以下に例を示します。削除されている部分を補ってみましょう。

例：“If you can’t fix the environment, please stop breaking it!” Severn
(1)sa__ to (2)t__ world (3)lea__ at (4)t__ Earth (5)Sum__ in Rio
(6)i__ 1992.

答え：(1)said, (2)the, (3)leaders, (4)the, (5)Summit, (6)in

やりかたはわかりましたか。動詞の時制や名詞の複数形などにも注意をする必要があります。

では、以下の問題をやってください。最初の2文は、この文章が何について書いてあるかヒントを与えるために、削除されていません。

問題：

Borrowing the name of their national bird, the kiwi, New Zealanders are proud to be “Kiwis.” They live in a country the size of Japan minus Hokkaido, with a population of only 3.5 million people. It’s (1)har__ ever (2)cro__ and (3)th__ is (4)lit__ pollution. (5)Y__ can (6)dr__ without (7)worr__ about (8)tra__ jams (9)a__ go (10)swim__, surfing, (11)o__ diving (12)a__ the (13)be__ — the (14)wa__ is (15)alw__ clean.

(16)Beau__ beaches, (17)moun__, and (18)for__ are (19)on__ part (20)o__ the (21)rea__ why New Zealand (22)i__ a wonderful (23)pl__. Its (24)peo__ are a (25)mix__ from (26)ma__ lands, (27)begi__ with (28)t__ Maori (29)s__-travelers (30)w__ arrived (31)i__ New Zealand one (32)thou__ years (33)a__. They (34)we__ joined (35)la__ century (36)b__ immigrants (37)fr__ Britain (38)a__ other (39)pa__ of (40)Eur__.

The New (41)Zeala__ have a (42)str__ spirit (43)o__ independence

(44) a _____ are (45) pr _____ of (46) th _____ way (47) o _____ life. (48) Mo _____ Kiwis
(49) wo _____ probably (50) ag _____ that the quality of a person's daily life
is more important than his or her job. Perhaps Japanese could learn
something from this.

学生番号 :

名前 :

得点 :

資料 2

アンケート

学生番号：

名前：

先ほど解いた問題について、次のアンケートに答えてください。

あの問題は、どんな英語力に関係していると思いますか。下の表の英語力について、「まったく関係ない」から「非常に深く関係ある」の5段階で該当する番号を丸で囲んでください。なお参考までに、先ほどの問題の例題と答えを下に再掲しておきます。また、表の下の1. ～4. の質問にも答えてください。

例：“If you can’t fix the environment, please stop breaking it!” Severn
(1)sa ___ to (2)t___ world (3)lea___ at (4)t___ Earth (5)Sum___ in
Rio (6)j___ 1992.

答え：(1)said, (2)the, (3)leaders, (4)the, (5)Summit, (6)in

	まったく関係ない	ほとんど関係ない	少し関係ある	かなり関係ある	非常に深く関係ある
文法力	1	2	3	4	5
読む能力	1	2	3	4	5
書く能力	1	2	3	4	5
話す能力	1	2	3	4	5
聞く能力	1	2	3	4	5
語彙力	1	2	3	4	5

1. 上の表にある能力以外に関係している能力がありますか。あると思う場合は、それを書いてください。

2. あの問題形式を授業の復習に使いたいと思います。どの程度賛成しますか。

まったく賛成しない

1	2	3	4	5
---	---	---	---	---

非常に賛成

3. あの問題を解いてみて、どんな感想を持ちましたか。自由に書いてください。

4. あの問題形式に手を加えていっそう試験らしくするには、どうしたらよいと思いますか。

ご協力ありがとうございました。

資料3

英語実力確認テスト

次の日本語に相当する英単語を書きなさい。ただし、与えられているアルファベットで始まる単語で答えなさい。全部で46問あります。

- | | | | |
|----------------------------|---------|-------------------------|---------|
| 1. 意見が合う | a _____ | 24. たいていの、最大多数の | m _____ |
| 2. 山 | m _____ | 25. 力が強い、(意志などが) 強い | s _____ |
| 3. (場所・運動などの起点) …から | f _____ | 26. (前述の名詞をさして) その | t _____ |
| 4. (今から) …前、以前に | a _____ | 27. 1,000、千個、千人 | t _____ |
| 5. そこに、そこで | t _____ | 28. (全体のうちの) 部分、一部分 | p _____ |
| 6. 混合、入り混じったもの | m _____ | 29. 泳ぐ、水泳する | s _____ |
| 7. (車などを) 運転する | d _____ | 30. …か…、…または…、あるいは… | o _____ |
| 8. 海 | s _____ | 31. ヨーロッパ | E _____ |
| 9. …の、…の性質を持つ | o _____ | 32. を始める | b _____ |
| 10. 込み合った、人口過密の | c _____ | 33. ほとんど…ない[しない] | h _____ |
| 11. 多くの、多数の | m _____ | 34. (位置) …の中に、…において | i _____ |
| 12. 小さい、ほとんどない | l _____ | 35. heやsheの後に来るbe動詞の現在形 | i _____ |
| 13. そして、および | a _____ | 36. (willの過去形) だろう | w _____ |
| 14. ニューゼaland人 | N _____ | 37. (人・車の) 往来、交通 | t _____ |
| 15. 場所、所 | p _____ | 38. あなた(たち)は [が] | y _____ |
| 16. youの後に来るbe動詞の過去形 | w _____ | 39. 人々 | p _____ |
| 17. (theyの所有格) 彼(女)らの、それらの | t _____ | 40. いつも | a _____ |
| 18. 砂浜、浜辺 | b _____ | 41. 誇りを持っている、…に満足している | p _____ |

- | | | | |
|--------------------------|---------|------------------------|---------|
| 19. 美しい、すばらしい | b _____ | 42. 〈媒介・手段〉によって、
…で | b _____ |
| 20. 森林、山林 | f _____ | 43. 水、水中 | w _____ |
| 21. 唯一の、たった1つの | o _____ | 44. 理由、わけ | r _____ |
| 22. すぐ [この] 前の、昨…、
先… | l _____ | 45. 心配する、気にする | w _____ |
| 23. …する (人) を示す関係
代名詞 | w _____ | 46. (場所) …に、…で | a _____ |

学生番号：

名前：

得点：

資料 4

脱文脈語彙テストでの項目難易度指数：IF降順 (Item NosはC-testのもの)

Item No	単語	IF	Item No	単語	IF	Item No	単語	IF
29	sea	0.83	39	part	0.39	8	traffic	0.08
36	by	0.78	50	agree	0.36	33	ago	0.08
49	would	0.74	13	beach	0.35	2	crowded	0.06
30	who	0.71	21	reason	0.35	23	place	0.06
25	mixture	0.69	6	drive	0.32	24	people	0.05
19	only	0.67	45	proud	0.29	5	you	0.05
11	or	0.63	37	from	0.28	32	thousand	0.04
46	their	0.62	7	worry	0.27	16	beautiful	0.03
12	at	0.53	34	were	0.26	10	swim	0.03
14	water	0.49	27	begin	0.22	4	little	0.02
35	last	0.49	15	always	0.21	40	Europe	-0.01
18	forest	0.49	22	is	0.17	41	Zealander	-0.26
3	there	0.49	48	most	0.15			
9	and	0.68	17	mountain	0.15			
38	and	0.40	31	in	0.14			
44	and	0.36	20	of	0.79			
avg	and	0.48	43	of	0.54			
42	strong	0.44	47	of	0.67			
26	many	0.43	avg	of	0.66			
1	hardly	0.40	28	the	0.10			