リーディングテスト項目の 2 次元分類における困難点 Difficulties in Two-dimensional Classification of Reading Test Items: What Made Them Disagree?

静 哲人 (関西大学)

Shizuka Tetsuhito (Kansai University)

This report is a follow-up to a study by Shizuka (1999a; 1999b), in which a new scheme for classifying reading test items was proposed and tried out. In this scheme, an item was to be placed in one of nine categories based on its position on two dimensions, one concerning the size of the relevant text portion and the other relating to the depth of cognitive processing required to respond correctly to the item. In the trial phase, four testing experts were asked to apply the taxonomy to Cambridge KET, PET and FCE. While acceptable agreement rate was observed for the two lower-level tests (KET & PET), rather low inter-classifier reliability was revealed for FCE. The present paper focuses on detailed qualitative and introspective examination of the FCE items over which classifiers' perception differed to a considerable extent in Shizuka's (1999a; 1999b) trial. The analyses revealed several different ways language testers view specific items, highlighting limitations of the adopted scheme.

1. 研究の目的

リーディング能力の重要性はインターネット時代の到来とともにますます高まっていると言える。しかしその能力を測定すべきリーディングテストにおいて、特定のアイテムがどのような（下位）技能を測定しているのかを決めることができるか、できるとすればどのような方法でか、という疑問に対する決定的な答えはまだない。Shizuka (1999a; 1999b)は、リーディングテスト項目を分類するための新しい枠組みとして「2次元分類法」を提案し、言語テスト研究の専門家による分類試行を実施した。その結果、本分類法が、ある程度の有効性ととも、分類上の困難点も抱えていることが明らかになった。本論は、この Shizuka (1999a; 1999b)の分類試行で、4人の分類者の認識が特に一致しなかった項目を具体的にひとつひとつ検証し、言語テスト専門家が個々のアイテムを見る見方の異同を探るとともに、2次元分類法の枠組みを改善するうえでの示唆を探ることを目的とする。

2. 研究の背景

2.1 リーディングテストアイテムは何を測定するのか

リーディング能力が、さらに細かい下位技能に分解できるかどうかについては未だに定説がないにも関わらず、市販の教材や問題集などには、ほとんど a priori に、「A という能力を測る問題」「B という能力を測る問題」に類する表現があふれている(Williams & Moran 1989)。また、いわゆる「低位技能」「高位技能」という表現は、厳密な定義がな

されないまま広く使用されている。

そのような分類は本当に実体をともなったものなのか、という疑問を第2言語能力の測定に関連して初めて投げかけたのは Alderson & Lukmani (1989)であろう。彼らの研究では、ランカスター大学の英語教師が 41 のリーディングテストアイテムを、その測定していると思われる技能によって 'higher' 'middle' 'lower' の3つに分類したところ、27もの項目において、ほとんど合意が得られなかったという。この結果から Alderson & Lukmani は、アイテムを「低位スキルを測るもの」とか「高位スキルを測るもの」などと分類することの妥当性に強い疑問を提示した。

この結論は Weir, Hughes, & Porter (1990)から分類者のトレーニングの欠如と 'higher order' 'lower order' という用語の定義の不十分さを指摘される。Lumley (1993)は、より注意深く定義された下位技能リストと、よく訓練された分類者を使うことでこの問題に決着をつけようと試みた。彼の研究では、注意深く定義された分類範疇を用い、綿密な打ち合わせを行った後に分類に当たった5人の英語教師は、メルボルン大学で使用されていた L2 リーディングテストアイテムの分類について、ほぼ完全な合意を得ることができたと報告されている。

彼の分類法のひとつの特長は、当該のアイテムに正解するために「どれほど深い読みが求められるか」（同じ語彙を探せばいいのか、パラフレーズしてあるものを理解するのか、さらには明示的には述べられていないものを自分で統合・推測するのか）という視点と、「答えがどれくらいの範囲に見いだされるか」（単文、複文の中にあるか、あるいはパラグラフ全体にあるのか）という視点を組み合わせている点である。しかし、Lumley のリストは、彼の研究で分析対象とした特定のテストの 22 アイテムを分類するという目的に特化したものであるため、十分な汎用性がないように思われる。

2.2 2次元分類法 (Shizuka 1999a; 1999b)

Lumley (1993)のリストの中に萌芽がみられる発想をさらに進めて、できる限り客観的で汎用性の高い分類を目的として Shizuka(1999a; 1999b)が提案したのが、2次元分類法である。この枠組みでは、すべてのアイテムを2つの次元上の位置で表すことを試みる。第1の次元は「問題を解く際に直接関わる言語単位の大きさ」である。まず便宜的に、すべての問題項目を次の3つに分類する。

1. SENTENCE: 問題文中の1つの文、あるいはそれ以下の単位（語、など）を理解すれば正解できるもの。
2. LOCAL DISCOURSE (LOCAL と略): 隣接する2つの文を理解すれば正解できるもの。
3. GLOBAL DISCOURSE (GLOBAL と略): 隣接していても3つ以上の文、あるいは、2つの文でも隣接していないものを理解しないと正解できないもの。

ここで問題にするのは、「問題を解く際に直接関わる」部分のみであり、いわゆる長文問題であっても、当該アイテムに正解するために理解しなければならないのが1文であれば、SENTENCE に分類する。文を範疇境界の単位に用いたのは、伝統的に「低位」と「高位」の境界に用いられることが多かったこと(Alderson 1978)、また文の境界を越えると verbatim memory が急激に劣化する(Jarvella 1971; Daneman 1991)など、リーディングにおいて文が特別な意味を持つ単位であることが理由である。

第2の次元は「関わる認知行動の深さ」である。正解を導くために必要な認知行動により、すべての問題項目を次の3つに分類する。

1. EXPLICIT VERBATIM (VERBATIM と略): 明確に述べられている意味のみを理解すればよく、しかも、問題文中の該当部分と問いの選択肢がまったく同じ語や表現を用いているもの。
2. EXPLICIT PARAPHRASE (PARAPHRASE と略): 明確に述べられている意味のみを理解すればよいが、問題文中の該当部分と問いの選択肢が別の語や表現を用いているもの。
3. IMPLICIT: 明確に述べられている意味を理解することに加え、いわゆる「行間」を補って判断する、あるいは推測する必要があるもの。

以上の2つの次元を組み合わせ、 $3 \times 3 = 9$ つの範疇を作る。詳細は Shizuka (1999b) を参照されたい。

2.3. 分類トライアル (Shizuka 1999a; 1999b)

本分類法がどの程度実際に機能するか検証するため、Shizuka (1999a; 1999b)は分類トライアルを行った。University of Cambridge Local Examination Syndicate から市販されている3つのレベルの past paper、Key English Test (KET)、Preliminary English Test (PET)、および First Certificate in English (FCE)のリーディングセクション部分（各35アイテム）を、日本人の言語テスト専門家3人に筆者を加えた4人が分類した。協力者の3人は、いずれも言語テストの分野で M.A.以上を持ち、かつ日本人学習者相手の教授経験も豊富な教師でもあった。まず、趣旨説明と2次元分類法の詳細を記述した資料を各分類者に郵送し、疑問点があれば文書で補足説明を行った。その上で、個別に、3テスト計105アイテムを、SENTENCE-LOCAL-GLOBAL (SLG 次元)と VERBATIM-PARAPHRASE-IMPLICIT (VPI 次元)の2つの次元ごとに分類した。物理的制約により、対面でのセッションは設けなかった。

SLG 次元については、SENTENCE を"1"、LOCAL を"2"、GLOBAL を"3"、VPI 次元については、VERBATIM を"1" PARAPHRASE を"2" IMPLICIT を"3"と換算し、分類者間信頼性を計算した結果が表1である。全体に低い数値であるが、特に FCE が低い。なぜこのようなことになるのかを、具体的なアイテムごとに検討することは、言語テスト専門家のアイテムの認識の異同を明らかにし、そのうえで分類法の枠組みの改善

につながる示唆が得られる可能性がある。

表 1：3 テストの次元別の分類者間信頼性係数 (Shizuka 1999b)

	SLG	VPI
KET (35 items)	.74	.44
PET (35 items)	.73	.57
FCE (35 items)	.57	.06
合計 (105 items)	.69	.42

3. 本研究：分類一致困難項目の分析

Shizuka (1999a; 1999b)の分類データを質的に分析し、分類者の判断がどのように異なったのかをアイテム毎に検討するのが本論の主目的である。スペースの制約により、最も分類一致が困難であった FCE (Appendix に掲載) に対象を絞ることとする。一致が困難であった度合いの指標として、各アイテムの各次元ごとに、分類結果を数値化(1, 2, 3)した4つの数値の標準偏差を計算した。例えば、あるアイテムにおいて、SLG 次元に関する4人の分類者の判断が、全員 SENTENCE であったとすれば、数値データは、"1, 1, 1, 1"となり、この4つの数値の標準偏差は"0"である。一方、このアイテムにおいて、VPI 次元の分類者の判断が、2人が VERBATIM で2人が IMPLICIT だったとすると、数値データは"1, 1, 3, 3"となり、標準偏差は、"1.0"となる。

3.1 FCE パート1

FCE パート1はパラグラフと要約文のマッチング問題である。分類結果と標準偏差を表2にまとめた。一番左の列がアイテム番号を示す。R1-4 は、4人の分類者を指し、R1 は筆者である。SLG次元とVPI次元を比べて気づくことはSLG次元では標準偏差が0、つまり全員が一致したアイテムもあるが、逆に判断が"1"(SENTENCE)と"3"(GLOBAL)の両極端に別れている場合もあるということである。4人の分類者の判断に、1と3の両方が含まれると、標準偏差は0.7を越える値となる。これを判断一致困難度のひとつの基準として、0.7を越えた値にはアスタリスク(*)をつけ、以下ではそのようなアイテムを個別に検討することにする。

SLG次元で一致困難だったアイテム4は、"These problems can become severe at large transfer airports, known as 'hubs', because of the large number of bags that are processed." ではじまるパラグラフに対して "H. Busy airports are likely to have more problems."を選択する問題である。このパラグラフの第1文のみでHを選択することは可能だと判断された

ので、筆者は"1"(SENTENCE)と分類し、他の二人の分類者も同様に判断したが、R2 は"3"(GLOBAL)と判断している。パラグラフの第1文に続く支持文も読む必要がある、という判断かと思われる。

表2 パート1のアイテム分類結果

item#	SLG 次元					VPI 次元				
	R1	R2	R3	R4	SD	R1	R2	R3	R4	SD
1	2	2	1	1	0.5	2	3	2	3	0.5
2	3	3	3	2	0.43	2	3	3	3	0.43
3	3	3	3	3	0	2	3	3	3	0.43
4	1	3	1	1	0.87*	2	3	2	3	0.5
5	1	3	3	1	1*	2	3	3	3	0.43
6	3	3	3	3	0	2	3	3	3	0.43
7	1	1	1	1	0	2	2	2	2	0

またアイテム5は、二人ずつ"3"と"1"に分かれた。"Even efficient transfer airports, such as Amsterdam, Copenhagen, Singapore and Zurich have their bad days."で始まるパラグラフの要約文として、"G. No airport can guarantee the safe handling of luggage."を選ぶ問題である。筆者を含む二人は、この第1文だけでもGを選ぶことが可能だと判断して"1"(SENTENCE)を選んだが、このアイテムの第1文と要約文Gの関係を、アイテム4の第1文と要約文Hの関係を比較してみると、アイテム5の関係のほうが、やや間接的かと思われる。"3"(GLOBAL)と判断した二人は、本パラグラフを構成する3文のすべてを読む必要がある、と感じたのだと思われる。

3.2 FCE パート2

表3 パート2のアイテム分類結果

item#	SLG 次元					VPI 次元				
	R1	R2	R3	R4	SD	R1	R2	R3	R4	SD
8	3	1	1	1	0.87*	2	2	2	3	0.43
9	3	3	1	3	0.87*	3	2	2	3	0.5
10	3	2	3	3	0.43	1	2	2	3	0.71*
11	2	3	2	1	0.71*	2	2	2	3	0.43
12	3	2	3	1	0.83*	2	2	2	3	0.43
13	3	2	2	3	0.5	2	2	2	3	0.43
14	3	2	3	1	0.83*	2	2	3	2	0.43

パート2は、問題文章に関する質問の答えとして最も適当なものを4つの選択肢の中から選ぶ、という多肢選択問題である。パート1で見られた、SLG次元の標準偏差がVPI次元よりも大きいという傾向はこのパート2でもやはり見られる。判断が両極に分かれたケースがVPI次元では1アイテムなのに対し、SLG次元では5アイテムもある。

アイテム8は、"Who is intended to benefit from the advice given in the article?"に対する答え"C. those who are interested in establishing a career."を選ばせる。問題文中の直接の該当箇所は、"we will be dealing with those of you taking a real step on the career ladder,..."である。筆者はその前にある、"We are not concerned..."の文、および"Nor are we concerned ..."の文も必要と判断したが、他の3人は、"we will be ..."の1文のみで正解を選択可能と判断したと思われる。振り返って考えれば彼らの判断のほうが適切かも知れない。

アイテム9は、"According to the writer, why is the choice of your second job important?"に対する答"A. It will affect your future job prospects."を選ばせるものであるが、筆者が見る限り、直接の該当箇所はないようである。第2、第3パラグラフ全体から間接的に読みとる必要があると判断し、"3"(GLOBAL)とした。他の二人も同様の判断をしたようだが、ひとりは"1" (SENTENCE)としている。第2パラグラフの第2文、"Much depends on your long-term aim."で判断できると感じたのかも知れない。

VPI次元でアスタリスクのついたアイテム10は、代名詞"it"の指す、"B. second job"を選ばせるものである。"it"に言及する最も距離の近い名詞句は、3行上の"This job"である。さらに1行上に"it"があり、さらに1行上に"your second job"とある。筆者はこの"second job"と正解選択肢の表現がまったく同一なので"1" (VERBATIM)としたのだが、分類者2と3は、本文中の"This job"と、正解選択肢の"second job"の関係として、"2" (PARAPHRASE)を選んだのかも知れない。分類者4の"3" (IMPLICIT)という判断は解釈するのが難しい。

3.3 FCE パート3

パート3は文章中の空所と文のマッチング問題である。分類結果の特徴(表4)は、SLG次元においてもVPI次元においても、ひとりの分類者の判断が他の3人と大きく違った結果、全体に標準偏差が大きくなっているということである。SLG次元では、分類者4の"1"(SENTENCE)という判断が突出しており、一方VPI次元においては分類者1の"1"(VERBATIM)という判断が突出している。分類者1(筆者)はいずれの選択

肢も、はっきりのべられている意味のみを理解すればよく、またパラフレーズも含まれないので"1"(VERBATIM)と判断した。しかし、このような「空所」と「選択肢」をマッチさせる形式においては、分類原則で言う、「問題文中の該当部分」が空所であるため、何と何を比べて「まったく同じ語や表現を用いているか」を判断すればよいかわからない、とも言える。分類者3は、空所に入る「意味」と選択肢の関係が「パラフレーズ」であると判断したのかも知れないし、分類者2と4は、空所の前後の文と、選択肢の文との間の「行間」を推測する必要がある、と判断したのかも知れない。分類基準と問題形式の不整合がこのような判断のばらつきにつながってしまったと言えよう。

表4 パート3のアイテム分類結果

item#	SLG 次元					VPI 次元				
	R1	R2	R3	R4	SD	R1	R2	R3	R4	SD
15	3	3	3	1	0.87*	1	3	2	3	0.83*
16	3	3	3	3	0	1	3	2	3	0.83*
17	2	3	3	1	0.83*	1	3	2	3	0.83*
18	2	3	3	1	0.83*	1	3	2	3	0.83*
19	2	3	3	2	0.5	1	3	2	3	0.83*
20	2	3	3	1	0.83*	1	3	2	3	0.83*
21	2	3	3	1	0.83*	1	3	2	3	0.83*

3.4 FCE パート4

パート4は、文章と、それに含まれる内容を表す文の、マッチング問題である。表5をみると、SLG次元において14アイテム中、実に9アイテムにアスタリスクがついており、しかもその8アイテムの分類に用いられているのは"1"と"3"のみで、中間の"2"(LOCAL)がないことに驚く。

最も判断の判断の割れているアイテム25, 26, 32a, 32bを詳しく見ることにする。アイテム25は、"I have to make myself exercise. "という内容を、"B. Linda Kelsy"の言葉から読みとる問題である。該当箇所を1カ所だけ選ぶとすれば、"I know exercise gives me energy, but it's finding the energy to do it!"であろう。分類者1と分類者2はこの1文のみで判断できると考えたのだろう。しかしこの文だけでは、決定的ではないのも確かではある。分類者3と4は、この文の前後の文を合わせた、"When I... ..at home. "の3文が必要と判断したのであろう。この判断も十分理由のあるものといえよう。

アイテム26も状況が似ている。"I enjoy behaving like a child. "に対して、"E. Deborah

Moggach"の、最後の文 "I like it because it's childlike. "だけでも選べるとも言えるが、完全に確実とも言い切れない。前の2文を加えた3文を読む必要があるとも言える。

表5 パート4のアイテム分類結果

item#	SLG 次元					VPI 次元				
	R1	R2	R3	R4	SD	R1	R2	R3	R4	SD
22a	1	1	1	1	0	2	3	2	2	0.43
22b	3	3	3	1	0.87*	2	2	2	3	0.43
23	1	1	1	1	0	2	2	2	2	0
24	1	3	1	1	0.87*	2	2	2	2	0
25	1	3	3	1	1*	2	2	3	3	0.5
26	1	1	3	3	1*	2	2	2	2	0
27	1	1	1	1	0	2	2	2	2	0
28	1	1	3	1	0.87*	2	2	2	2	0
29	1	1	1	1	0	2	3	2	2	0.43
30	1	3	1	1	0.87*	2	3	2	3	0.5
31	1	1	1	1	0	2	2	3	3	0.5
32a	1	3	1	3	1*	2	3	2	3	0.5
32b	1	3	1	3	1*	2	3	3	3	0.43
33	1	3	3	1	1*	2	3	2	2	0.43

アイテム 32a は、"Being cheerful gives you energy."を"C. Annie Nightingale"のことばかり、32b は、"D. Floella Benjamin"のことばかり読みとる問題である。C のなかで最も近い文は、"I don't eat healthily to get my energy - I think it's natural enthusiasm and being positive."であり、分類者 1 と 3 はこの文のみで可能と判断し、"1"(SENTENCE)としたのだろう。分類者 2 と 4 はこれでは不十分と考え、あとに続く 3 文程度も必要と判断し、"3"(GLOBAL)としたのだろう。これもやはりどちらとも決めがたいところであった。

4. 考察

4.1 ふたつの次元の間のパタンの差

Shizuka (1999a; 1999b)は、分類者間の判断の関係を、積率相関係数を元にして比較した結果、SLG 次元のほうが、VPI 次元よりも判断が一致しやすかったと結論づけた。しかしながら今回の標準偏差による分析では、むしろ VPI 次元のほうが、ばらつきが小さいという結果がでた。この一見相矛盾する現象は実は矛盾ではない。積率相関係数は、二つの変数の値が同じ方向で変動（増減）する傾向の強さの指標である。一方、今回の分析での標準偏差は、同じ傾向で増減するか否かを表すものではない。表すのは、複数の変数の値の分散の度合いである。たとえば、お互いにまったく無関係に増減するが、

値の差は小さいような複数の変数間の積率相関係数は限りなく 0 に近いだろうが、標準偏差も比較的小さくなる。逆に、お互いにまったく同じ方向で増減するが、値の差は大きいような複数の変数であれば、積率相関係数は限りなく 1 に近いだろうが、値の分散を表す標準偏差は比較的大きくなる。この点を確認した上で、今回の分析結果を解釈するならば、SLG 次元は、分類者間の判断のばらつきの幅は大きい、全体としては同じ方向で変動し、一方 VPI 次元は、分類者間の判断が異なる幅は小さかったが、その小さな幅での変動の方向は、分類者間で異なる方向であることも多かった、ということになる。

4.2 SLG 次元

SLG 次元では分類の標準偏差が 0 つまり完全に一致した場合も 9 アイテムあり、VPI 次元の 6 アイテムよりも多い反面、同一アイテムの分類として"1"も"3"も含まれる、つまり判断が両極に別れた場合も、なんと 20 アイテムに上った。上述したようにそのような場合には標準偏差にアスタリスクを付して上で個別に再検討を加えたわけであるが、その結果、この次元の分類に内在する困難点が明らかになったと言える。確認すると、当該アイテムに正解することが、単一の文の理解のみで可能なら SENTENCE, 3 つ以上の文もしくは隣接していない 2 つの文の理解が必要なら GLOBAL というのが定義であったわけだが、ある文「のみ」で正解できるか、その周辺のいくつかの文も「必要」か、という判断には、かなりのグレーゾーンが存在する、ということが改めて判明した、というわけである。複数の分類者の判断が"1"と"3"に分かれた場合、改めて考察してみても、一方の判断の方が圧倒的に妥当と思われるケースは皆無と言ってよく、どの場合も、正解への鍵となる中心的 1 文のみで正解を選べるとも言えるが、より確実には、それ以外の文も読む必要がある、といったケースがほとんどであった。

つまり SLG 次元上での位置を決定するための、「正解に到達するために必要最低限なテキスト部分」の確定という作業には、予想以上に複雑な要因が絡むということが判明したと言える。まず、「誰がたどり着くのか」という問題がある。例えば英語の初学者と英語教師では、同一のアイテムにおいて、どれが正解かを見極めるのに必要なテキスト情報の量は明らかに異なるであろう。これに影響する変数として、まず英語そのものに対する知識の量がある。また、英語のリーディングテストというものに対する知識の量、たとえば選択肢の設定のされ方等についての知識の量、がある。そして、当該英文の内容に関わる背景知識の量がある。これらの変数に関して極端に有利な受験者、たとえば母語話者の言語テスト専門家等、であれば、アイテムのリードと選択肢を見ただけ

で、かなりの確率で正解を見抜いてしまうことも十分あるだろう。一方、これらの変数において不利な受験者にとっては、直接の該当個所のみならずそれを取りまく箇所までも理解してはじめて正解にたどりつくことができることもあるかもしれない。そして、忘れてならないのは、今回の分類を行ったのはこれらの変数に関して非常に有利な言語テスト研究者であったという点である。4人の専門家の間でさえ判断が大きく割れた場合が35アイテム中20アイテムもあった、という事実は深刻に受け止める必要がある。

さらに「どのレベルの自信をもってたどり着くのか」という問題がある。FCE や多くの標準テストのように、解答を限られた数の選択肢の中から選ぶ形式のアイテムの場合にはとくにこの問題が大きい。結果として正解にたどり着くにしても、100%の確信をもって選択する場合と、そうでない場合では、必要となる情報量が異なるだろう。「おそらくこれだろう」という程度での選択が可能となる「必要最低限のテキスト部分」のほうが、「絶対に間違いなくこれである」という選択を可能とするテキスト部分の量より小さいことがあるだろう。また正解の選択肢に直接たどりつくのではなく、それ以外の錯乱肢を除外できるための情報を得られるためのテキスト部分、となるとまた微妙に異なってくることも考えられる。

4.3 VPI 次元

一方の VPI 次元上での分類において明らかになった大きな問題は、判断に、問題形式が大きく影響していたという点である。パート3は文をディスコースの空所に補充するマッチング形式だったわけだが、このパートの7つのアイテムすべてを、分類者1は VERBATIM、分類者3は PARAPHRASE、分類者2と4は IMPLICIT に分類したのである。この3つの分類基準は、「問題文の該当個所」と「正解の選択肢」のふたつの言語形式の間の関係がどの程度直接的か、によって3つに分類することを意図したものであった。すなわち両者がキーワードまで共有してほとんど同一の表現を用いている場合が VERBATIM、異なる表現を用いながらも同じ内容を表してれば PARAPHRASE、表している内容までもやや異なっており、2者を結びつけるためには推測を必要とするなら IMPLICIT としたのである。この基準にしたがって、パート1においては、問題文中の該当個所と要約文の関係、パート2においては問題文中の該当個所と選択肢の関係、パート4においてもやはり、問題文中の該当個所と選択肢の関係、を分類したはずである。この3つのパートにおいては（アイテム10を唯一の例外として）分類者の判断が"1"と"3"の両極に分かれた場合はない。これは問題形式上、何と何の関係を分類するのかという認識が一致比較的明確だったためではないかと考えられる。

一方、唯一パート3は、本来は問題文ディスコース中にあった文を抜き出して選択肢としているため、何と何の関係を分類するのか、という点が分類者によってまちまちになってしまった可能性がある。片方が選択肢であることは明白だが、その一方が、「空所」であるはずはないにしても、「当該の空所に入るということを決定づけるための該当箇所」なのか、「当該の空所の前後の文」なのかが分類基準を読んだだけでは明らかでない。このどちらであるにしても、この空所補充の形式とこれ以外のパートの形式と決定的に異なる点がある。それは他のパートでは比較するふたつの言語形式が、基本的には同じ意味を表していたのに対し、この空所補充の形式では、同じ意味を表すものではなかったという点である。つまり、例えば、パート1では要約文と本文の該当箇所が同じ意味を表しているものを探し出してマッチングする、という作業を行う。これに対しパート3では、正解の選択肢と、当該の空所の前後の文は、同じ意味を表す関係ではなく、前後に配置してうまく接続する、という関係にある。このような関係の場合、AとBという文がうまく接続するかどうかという判断は、AおよびBの「明示的」意味のみを材料にして下すことが可能であり、そこに「行間」を補う行為や、「推測」は必要ない、というのが、筆者の「行間」および「推測」の語感である。この理由によって分類者1はこのパートのアイテムを IMPLICIT を分類しなかった。しかしながら、このパートのアイテムすべてを "3" IMPLICIT と判断した分類者2と4は、AとBの文がうまく接続するかという、AとBというふたつの文の間の関係に関わるものだから、「行間」に関わるのだ、と考えたのかも知れない。こうしてみると、「行間を読む」などの日常の慣用句や、「推測」などの多様な内容を指す可能性のある用語を分類基準に用いる場合は、細心の注意を払う必要があった、ということになる。

4.4 まとめ

2次元分類法の出発点は、分類の基準を可能な限り単純化して、ある程度客観的な分類が可能になる、汎用性のある分類法を作り出すことにあった。しかしながら今回のデータから判断する限り、その目標に照らして満足のいく結果が得られたという結論は出し難い。「正解を導くために必要な最低限のテキスト部分」「明示的意味のみで処理できる場合と、推測が必要な場合」等の、かなり明確であると期待された概念でも、言語テスト研究者の間でさえ、かなりの異なる判断を生み出した。これは郵送または E-mail という間接手段によったとはいえ、例を含むかなり詳細な解説資料を全員によんでもらい、質疑応答を行った上での結果である。楽観的に解釈するならば、この種類の試みでよい結果を出すには、やはり対面でのトレーニングセッションを持ち、例えば各自が「行

間」などという用語で意味するもののすりあわせを行うのが不可欠である、と結論づけることができよう。しかし悲観的な解釈としては、Alderson (1990b) が到達した、「テスト項目が測るものは受験者によって異なるようだ」という結論と、残念ながら整合する結果となった、とも言えそうである。

謝辞

本研究の一部は、平成 11 年度関西大学学部共同研究費によって行った。問題項目の分類作業に貴重な時間を割いて協力してくださり、また詳しいコメントをくださった安間一雄、松澤伸二、及川賢の各氏に深く感謝申し上げる。また根岸雅史氏には分類方法について貴重なコメントをいただいた。あわせて御礼申し上げます。

引用文献

- Alderson, J. Charles. "A Study of the Cloze Procedure with Native and Non-native Speakers of English." Diss. University of Edinburgh, 1978.
- Alderson, J. Charles. "Testing Reading Comprehension Skills (Part Two): Getting Students to Talk about Taking a Reading Test (A Pilot Study)". *Reading in a Foreign Language* 7. 1 (1990b): 465-503.
- Alderson, J. Charles, and Yasmeen Lukmani,. "Cognition and Reading: Cognitive Levels as Embodied in Test Questions." *Reading in a Foreign Language* 5. 2 (1989): 253-270.
- Daneman, Meredyth. 1991. "Individual Differences in Reading Skills." In Barr, R., Kamil, M. L., Mosenthal, P., & Pearson, P. D. (Eds.). *Handbook of Reading Research* (Vol. 2). New York, NY: Longman, 512-538.
- Jarvella, Robert. "Syntactic Processing of Connected Speech." *Journal of Verbal Learning and Verbal Behavior* 10 (1971): 409-416.
- Lumley, Tom. "The Notion of Subskills in Reading Comprehension Tests: An EAP Example." *Language Testing* 10. 3 (1993): 211-234.
- Shizuka, Tetsuhito. "Two-dimensional Classification of Reading Rest Items." Paper presented at the 23rd Convention of Kanto-koshin-etsu English Language Education Society, held in Kofu on August 10. 1999a.
- Shizuka, Tetsuhito. リーディングテスト項目の2次元分類法. *The JLTA Journal* 2 (1999b) [In press]
- Weir, Cyril, Arthur Hughes, and Don Porter. "Reading Skills: Hierarchies, Implicational Relationships and Identifiability." *Reading in a Foreign Language* 7. 1 (1990): 505-510.
- Williams, Eddie, and Chris Moran. "Reading in a Foreign Language at Intermediate and Advanced Levels with Particular Reference to English." *Language Teaching* 22. 4 (1989): 217-228.

Appendix: FCE Reading Section used for the classification trial

Part 1

(問題番号 1～7)

右のページに載せてあるのは、飛行機の旅についての雑誌記事です。パラグラフ 1～7 の内容を最もうまく要約している 1 文を、下の A～I からひとつずつ選びなさい。なお、例として、パラグラフ 0 を最もうまく要約しているのは選択肢 I です。

- A. It is important to report missing luggage.
- B. Airlines make special efforts to avoid errors.
- C. Airlines sometimes blame travellers for lost luggage.
- D. No airline is free from mistakes.
- E. Plan ahead when travelling.
- F. Unexpected events can result in luggage going missing.
- G. No airport can guarantee the safe handling of luggage.
- H. Busy airports are likely to have more problems.
- I. Losing your luggage can be very inconvenient.

In case you lose your luggage

Alex McWhirter looks at airlines that lose his bags.

0. There is nothing more disappointing than arriving at an airport overseas to discover that your baggage has been left behind. At best you will have to put up with wearing the clothes you stand up in for hours or days, until the airline reunites you with your luggage. At worst, you may be in a different climate zone, thousands of miles from home and forced to wear wholly unsuitable clothes.
1. Although airlines rarely reveal how many cases they lose, it is a fact of life that sooner or later regular travellers will be parted from their luggage. Even the best airlines slip up from time to time, and it is impossible for any carrier to guarantee that a passenger's checked luggage will go on the same flight, particularly when a journey calls for one or more changes of aircraft.
2. The system works like this. Airlines insist on exaggerated check-in times (which require passengers to report to the airport at a given time before departure) designed to allow sufficient time for baggage to pass through the airport and be loaded on to the plane. Minimum connecting times (MCTs) are the shortest time it takes to transfer between two flights. These, too, are exaggerated to allow for baggage transfers.
3. In normal circumstances the system works well. But extra security checks at airports and problems with air traffic combine to cause delayed flights. All this can cause the baggage system to fail. Then there is the possibility of human error, or an accident in which the destination label is torn off.
4. These problems can become severe at large transfer airports, known as 'hubs', because of the large number of bags that are processed. Last year, for example, London's Heathrow airport handled more than 41 million passengers, of whom nine million were changing planes. British Airways alone handled two million transfer passengers at Heathrow, with most making the one-mile transfer between Terminal 1 (for Domestic and European flights) and Terminal 4 (for long-distance flights).
5. Even efficient transfer airports, such as Amsterdam, Copenhagen, Singapore and Zurich have their bad days. The risk of baggage being lost when changing planes is higher than average at certain airports. Even the United States has problems - Miami airport is well known for luggage going missing when transatlantic passengers make immediate connections for destinations for Latin America.
6. You should choose direct flights whenever possible and check in well before the official time. If a change of plane is unavoidable, or makes your flight less expensive, then try to fly the same airline throughout. Try to allow more connecting time by taking an earlier flight to the transfer airport, and make sure you label your luggage inside and out with your home and holiday addresses. Don't forget to include the flight numbers.
7. If, after all this, your luggage still goes missing, you must contact the appropriate airline official in the baggage hall and complete a *property irregularity report* (PIR). This must be done before leaving the airport.

Part 2

(問題番号 8~14)

次の文章は新聞記事からの抜粋です。問いに対するもっとも適当な答えを選びなさい。

A lot of advice is available for college leavers heading for their first job. In this article we consider the move to a second job. We are not concerned with those looking for a second temporary position while hunting for a permanent job. Nor are we concerned with those leaving an unsatisfactory job within the first few weeks. Instead, we will be dealing with those of you taking a real step on the career ladder, choosing a job to fit in with your ambitions now that you have learnt your way around, acquired some skills and have some idea of where you want to go.

What sort of job should you look for? Much depends on your long-term aim. You need to ask yourself whether you want to specialise in a particular field, work your way up to higher levels of responsibility or out of your current employment into a broader field.

Whatever you decide, you should choose your second job very carefully. You should be aiming to stay in it for two to three years.

This job will be studied very carefully when you send your letter of application for your next job. It should show evidence of serious career planning. Most important, it should extend you, develop you and give you increasing responsibility. Incidentally, if the travel bug is biting, now is the time to pack up and go. You can do temporary work for

a while when you return, pick up where you left off and get the second job then. Future potential employers will be relieved to see that you have got it out of your system, and are not likely to go off again.

Juliette Davidson spent her first year after leaving St. Aldate's College working for three solicitors. It was the perfect first job in that 'OK ... they were very supportive people. I was gently introduced to the work, learnt my way round an office and improved my word processing skills. However, there was no scope for advancement. One day I gave in my notice, bought an air ticket and travelled for a year.'

Juliette now works as a Personal Assistant to Brenda Cleverdon, the Chief Executive of Business in the Community. 'In two and a half years I have become more able and my job has really grown,' she says. 'Right from the beginning my boss was very keen to develop me. My job title is the same as it was when I started but the duties have changed. From mainly typing and telephone work, I have progressed to doing most of the correspondence and budgets. I also have to deal with a variety of queries, coming from chairmen of large companies to people wanting to know how to start their own business. Brenda involves me in all her work but also gives me specific projects to do and events to organise.'

- 8 Who is intended to benefit from the advice given in the article?
- A students who have just finished their studies
 - B. people who are unhappy with their current job
 - C. those who are interested in establishing a career
 - D. people who change jobs regularly

- 9 According to the writer, why is the choice of your second job important?
- A It will affect your future job prospects.
 - B It will last longer than your first job.
 - C It will be difficult to change if you don't like it.
 - D It should give you the opportunity to study.
- 10 underlined 'it' in paragraph 4 refers to your
- A first job.
 - B second job.
 - C application.
 - D career.
11. If you have a desire to travel, when does the writer suggest that you do it?
- A straight after you have left college
 - B when you are unable to find a permanent job
 - C after you have done some temporary work
 - D between the first and second job
- 12 What is meant by 'you have got it out of your system' in paragraph 4?
- A You have planned your career sensibly.
 - B You are an experienced traveller.
 - C You have satisfied your wish to travel.
 - D You have learned to look after yourself.
- 13 How did Juliette Davidson benefit from the experience of her first job?
- A It was a good introduction to working in an office.
 - B She met a variety of interesting people.
 - C It enabled her to earn enough money to travel.
 - D She learnt how to use a word processor.
- 14 In what way is Juliette's current job better than her first job?
- A She has a more impressive job title.
 - B She now knows how to start her own business.
 - C She has been able to extend her skills.
 - D She is more involved in the community.

Part 3

(問題番号 15～21)

右ページに載せてあるのは、チョコレートについての雑誌記事です。文中の空所 15～21 に当てはまる、最も適切な文を A-I からひとつずつ選びなさい。なお例として、最初の空欄 0 に当てはまるのは選択肢 I です。

- A. A Swiss company then introduced milk solids to the process which gave us milk chocolate.
- B. They also used them to make a drink called *xocoatl*.
- C. Until the last century, the chocolate drink was made from solid blocks of chocolate which had to be melted down in hot water.
- D. When dried they become cacao beans, which can be used to make chocolate.
- E. Clever advertising which associated it with the healthy qualities of milk from the English countryside quickly established the bar as a rival to the more decadent French brands.
- F. British manufacturers include up to 5 per cent vegetable fat in their chocolate, something forbidden elsewhere.
- G. As most cacao farmers operate on a very small scale, many were forced out of business.
- H. This has forced manufacturers to look for new ways to attract customers.
- I. Only at Christmas do people eat more of the cocoa-based foodstuff.

Bitter water hits the big time

Chocolate, which has its origins in South America, is now part of a multi-million pound worldwide business.

At Easter, British people spend over £230 million on chocolate. A massive eight per cent of all chocolate is bought at this time.

0 Although the large-scale industrial production of chocolate began in the last century, the cacao plant was first cultivated by the Aztec, Toltec and Mayan civilisations of Central America over three thousand years ago.

The cacao tree is an evergreen, tropical plant which is found in Africa, South and Central America, the West Indies and South-East Asia. The fruit of this tree is melon-sized and contains 20-40 seeds.

15 In English-speaking countries, they are often called cocoa beans. This is a misspelling from the 17th century when they were also called cacao and cacao beans.

The Aztecs used cocoa beans as money.

16 This is from the word in the Aztec language, Nahuatl, meaning 'bitter water'. In Aztec times the chocolate drink was flavoured with spices and used on ceremonial occasions and for welcoming visitors. The Spanish found the drink more palatable mixed with cinnamon and sugar, but the recipe did not spread to the rest of Europe for another century. In the late 17th century, chocolate houses were set up in Europe's capital cities, where people gathered to drink chocolate.

17 But in 1826, C J van Houten of the Netherlands invented chocolate powder. This was made

by extracting most of the cocoa butter from the crushed beans.

The age of the chocolate bar as we know it began in 1847 when a Bristol company, Fry and Sons, combined cocoa butter with pure chocolate liquor and sugar to produce a solid block that you could eat.

18

At the turn of the century, the British chocolate market was dominated by French companies. In 1879 the English company Cadbury even named their Birmingham factory Bournville (*ville* is the French word for town) in the hope that a little French glamour would rub off. But then came Cadbury's famous Dairy Milk bar which began life as Dairymaid in 1905.

19

It seems that, for the time being at least, chocolate intake in Britain has stabilised at about four bars each week.

20 The latest marketing trick is the so-called 'extended line'. This is when the humble chocolate bar becomes an ice cream, a soft drink or a dessert, to tempt chocoholics who have grown tired of conventional snacks.

At the other end of the production process, cacao farmers are still feeling the effects of a crash in cocoa bean prices at the end of the 1980s.

21 Perhaps you could spare a thought for them as you munch your next chocolate bars.

Part 4

(問題番号 22～33)

様々な女性がインタビューされています。次の 22～35 の文の内容を言っている女性を、ひとりずつ (22 と 32 だけは、ふたりずつ) 選んで記号で答えなさい。同じ女性を 2 度以上選んでも構いません。なお、例として、最初の 0 の文の内容を言っているのは女性 F です。

- 0. I don't remain interested in things for long.
- 22. People I work with give me energy. (ふたりいる)
- 23. Being forced to do things uses up your energy.
- 24. Doing things for the first time gives me energy.
- 25. I have to make myself exercise.
- 26. I enjoy behaving like a child.
- 27. Certain kinds of people annoy me.
- 28. It is important to me that others are interested in my work.
- 29. You should use your energy carefully.
- 30. Sometimes I lack energy at work.
- 31. I like creating something from nothing.
- 32. Being cheerful gives you energy. (ふたりいる)
- 33. You should take every opportunity to have fun.

WHERE I GET MY ENERGY

Emma Marsden asked six women who live life to the full to tell us how they do it

A. Jeanette Kupferman

Journalist and author: 'I think it's excitement. I've got to be doing something that's a challenge. If I'm absorbing something new and learning, I get tremendous energy. Anything to do with rhythm gives me energy, too - at the moment I'm learning flamenco dancing. Dance is wonderful, very expressive and energising. I think you can feel drained if you're bored, maybe feeling life is pointless. Being effective in the world - even in a small way - gives you energy.'

B. Linda Kelsy

Magazine editor: 'I have so much work to do but I enjoy it all - I love being at work, it isn't a chore. If you enjoy something, it doesn't make you feel low even though you may be feeling tired. I need a fair bit of sleep so when I do get it I feel really good. When I'm running at six hours' sleep a night, I feel a bit wet. I know exercise gives me energy, but it's finding the energy to do it! I do a yoga class once a week before work and I try to do a workout tape at home. But if I've done exercise, my spirits rise amazingly. I get very low patches in the office between 3 and 5 in the afternoon and think it's probably good to eat something then.'

C. Annie Nightingale

Disc jockey: 'I don't eat healthily to get my energy - I think it's natural enthusiasm and being positive. I really can't stand people with negative attitudes. I'm quite naive - I think we should all enjoy our lives, grab the moment. Enjoyment comes first. You've got to achieve things and set yourself various goals. I enjoy what I do and the last two years have been very exciting. In my line of business there are lots of lively young people and you can't help being affected by that. I love tearing around the place on jobs, having a mad life. I'm very sociable. I enjoy taking photographs but I'm not sure if I'm good at it. I rush to get them processed and can't wait to get them back. I've also started writing words for imaginary songs - it's just for

fun and it's a good release for my mind when I'm doing a lot of travelling.'

D. Floella Benjamin

Author and children's TV presenter: 'My energy comes from things I do and enjoy. I believe in what I'm doing, both at work and at home with my family and I think that being positive about life helps. That's why I love working with kids, they give out so much and have an inbuilt resilience. Energy is like a natural chemical, triggered off by communicating with others - just try smiling at people and feel the thrill you get when they smile back. It's far better than feeling sorry for yourself. Although I try to eat healthily and don't smoke or drink, I don't have an exercise plan. The most energetic I get is running around with the kids and playing with them.'

E. Deborah Moggach

Author and playwright: 'What I love doing more than anything in the world is making a garden out of complete wasteland. I did that once. The garden was solid concrete - I smashed up all the concrete and dug in loads of soil. It was far better than any amount of workouts or tennis. The other thing that gives me energy is knowing that somebody wants to read what I'm writing - I find it difficult to write in a void. And I like sneaking into a farmer's land or a wood. If I'm found, I say "Oh, what a wonderful wood, I didn't know it was private property," and so on. I like it because it's childlike.'

F. Katherine Monblot

Therapist: 'I believe you must have an interest in and respect for what you do in life. I like to take risks and I get bored quickly, which keeps me motivated. Doing things you don't want to do wears you out. I used to be a member of various committees and resented the demands they made on my time. You only have a certain amount of energy, so direct it into the things that are most important to you.'