

# 子どもの英語力を測る：語彙テスト 開発の試み<sup>1</sup>

小林 美代子  
熊本大学

本稿では、子どもの英語力測定のためのテスト開発の過程と試行テストの分析結果を報告する。2011年から公立小学校の5、6年生を対象に正課として英語活動が開始する。英語活動実施の実態を踏まえ、音声に焦点を当てるとともに、子どもの言語発達における語彙の役割の重要性を考慮し（Cameron 2001、小椋 1999）、音声を中心とする英語語彙テスト（試行版）を開発した。Nation（2001）の語彙知識の理論的枠組みに基づく語彙知識評価の枠組みの作成、語彙の選定、テスト形式の選定などの観点からテスト開発過程を報告し、800余名の小学生による試行テストの分析結果に基づき、テストの信頼性・妥当性について討議するとともに、今後の研究課題についての示唆を考察する。

## はじめに

本稿は、子ども用英語力測定のための語彙テスト開発の試みについて、その開発過程とテストの概要、及び試行テストで得られたデータ分析結果の一部を報告するものである。本研究は、言語の発達と脳の働きの関係を探る大規模な研究の一環として<sup>2</sup>、子どもの英語力の発達段階を正確に測定する必要

---

<sup>1</sup> 本稿は、日本語学テスト学会第12回全国研究大会（2008年9月14日、於常磐大学、水戸）における口頭発表を基に加筆修正したものである。

<sup>2</sup> 神田外語大学言語科学研究センターでの委託研究（（独）科学技術振興機構、社会技術研究開発センターの研究開発プログラム「脳科学と教育」（タイプⅡ）『言語の発達・脳の成長・言語教育に関する統合的研究』（研究代表者：首都大学東京 萩原裕子）のサブ研究領域「言語学・応用言語学に基づく、外国語能力の検査、判定、評価法の開発」（研究機関代表者：長谷川信子；研究分担者：井上和子、小林美代子、堀場裕紀江））

性から始められた。そのためにまず、現在どのようなテストが存在するか俯瞰したが、その過程の中で、子ども用の英語力測定試験は、国際的に見ても非常に数が限られていること、また、熟達度テストが中心で、テスト結果は大まかなレベルを示唆するようなものが多く、子どもの英語力発達の細やかな段階に対する診断的な情報を得ることが難しいことが判明した（Cambridge ESOL, McKay 2006, Rixon 1999 他）。

2011年には、小学校5、6年生を対象とし、英語活動が正課として導入されることになり、このほど参考教材の『英語ノート』が全国の拠点校に配布された。英語活動の準備は少しずつ進められているが、教科としての導入ではないこともあって、統一的な学習の到達目標は示されておらず、教材や評価法は開発途上である。今後の小学校英語への先鞭をつける意味でも、診断的情報を提供するような独自のテストを開発することは有意義であると考えた。

テスト開発に先立って、以下のことを目指した。

- 1) 近年の言語能力モデルに基づく、信頼性と妥当性の高い子ども用テストを作ること。
- 2) 子どもの英語能力発達の段階を十分に反映し、テストの結果から、学習段階に応じた詳細なフィードバックを提供できること。
- 3) 今後、日本における小学校英語教育のための統一した到達目標設定のための参考となり、英語学習の効果を検証するための有効な道具の一つとしての役割を担えるようなものになること。

実際のテスト開発に当たり、まず、言語能力のうち、子供の言語発達上、最も顕著な部分に焦点を当てることとした。すなわち、子どもの母語習得及び外国語学習においては、語彙が大きな役割を果たすと報告されているため（小椋 1999、

McKay 2006 他)、本テストでは、語彙力に焦点を当てることにした。加えて、現在の小学校英語活動の現状を反映し、英語の4技能のうち、音声面を重点的に測定することとした。同時に、文字の学習が進んでいる児童には、文字に対する知識も測定できるよう工夫した。語彙の知識については、Nation(2001)や Read(2000)らの最近の語彙研究の理論を応用して、語彙の各側面を網羅するようにした。この点を次に説明する。

### 語彙テストの理論的枠組み

語彙知識は、語彙の「サイズ」と「深さ」という二つの側面から考えられることが多い(Nation 2001, Read 2000 他)。英語学習を始めて間がない子どもたちの語彙知識を測定しようとするとき、日本の英語教育の現状を考えると、子どもたちが習う単語の総数はせいぜい100語から300語程度で、非常に限定されている。そのような限られた語彙を評価の対象とするとき、知っている単語の数、つまり「サイズ」を測るだけでは、情報としてあまり意味をなさない。それよりも、ある単語をどの程度知っているのか、聞いてわかるだけなのか、自分でも発音できるのか、あるいは書くこともできるのかといった「深さ」の観点から詳細な情報を得られれば、個々の子どもの語彙力の発達段階を正確に特徴づけることができると考えられる。

そうした観点から、Nationの語彙知識の枠組みに準拠して、テストの枠組みを考えることとした。これによって、語彙知識を体系的に測定することが可能となる。この枠組みでは、語彙の知識を、まず認識と産出の領域の二つに大きく分け、それぞれに対して、音声と文字の二つの領域を設定した。これは、いわゆる言語の4技能に相当すると考えられる。それぞれに対して、言語知識の下位区分として、文字または音

素レベル、単語レベル、文レベル、そして、談話レベルの 4 つに区分した。これらのレベルのそれぞれをひとつひとつ確認していくことによって、子どもたちの英語語彙習得状況をかなり厳密に測定できることになり、よりの確なフィードバックを与えることも可能になる。次の表 1 は今回開発のテストで測定することを企図した項目を示したものである。

表 1：測定対象項目

一斉テスト

	認識		産出	
	音声（聞）	文字（読）	音声（話）	文字（書）
文字・音素レベル	✓	✓	—	✓
単語レベル	✓	✓	—	✓
文レベル	✓	✓	—	—
談話レベル	✓	—	—	—

個別テスト

	認識		産出	
	音声（聞）	文字（読）	音声（話）	文字（書）
文字・音素レベル	—	—	—	—
単語レベル	✓	—	✓	—
文レベル	✓	—	✓	—
談話レベル	✓	—	(✓)	—

この枠組みのうち、音声の産出は、4 技能のうち、言わばスピーキングに当たるものなので、一斉テストで測定することは困難である。この部分は、個別面談を行うことで補うこととした。一斉テストと個別面談の両方を実施することによって、語彙知識の枠組みのほぼすべてをカバーすることができると考えた。表中、<—>が入っている部分は、現在の日本の大方の小学生の英語学習の状況を考えて、今回の子ども用語彙テストには不適切であろうと判断した項目である。

本稿では、このうちの一斉テストについての開発過程と試行テストの結果を報告する。

## 試行テスト開発過程

Bachman & Palmer(1996)のテスト作成手順に従い、以下のよう  
な過程で試行テストを作成した。

- 1) まず、早期英語活動の実態を把握するため、国内外で出版されている英語教材、日本の公立小学校のうち特定学区として英語活動を行っている小学校の独自教材の分析を行った。また、公立小学校や民間英語教室での授業見学を行うと共に、個別の面談も行い、現状調査を進めた。
- 2) 上記のような様々な準備調査をした後、児童が知るべき語彙とはどのようなものなのかを特定するために、語彙リストを作成した。このリスト開発の過程に関しては、町田他（2008）を参照されたい。
- 3) 次に、テストの概要として、測定する内容、測定方法、問題数等を規定した。
- 4) 上記の枠組みに基づき、測定対象とする語彙項目を上記2)で作成した語彙リストの中から無作為抽出すると同時に、多肢選択問題の錯乱肢となる語彙項目も抽出した。その後、設問やイラストを作成した。
- 5) 子どもにもわかりやすい発音と声の調子などを十分に考慮して、英語母語話者に依頼して音声の録音をした。
- 6) テストの実効性等を確認するためにミニ試行テストを実施後、2008年2月より大規模な試行テストを開始し、順次結果の分析を実施した。

## 試行テストの概要と特徴

上述の過程を経て作成したテストの概要と特徴をまとめると

以下の通りとなる。

1. 筆記形式の一斉テストと個別面談の2部から成る。個別面談は、特に音声面での産出の力を測定し、一斉テストだけでは得られない情報を補うために実施する。
2. いずれのテストも、幅広い年代の子どもが受験することを想定して、絵をふんだんに使い、日本語の文字を読めない子どもにも対応できるようにし、子どもの日常生活に基づく、馴染みやすい場面設定をしている。
3. 語彙知識の理論に基づき、体系的・段階的に設問を作成したため、どの程度まで習熟しているかを診断するための情報を提供できるような形式になっている。

一斉テストは、音声のみの版、アルファベットを追加した版、さらに文字の読み書きに関する問題を追加した版の3種類を用意した。これは、協力者の時間的制約という現実的な問題に対応することと、文字学習を既に始めている児童と始めている児童に適切なテストを提供することが目的である。音声のみの版は、3部門から成っており、所要時間約15分、アルファベットを加えた版は、全部で5部門、所要時間約20分、読み書きまで含めた版は、全部で9部門、所要時間約30分である。便宜上、3種類のテストにSS、S、Lという名称をつけ、以下、結果を説明する際にこの呼称を用いる（設問例は稿末資料を参照のこと）。

## 試行テスト実施手順

### 協力者

試行テスト協力者は、総計779名、男子383名、女子396名で、性別比はほぼ均等であった。試行テスト実施期間は、2006年10月～2008年3月に亘り、協力者の内訳は東京都渋谷区

の公立小学校3校の児童、首都大学における大規模調査データ収集協力者、インターネット経由で応募してきた団体（英語塾等）及び個人である。

受験したテストの種類別による協力者数は、聞き取り部分のみ（SS版）が412名、聞き取りとアルファベット部分（S版）は122名、読み書きまで含めたすべてのテスト（L版）を受けた児童は245名であった。読み書きを含む版は、原則として、既に文字学習を始めている児童を対象に作成したが、実際には、文字学習を始めている児童もかなり含まれていた。聞き取り部分は、3種類すべてのテストに含まれているため、すべての協力者が受験しており、総計779名のデータが得られた。またアルファベットを含む版についても同様の理由から、367名のデータが得られた。学年別の協力者数は、以下の通りである。

表 2：協力者数（学年別）

	1年	2年	3年	4年	5年	6年
計	39名	133名	319名	165名	104名	19名
男子	17	72	162	74	55	3
女子	22	61	157	91	49	16

### 試行テストの結果

以下、試行テストの結果を報告する。まず、信頼性と弁別係数を見てみる。

表 3：信頼性係数及び弁別係数

	$\alpha$	$\alpha' *$	Mean Biserial
SS (k=15)	.622	.832	.593
S (k=25)	.798	.877	.680
L (k=45)	.901	(.901)	.662

\*Spearman-Brown Prophecy formula により、項目数を45問とした場合の信頼係数

表 3が示すとおり、テストの信頼性については、ほぼ満足の

いく値となっている。3種類のテストは項目数が異なるため、Spearman-Brown Prophecy formula を用い、便宜上項目数を揃えて推定計算したところ、どの種類のテストもかなり高い信頼性が想定された。また弁別係数も非常に高いものとなっており、学習の進んでいる児童とそうでない児童を明確に弁別することができるテストであることがわかる。15項目のみのSS版でも、推定前の値で、アルファ値が.622、弁別係数が.593となっており、項目数が非常に少ないことを考慮すると、満足のいく値と言えよう。

次に記述統計を示す。

表 4：記述統計

SS（聞き取りのみ）（k=15）

n	Max.	AVG (%)	SD (%)
779	15	10.58 (70.5)	2.51 (16.7)

S（聞き取り＋アルファベット）（k=25）

n	Max.	AVG (%)	SD (%)
367	25	19.77 (79.1)	3.89 (15.6)

L（聞き取り＋アルファベット＋読み書き）（k=45）

n	Max.	AVG (%)	SD (%)
245	45	29.67 (65.9)	7.66 (17.0)

聞き取り部分のみのテスト（SS版）の平均得点は約70%、アルファベットを加えたテスト（S版）は約80%、読み書きを含むテスト（L版）は約66%であった。読み書きが含まれるものは当然のことながら最も難易度が高くなっているが、全体的に本試行テストは易しめだったと言えよう<sup>3</sup>。

<sup>3</sup> すべての部門(L版)を受験した245名のみ得点を見た場合、SS部分は11.97(79.8%)、S部分は20.53(82.1%)で、ここに提示してある数字よりも若干高い結果となっていた。これは、L版を受験した学習者の多くが文字まで学習しており、また学年も高かったことに起因していると思われる。



次に、すべての部門を受験した 245 名の結果を詳細に見てみよう。

表 5：記述統計（部門別）（n=245）

			Max	AVG	SD
音声 （聞き取り）	単語認識	R	5	4.41	1.00
	文理解	R	5	4.36	0.91
	談話理解	R	5	3.20	1.21
アルファベット	聞き取り・認識	R	5	4.36	0.97
	聞き取り・産出	P	5	4.20	1.22
文字 （読み書き）	単語認識	R	5	3.33	1.37
	文理解	R	5	2.54	1.57
	単語綴り並べ替え	P	5	2.25	1.45
	単語産出	P	5	1.03	1.23

R: 認識; P: 産出

図 1は同じ情報をグラフにしたものである。

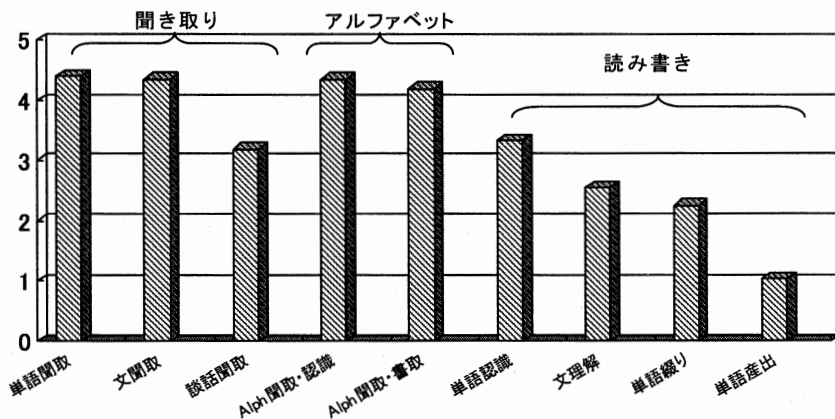


図 1：記述統計（部門別）（n=245）

扱う発話の量が増えるにつれて、また認識よりも産出の方が、難易度が高くなっている傾向が窺われる。聞き取りの単語レベルと文レベルの得点の間の差を除いて、すべて統計的にも有意な差となっていた。

文レベルの聞き取りと単語レベルの聞き取りとの間にほとんど違いがなかったことに関して、その理由を探究してみたところ、文レベルの聞き取り問題の中に、極端に易しい項目がひとつ含まれていることが判明した。そのため、総じて文レベルの聞き取りが易しくなってしまったことが考えられる。この項目を除外して再計算すると文レベルの平均点は 4.2 となり、単語レベルとの間の差は統計的に有意となった。つまり、当初想定した語彙知識の理論的枠組みが、項目の難易度として反映され、学習の度合いがほぼ正確に測定できるテストとなっていることが窺われる。

次は、学年別の得点を見てみよう。下の図 2 は音声とアルファベットの問題から成る S 版のデータを示すが、他の版もほぼ同じ傾向が見られた。学年が上がるにつれて得点が高くなっていることが見てとれる。

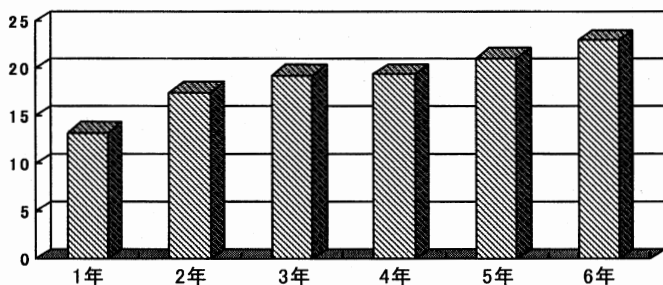


図 2：記述統計（学年別）(n=367)

次に英語学習の経験と得点との関係を見てみる。学年別得点同様、表 6 は S 版の結果を提示するが、他の版もほぼ同様

の結果であった<sup>4</sup>。

表 6：学習歴の有無（S版）（n=220）

経験あり (n=188)	経験なし (n=32)	<i>t</i>
20.70	17.97	3.888**

学校+私塾 (n=53)	学校のみ (n=87)	私塾のみ (n=48)	経験なし (n=32)	<i>F</i>
22.21	20.01	20.29	17.97	9.698**

\*\**p*<.01

Scheffe: 学校+私塾>経験なし\*\*；学校+私塾>学校のみ\*\*；  
私塾のみ>経験なし\* (\*\**p*<.01；\**p*<.05)

英語学習の経験がある児童と経験がない児童との間では得点に有意な差があった。ただ、経験のない児童の得点平均が、25点満点中18点となっている点に着目したい。「経験なし」の児童の多くが回答できた設問を検討したところ、日本語の中で日常的に使われている外来語の聞き取りやアルファベットの認識・産出の問題であった。これらの設問は、英語学習の経験がほとんどない児童でも正答できてしまう、ということを見ると、英語学習の効果を測定する目的にはあまり望ましい設問ではないと言えよう<sup>5</sup>。

次に、学習経験のタイプの違いを見てみる。「学校プラス英語教室などの私塾」で英語を学習している児童と、「学校のみ」、「私塾のみ」、「経験なし」の児童の平均得点を比較したところ、有意な差があった。「学校のみ」と「私塾のみ」の間には有意な差はなかったが、両方に行っている児童は、経験がな

<sup>4</sup> ここでのサンプル数は220で、これまでのテスト受験者数 (n=367) よりも少なくなっている。これは、学習歴に関するアンケートに回答しなかった児童がいたためである。

<sup>5</sup> ただし、今回の調査では、学習歴3ヵ月未満の児童は、「学習経験なし」とみなしたため、英語にまったく触れていないというわけではないことを断っておく。また、「学習経験なし」と回答した児童の10%程度は、片親が外国人であるなど、日頃英語に触れる環境にある子どもたちであった。

い児童だけでなく、学校のみで経験している児童よりも高い得点を取っていることが判明した。もともと学校以外で英語を勉強しようという児童は動機付けが高いこともあるかもしれない。英語教室などの方が、学校よりも学習内容が進んでいることも考えられる。さらに今後、学習期間、学習内容との関係なども調査していきたいが、少なくとも、今回の分析結果から、本試行テストは、学習の進んでいる児童とそうでない児童を弁別することのできるテストであるということはいえそうである。

次に IRT の分析結果を見てみる。

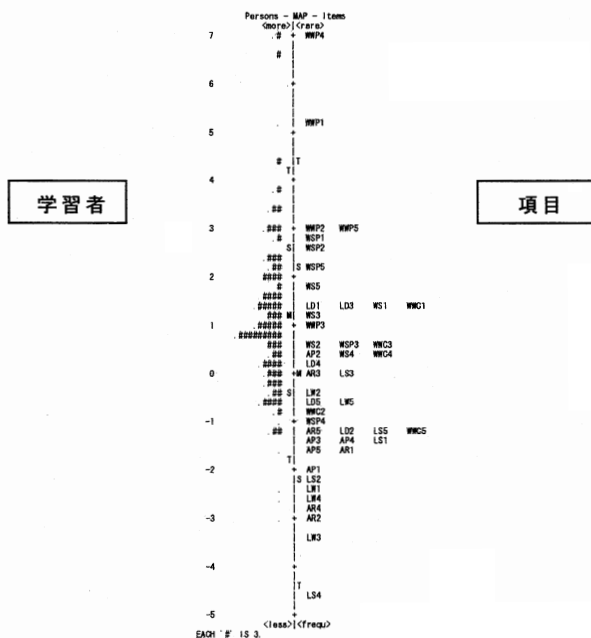


図 3 : IRT 分析結果 (n=245)

図 3が示すとおり、テスト項目は、全体的に学習者の能力に比べて易しめであったことがわかる。ただ、子ども用のテストであることを考えると、達成感を与え、動機付けを与えることも重要なので、難易度が低かったことはさほど問題とするべきではないかもしれない。ただ、極端に易しすぎて余剰な項目や、レベルが重複している項目は、削除したり、レベルを調整するために選択肢を修正するなどの工夫が必要であろう。極端に難しい項目が二つあるが、学習者の側を見ると、同様のレベルの児童が数名いるので、このレベルの難易度の項目は削除せず、残しておいてよいのではないかと推測される。

今後は、S版とSS版のみを受験した児童のデータも繋留することで、同じIRTマップの中に分布させ、今回開発したテストの難易度の適性についてさらに精査したいと考えている。

項目分析の結果、弁別度の高い項目として、以下のような項目が挙げられる。

- 単語の文字認識 例：door (.50), stone (.50)
- 文の文字認識 例：bank (.50), little (.55), climb (.71)
- 馴染みのある単語の綴り（並べ替え）例：banana (.80), lunch (.73), baby (.80)
- 既知語の文字産出 例：belt (.50), eat (.57)

これらの設問は、学習の進んでいる児童とそうでない児童を識別するのに有用な項目と言える。

また、弁別度の低い項目には、以下のようなものが挙げられる。

- 馴染みのあるアルファベット（馴染みのないアルファベットは問題なし）
- カタカナ語（英語学習経験がない子どもも知っている

語) の音声認識 pink (.18), sofa (.12), shirt (.48)

- 未知語の産出 build (.22), picture (.13)<sup>6</sup>

これらの項目は学習の程度を識別するのにあまり効果的でないことがわかる。

次にパート別・学年別平均点を見てみよう。音声部分の結果を図4に示す。

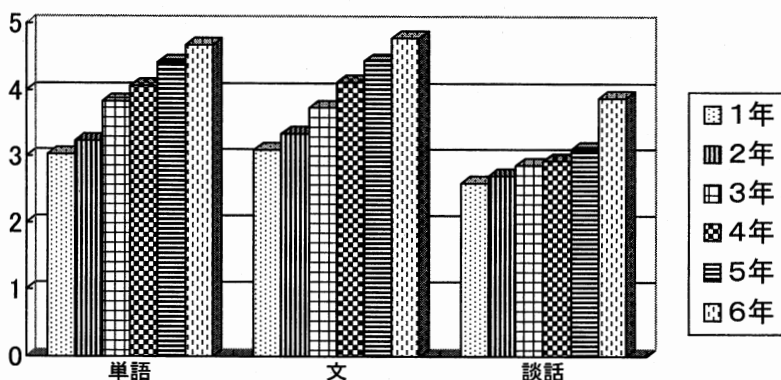


図4：パート別・学年別結果(音声部門) (n=779)

全体的に学年が上がるにつれて得点が高くなっているが、談話レベルは、全学年を通じて概して低い得点となっている。談話レベルの設問は、今回の試行テストに協力してくれた児童たちにとっては、学年に関わらず、困難なものだったということが言えよう。

アルファベットの認識に関しては、図5に示す通り、低学年でも比較的得点が高く、学年の差はあまりない。それに対して、同じアルファベットでも、聞いて書き取る問題(文字の産出)は学年によって大きく異なる。このことから、アル

<sup>6</sup> McNamara(1996)の基準 (Infit<0.7 or >1.3) に照らして検証した結果、misfitの項目が4問あった。談話の聞き取りが2問;文字認識文レベルが1問;単語の文字産出が1問であった。対象語が完全な未知語であった可能性がある。

ファベットを書かせる設問の方が児童たちの学習の度合いをより効果的に弁別できると言えよう。

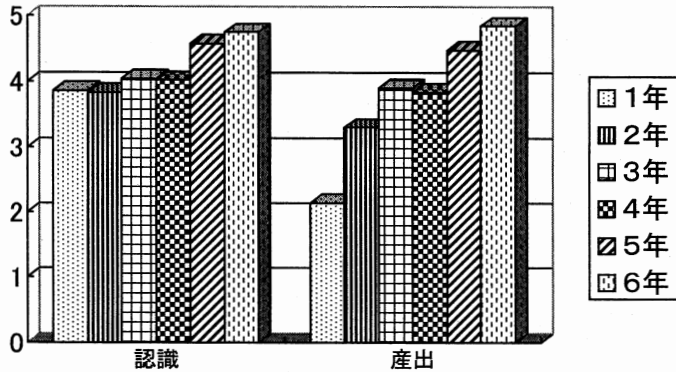


図 5: パート別・学年別結果(アルファベット部門)(n=367)

図 6は文字部分のパート別・学年別の平均点を示す。

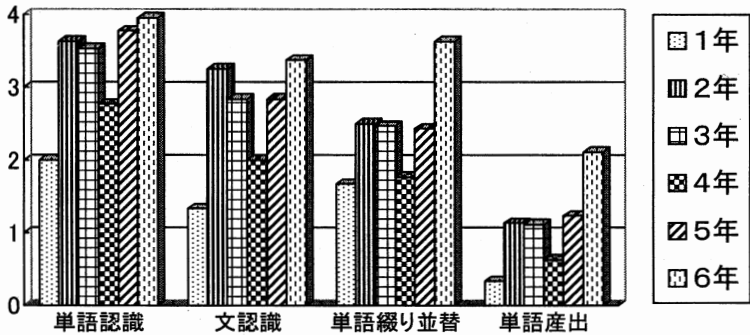


図 6: パート別・学年別結果(文字部門) (n=245)

単語の綴りの並べ替えと単語を書く問題で6年生が他の学年に比べて高くなっていることを除くと、学年による明確な傾向は見られない。特に、単語を書く問題は、学年を通じて、

非常に低い得点となっている。これは、まだ文字学習を始めていない子どもたちもこの部分のテスト（L版）を受験していることも一因となっているかもしれない。このデータに関しては、L版を受験した児童のみのデータ（ $n=245$ ）で、全体のサンプル数が少なく、各学年の人数に非常にばらつきがあるため、順当な比較が難しいことに留意する必要がある。今後、文字学習を始めている児童とそうでない児童を区別して、各々どのような得点を取っているか等、詳細な吟味をしていく必要があるだろう。

### 考察と結論

試行テストの分析結果から、今回開発した子ども用語彙テストについてまとめると、以下のようなことが言えよう。

- 概ね信頼性が高く、弁別力のある語彙テストである。
- 難易度も、少し易しめながら、子ども用ということを考えれば適切な難易度と言ってよいだろう。ただし、易しすぎる項目の重複部分を削除することでより適切なテストとなるであろう。
- 学年や学習歴を反映していたことから、学習の度合いをほぼ正確に測定できていると言えるだろう。
- 音声・文字、認識・産出といった語彙知識の枠組みに照らして、語彙知識の発達の程度を仔細に測定できている。
- 信頼性と弁別係数の分析結果から、音声部分のみの15分からなる15分程度のテストでも、ほぼ適正な評価ができることが判明した。これは、テスト実施の実際面でかなり意義のある示唆となるだろう。

最後に今後の課題について述べる。冒頭にも述べた通り、子どもの外国語能力を測定する評価法は、世界的に見てもまだ開発途上である。特に現在の日本のように英語指導の到達



目標が明確に示されていない場合、何を測定したらよいのかさえも不明で、評価を考えることは非常に難しい状況である。今回は、子どもの言語発達の初期段階で、母語習得においても外国語習得においても重要な役割を果たしていると思われる語彙の知識を測定する評価法開発を試みたが、より妥当性のあるテストを開発するために、今後次のようなことを検討していく必要がある。

まず、今回の試行テストでは、ある程度満足のいく高い信頼性を得ることができたが、最小限確保すべき問題数はどのぐらいなのか。これは、子どもを対象としたテストの場合、できるだけ拘束時間を短縮することが望ましいので、実際に意義のある課題である。

また、音声と文字のバランスはどうあるべきか、十分に検討する必要がある。文字の認識・産出の設問に関しては、総じて音声問題に比べてあまり得点は高くなかったが、弁別度は高いという結果が出た。現在、小学校では、原則として、文字指導は行わないことになっているが<sup>7</sup>、読み書きに関する設問に正答した児童がこれだけいたことを考慮すると、ある程度文字の問題を含めておいた方が、幅広い子どもの語彙知識を正確に測定することのできるテストになると思われる。実際に教室でどのような指導や活動がなされているのかといったことも検証しつつ、どのような問題タイプを含めるべきかさらに検討を重ねたい。

それと関連して、子どもの場合、学習している語彙の総数が非常に限られている中、既習の語がたまたまテストに出れば、その問題は非常に易しいものになってしまう。異なる題材やシラバスを使用してきた子どもたちに対して公平さを確保しつつ、妥当性のあるテストを作成していくことが求

---

<sup>7</sup> 文字指導の導入に関しては、まだ議論が分かれており、まったく扱わないという厳格な姿勢ではなく、文字に親しませることは許容されるようになってきているようである。

められる。同じことが、カタカナ語の扱いについても言える。今回、音声認識の問題で、pink という語を対象語とした問題が極端に易しかったのは、英語を学習していなくても、誰もが音声を聞いて認識できてしまったからであろう。それに対して、banana や lunch といった単語の綴りを書かせる設問は弁別度が高いという結果が出た。このことから、カタカナ語の扱いには留意する必要があることがわかる。今回の試行テストの結果からの示唆としては、カタカナ語は音声認識の問題からは排除することが望ましく、綴りを書かせる問題としては良問になる可能性が高いと言えるだろう。

今回提示した試行テストのデータをさらに詳細に分析すると同時に、個別面談のデータをもっと集めて分析することで、より精密な妥当性の検証をすることも有意義であろう。題材やテスト形式などについて、成人や子ども用の他の語彙テストとの比較検討を行うことや、対象語を拾い出す基になる語彙リストを充実させていくことも必要である。テスト項目の一つ一つを確認し、使用した語の品詞や意味タイプによる難易度や弁別力を検証することで、より汎用性のあるテスト開発に貢献したい。

## 謝辞

本研究は、独立行政法人科学技術振興機構社会技術研究開発センターの研究補助金、研究委託費による支援を受けて実施したものである。テストを受験した児童たち及び教育機関等、研究に協力してくれた多数の方々に感謝する。特に、神田外語大学教授長谷川信子氏及び研究助手町田なほみ氏には心からの感謝の意を表したい。また、テスト開発の初期段階でテストの枠組み作成に貢献してくれた堀場裕紀江氏及び Paul Joyce 氏にも感謝の意を表したい。

## 参考文献

- 小椋たみ子 (1999). 『初期言語発達と認知発達の関係』 東京 : 風間書房
- 町田なほみ・小林美代子・長谷川信子(2008). 「早期英語教育のための語彙リスト開発過程」『神田外語大学言語科学研究センター紀要』 7: 241-268.
- Bachman, L.E. & Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Cambridge ESOL <http://www.cambridgeesol.org/exams/young-learners/yle.html> (2009年12月25日アクセス)
- Cameron, L. (2001). *Teaching Language to Young Learners*. Cambridge: Cambridge University Press.
- McKay, P. (2006). *Assessing Young Language Learners*. Cambridge: Cambridge University Press.
- McNamara, T. (1996). *Measuring Second Language Performance*. Longman.
- Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Rixon, S. (1999). *Young Learners of English: Some Research Perspectives*. Essex: Longman.

資料：問題例

設問例：. 認識，音声，単語レベル

聞こえてくることばが表している絵を4つの中から1つ選び、□の  
中に○を書きましょう。



① □



② □



③ □



④ □

(script:potato)

860-8555

熊本県熊本市黒髪 2-40-1

熊本大学大学院

社会文化科学研究科

*mkobay@kumamoto-u.ac.jp*