

# 適応型言語能力検査 (ATLAN) の作成とその評価

高 橋 登\* 中 村 知 靖\*\*

本研究では、主に学童期の子ども達を対象とした、項目反応理論に基づく適応型言語能力検査を開発した。その際、幼児期までとは異なる、学童期に固有の言語能力である読み能力に焦点を当て、それと密接に関連する語彙と漢字の2種類の検査を作成した。語彙検査は幼稚園年中児～中学3年生まで計3,076名、漢字検査は小学校1年生～中学3年生まで計4,463名の結果に基づいて作成された。困難度・識別力という2つの項目パラメータを持った項目群からなる項目プールを用意し、スタンドアロンのパソコン上、およびネットワークを介してWeb上で動作する適応型言語能力検査(ATLAN)を開発した。項目プール作成時に実施した紙版検査の受検者の解答パターンを利用してシミュレーションを行ったところ、十分な再現性が得られ、ATLANが能力推定の面で問題がないことが示された。また、既存の語彙、漢字課題との間にも中程度の相関が見られ、検査として妥当性があることが示された。さらに、紙版の課題に比べ少ない問題数でありながら、受検者ごとに異なる問題、問題パターンが出題され、受検者の解答に応じて出題する問題を変える適応型の検査として機能していることが確かめられた。最後に、実用に向けた今後の課題について議論された。

キーワード：語彙、漢字、言語能力、適応型検査、項目反応理論

学童期の子ども達の言語能力を査定するには、従来 WISC-III, K-ABC 心理・教育アセスメントバッテリー、新版 K 式発達検査 2001 などの発達検査・知能検査や ITPA, 絵画語彙発達検査 (PVT) などの言語能力検査が広く用いられてきた。語彙に特化した PVT を除くと、これらの検査は多くの下位項目を含み、プロフィールから受検者の特徴を描き出すことができる。これは受検者である子ども達の個別支援を考える上で重要な情報を提供するものであり、これらの検査が広く教育や療育の場で利用されてきた理由であろう。その一方で、検査の実施には長い時間を要し、検査をする側、される側の双方にとっての負担は大きい。また、これらの検査は言語能力の測定という視点から見ると、測定論的な観点および現在の言語発達研究の成果との関係という点で必ずしも十分なものではない。

これらの検査は、いずれも古典的テスト理論に基づいて作られている。従来型の古典的テスト理論に基づく検査では、能力の幅の大きい子ども達に対し、基本的には単一の問題のセットを用意するだけなので、子ども達は困難度の幅が大きく、かつ数多くの問題に回答しなければならず、負担が大きい。また、テストで

は、項目は易しいものから難しいものまで一様に含まれるように工夫されるが、その結果、項目数の限界から受検者すべてに対して高い精度で測定を行うことは難しい。特定の難しさの項目を集めれば、一部の受検者に対して高い精度の測定を行うことは可能であるが、それ以外の受検者に対しては不十分な精度での測定となってしまうからである。

こうした古典的テストの持つ難点を改善する方法として、項目反応理論に基づいた適応型検査が考えられている(野口, 1985)。この検査では、予め求められた項目の特徴を表す項目パラメータを利用し、子ども達の反応に応じて最適な困難度の項目を出題することができる。そのため、古典的テスト理論に基づく検査に比べると子ども達の負担は少なく、しかも測定の精度を高めることが可能である。

項目反応理論を利用した適応型検査では、問題の提示とリアルタイムな能力推定が必要なため、コンピュータの利用が不可欠である。近年、コンピュータネットワーク環境が充実し、ウェブブラウザを利用した検査システムの開発が可能になった。このことにより、コンピュータネットワーク環境があれば、どこでも検査を受けることが可能になるとともに、検査結果の蓄積も容易になる。

項目反応理論を利用した言語能力検査については、すでに芝(1978)、芝・武井・荻野(1985)が幼児～学童

\* 大阪教育大学教育学部  
noborut@cc.osaka-kyoiku.ac.jp

\*\* 九州大学大学院人間環境学研究院  
tom@hes.kyushu-u.ac.jp

期を対象とした語彙検査を、平・小野・前川・林部・米山 (1995) が高校生程度の日本語能力測定のための語彙・漢字検査を開発しているが、コンピュータに実装した適応型検査の実現までは至っていない。そこで本研究では、学童期の子ども達を主な対象とする、項目反応理論に基づいた適応型言語能力検査を開発し、その実用的な可能性を評価することを目的とする。

ところで、従来の言語発達研究は、乳幼児期の言語獲得に焦点が当てられることが多く (例えば岩立・小椋, 2005 ; 小林・佐々木, 2008 ; 小椋, 2006 など参照), 発達支援もこの時期を中心に組み込まれてきた (西村, 2001 など)。学童期は、乳幼児期の言語獲得で主要なテーマとなる文法や音韻の処理に関しては、基本的な点についてはすでに獲得されているだけでなく、語彙獲得に関しても幼児期までの獲得のプロセスと同じ過程を想定することは困難である (高橋, 2006)。そうした意味で、学童期の言語能力は、この時期に固有の問題として考える必要がある。

幼児期とは異なる学童期に固有の言語能力として、すなわちこの時期に熟達する言語能力として読み能力を挙げることができる。高橋 (1999) は読みの過程が複数の処理が併行して進行する過程であり、そうした下位過程における処理の効率性が読み能力を説明するものであることを示してきた。さらに、高橋 (1999, 2001) は就学前から小学校5年生までの読み能力と、関連する要素となる諸能力との関係について縦断的な分析を行い、ひらがなの読みを習得して間がない幼児期～1年生段階では、文字刺激の符号化段階での処理の効率性、すなわち単語のスムーズな読みが読解に大きく関わるが、学年の上昇とともに読みの速度の個人差は少なくなり、また読解との関わりも薄れること、その一方で、一貫して読解に強く影響し続けるのは語彙であることを明らかにしている。そして、語彙が読解に影響するだけでなく、読解の能力が次の時期の語彙に影響するというように、語彙と読解は、語彙力が読解の能力を説明し、読解の能力は後の語彙力に影響するというようにスパイラルな関係にある。そうした点において、学童期の言語能力を査定する上で、もっとも基本となるものとして語彙を考える必要があるだろう。

一方、読み書きの能力に固有の障害として、発達性読み書き障害 (developmental dyslexia) が知られている。英語圏では、発達性読み書き障害は音韻意識 (phonological awareness) など、言語音の処理の問題と強く関連があることが明らかになっている (Frith, 1997 など) が、中国語児や日本語児の場合は、音韻意識の問

題はあるとしても英語圏ほど深刻であることは少なく、それとは別に、文字の形態的な処理の問題が関わっている可能性があることが指摘されている (Ho, Chan, Lee, Tsang, & Luan, 2004 ; Ho, Chan, Tsang, & Lee, 2002 ; 宇野・金子・春原・松田・加藤・笠原, 2002)。日本語や中国語の場合、読みの獲得に英語ほど洗練された音韻意識は必要とされないこと、その一方で、漢字のような複雑な形態を持った文字の処理が必要とされることがその原因となっていると考えられる。もしそうであるならば、日本語を母語とする学童期の子ども達にとっての言語能力として、漢字の問題も欠くことはできないだろう。

そこで本研究では、学童期の言語能力の基礎として、語彙と漢字を取り上げ、項目反応理論に基づく適応型の言語能力検査を開発し、その特徴を明らかにした上で、実用の可能性について評価することを目的とする。本論文では、最初に語彙、漢字それぞれの検査の作成過程について説明する。次に妥当性の検討を行った上で、読み能力との関連について検討する。

## 語彙項目プールの作成

### 項目の作成

国立国語研究所 (2001) に収録されている阪本教育基本語彙、新阪本教育基本語彙、田中教育基本語彙、池原基本語彙、教育基本語彙、学習基本語彙、および日本語教育のための基本語彙から幼児～中学生用の単語を抽出した。単語の抽出に当たっては、複数の語彙データベースに含まれる単語を選択するようにした。また、各データベースのそれぞれの単語には目安となる習得時期が示されているので、これを参考にしながら幼児用、小学校低学年 (1-3年) 用、高学年 (4-6年) 用、および中学生用の単語を選択した。さらに抽出の際には、(1)品詞が偏らないようにする、(2)問題の作成が容易なものを選択する、(3)幼児・1年生に関しては絵による提示が可能な語を選ぶ、といった点に留意した。

抽出した単語をもとに、幼児・小学校低学年・高学年・中学生用に合計 310 項目の問題を作成した。こうして作成された問題項目から、幼児版 2 種類、小学校低学年版・高学年版各 3 種類、および中学生版 2 種類を用意した。それぞれの版は対応する習得時期の単語で構成されていたが、それぞれの版には等化のために使用される共通の項目が含まれており、その中には習得時期の異なる項目も含まれていた。使用された版により、共通の項目に異同はあるが、全体の 49 項目は等化のための共通項目であった。

項目の内容は同義語、反対語を問うものであり、イラストを選択肢とする問題が 45 項目含まれている。回答形式は選択肢が 5 つの多肢選択形式であった。

#### 調査協力者

大阪府内の私立幼稚園児、公立小・中学校の児童・生徒計 3,076 名。Table 1 に調査協力者の内訳を示した。

#### パラメータの推定

調査協力者の項目への反応について正答を 1、誤答を 0 として得点化し、版ごとにデータをまとめ、項目間の四分相関係数を用いた因子分析(ミンレス法)を行った。計算には TESTFACT 4 を利用した。その結果、第 1 固有値が第 2 固有値と比べ、10 版の平均で 3.5 倍大きいことから、一次元性を仮定することができると判断し、次の項目反応モデルによる分析を進めた。

項目反応モデルの分析では、すべての学年ならびに版のデータをまとめ、調査協力者が解答していない版の項目に対する反応を欠測値として扱い、多集団項目反応モデル(2 パラメータロジスティックモデル)を利用して項目パラメータを推定した (Bock, Thissen, & Zimowski, 1997)。計算には BILOG-MG 3 を利用した。想定した母集団の構成は、幼稚園児・小学生低学年 (1-3 年生)・小学生高学年 (4-6 年生)・中学生である。通過率 (.02 以下ならびに .98 以上を削除)、識別力の値 (.3 未満を削除)、項目とテスト得点との双列相関係数 (.1 未満を削除)、項目特性曲線のデータとの当てはまり程度を考慮し、最終的に 268 項目を項目プールとした。学年ごとの能力値の平均と標準偏差を Table 2 に示す。

#### 漢字項目プールの作成

#### 項目の作成

「常用漢字の習得と指導」(国立国語研究所, 1994)、国語

**Table 1** 語彙課題項目プール作成の調査協力者

		人数	版数	各版の項目数
幼稚園	年中児	135	2	28
	年長児	145		
小学校	1 年生	368	3	30
	2 年生	325		
	3 年生	339		
	4 年生	360	3	45
	5 年生	390		
	6 年生	384		
中学校	1 年生	219	2	46
	2 年生	199		
	3 年生	212		

**Table 2** 語彙課題・漢字課題の学年別能力値の平均と標準偏差

学年	語彙課題	漢字課題
幼稚園		
年中児	-4.07 (0.75)	
年長児	-3.36 (0.57)	
小学校		
1 年生	-2.25 (0.63)	-2.47 (0.48)
2 年生	-1.83 (0.72)	-2.02 (0.72)
3 年生	-1.16 (0.92)	-1.45 (0.73)
4 年生	-0.47 (0.85)	-0.71 (0.69)
5 年生	0.03 (0.98)	-0.08 (0.88)
6 年生	0.42 (0.96)	0.06 (0.69)
中学校		
1 年生	1.02 (1.02)	0.24 (0.94)
2 年生	1.33 (1.14)	0.67 (1.17)
3 年生	1.68 (1.03)	1.21 (1.45)

教科書などを参考にして、小学校各学年配当の教育漢字および、中学校で学習する常用漢字の中から漢字を選択した。選択する漢字の数は、それぞれの学年で学習する漢字の数におおむね対応するようにし、その上で項目の作成が容易なものを選択した。ただし、実施できた調査との兼ね合いから、中学生用の漢字の割合は実際よりは少なくなっていた。こうして選択した漢字をもとに、合計 418 項目からなる問題を作成した。項目内容としては、大きくは漢字で表記された単語の読みを問うもの、ひらがなで表記された単語の漢字表記を問うものの 2 種類である。この他に、低学年用には絵を提示し、それに対応する単語の漢字表記を選択させる問題も含まれていた。語彙課題同様、回答形式は選択肢が 5 つの多肢選択形式である。漢字課題では、問題数を確保するために 2 つの時期に分けて問題の作成、調査の実施を行った。第一期では小学校低学年版、高学年版各 3 種類、および中学生版 2 種類 (こちらを旧版と呼ぶ) を用意し、第二期ではそれぞれ 3, 3, 1 種類 (こちらを新版と呼ぶ) を用意した。小学校低学年版・高学年版はそれぞれ小学校 1 年生～3 年生、4～6 年生までの配当漢字に基づいて作成された項目で構成され、中学生版も同様であった。ただし、それぞれの版には等化のために使用される共通の項目が含まれており、その中には配当学年以外の項目も含まれていた。調査時期、使用された版により、共通の項目に異同はあるが、全体で 26 項目が共通項目として使用された。

#### 調査協力者

大阪府内の公立小・中学校の児童・生徒計 4,463 名。Table 3 に調査協力者の内訳を示した。

### パラメータの推定

旧版についての因子分析による1次元性の確認では第1固有値は第2固有値に比べ、8版の平均で6.1倍大きかった。項目パラメータの推定についても基本的に語彙と同様であるが、漢字では新版も作成したため、旧版、新版それぞれで多母集団項目反応モデルによる項目パラメータを求めた後、旧版と新版との間で線形等化を行った。等化は旧版と新版に共通する13項目を利用し、旧版の項目パラメータを新版のスケールに変換する形で行われた。等化後、語彙と同様の基準で項目を選択し、最終的に405項目を項目プールとした。学年ごとの能力値の平均と標準偏差をTable 2に示す。

### 語彙・漢字問題項目の特徴

語彙、漢字問題の特徴を比較するために、それぞれについて、困難度と識別力について散布図を描いた (Figure 1, 2 参照)。漢字問題と語彙問題では識別力の分布に大きな違いが見られる。これは、語彙に比べ、漢字では学年による正答率の変化が著しい項目が多く含まれていることを示している。このことは、漢字の学習が学年の配当に厳密に従っているため、漢字問題の正答率が主に学校で未学習か学習済みかによって決まることによると考えられる。

### 適応型言語能力検査 (ATLAN: Adaptive Tests of Language Abilities)<sup>1</sup>の作成

語彙・漢字項目プールの問題を用いて、スタンドアロンのパソコン上、および Web で動作する適応型言語能力検査 (以下 ATLAN と略する) を開発した。パソコン版はデータ管理を Access 2003 (マイクロソフト社製) で行い、問題の提示に関しては Visual Basic (マイクロソフト社製) で行った。Windows XP 上で動作する。Web

Table 3 漢字課題項目プール作成の調査協力者

	人数	版数	各版の項目数
小学校	1年生	445	
	2年生	653	6
	3年生	581	45(30) <sup>注)</sup>
	4年生	660	
	5年生	643	6
	6年生	391	50(36)
中学校	1年生	368	
	2年生	355	3
	3年生	367	50(50)

注) 括弧内の数字は第1期で使用された版の項目数

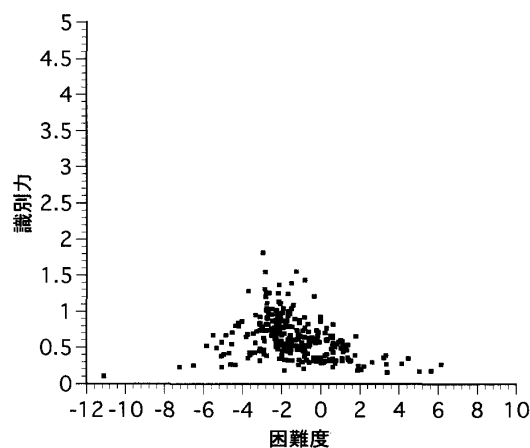


Figure 1 ATLAN 語彙項目の識別力と困難度の散布図

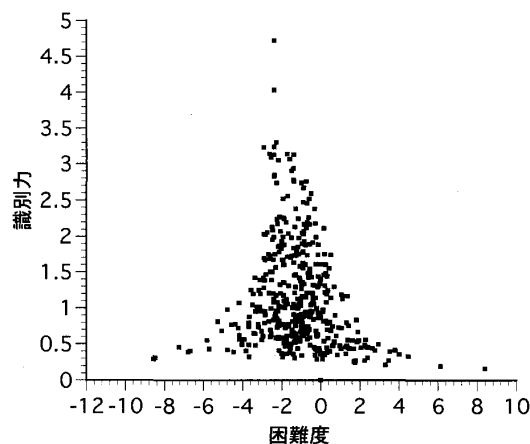


Figure 2 ATLAN 漢字項目の識別力と困難度の散布図

版は Perl および JavaScript で開発された。

ATLAN の流れをフローチャートで示すと Figure 3 のようになる。基本的な操作手順は以下のようになっている。1) オープニング画面 (検査の説明), 2) 受検者情報入力ならびに検査コース指定画面: 氏名・性別・生年月日・学年の入力と語彙・漢字など受検するコースを指定する。3) マウスクリック練習 (任意)。4) 項目提示: 終了基準が満たされるまで受検者

<sup>1</sup> 本論文では、語彙と漢字の検査しか作成しておらず、その点からすると、「言語能力検査」という名称は不適切であるように思われる。ただし、われわれの研究プロジェクト (高橋・大井・大伴・中村・小松, 2005) では、この他に文法 (大伴, 2005), 語用 (大井・高橋, 2005) についても適応型検査の可能性について検討し、試行的には Web 上で動作するようになっている。このような将来的な拡張の可能性も踏まえ、こうした包括的な名称を採用した。

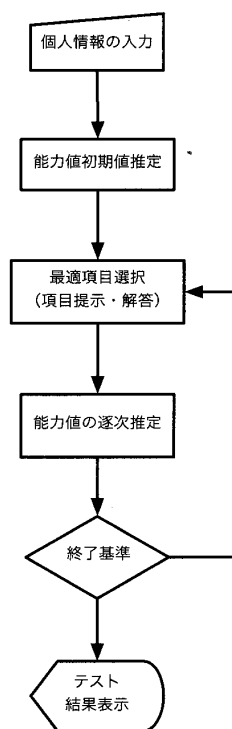


Figure 3 ATLAN のフローチャート

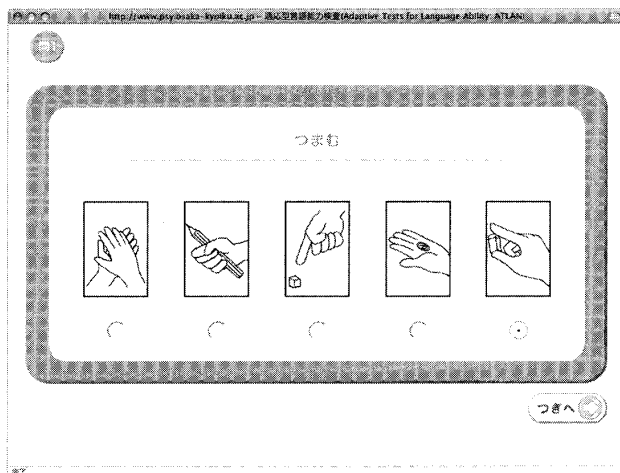


Figure 4 ATLAN 語彙検査 (イラスト提示) の問題例

の能力に適した項目が提示される。Figure 4, 5 に語彙課題の問題例を示す。5) 検査終了画面：教育現場において教師のみが結果を閲覧できるようにするため、検査結果を検査終了直後に表示しないようにしている。ここで結果表示を選択すれば画面上に検査結果が表示される。その際には、受検者の能力値と、各学年での平均能力値などの情報が出力される。また、各項目への解答状況と困難度も出力される。検査結果を印刷したり、ファイルに保存したりすることが可能である。

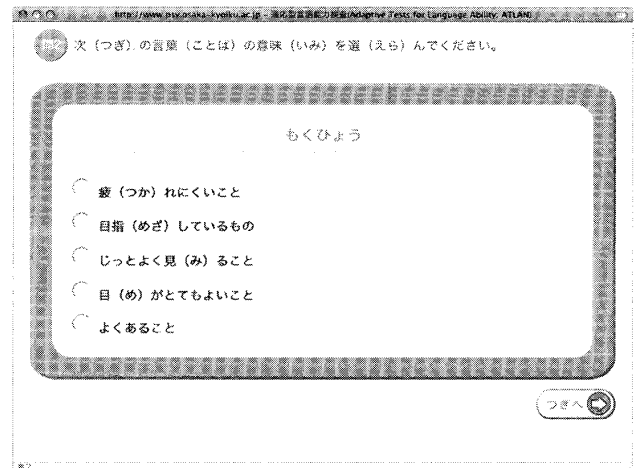


Figure 5 ATLAN 語彙検査 (文字提示) の問題例

この適応型検査システムにおいて、受検者能力の初期値は、受検者情報入力画面で得た学年をもとに、項目プール作成時に算出された学年ごとの能力分布の平均値としている。毎回の項目の選択は受検者の能力値に基づいて情報量が最大となるように行われる。また、受検者の能力値は各回の項目に対する反応をもとに逐次的に推定される。推定方法としてはベイズ推定(EAP推定量)が用いられており、第1問目における能力値推定のための事前分布は学年情報をもとにした能力分布が用いられている。検査の終了は、事後標準誤差が0.3以下あるいは出題数が50を超えた場合としている。

語彙検査・漢字検査とも同一のアルゴリズムに基づいて動作するが、両検査の項目の識別力の特徴から、収束までに要する問題数に大きな違いが見られた。動作確認を行うと、語彙検査では10～20問で収束するのにに対し、漢字では数問で、場合によっては1問正解ただけで収束してしまった。漢字検査は検査の実用性の点で問題があるので、識別力が極端に低い、あるいは高い項目は項目プールから除外し、また細部の調整を行い、極端に少ない問題だけで収束してしまうことのないようにした。この結果、ATLAN 語彙検査は最終的に268項目、漢字検査は256項目の問題から構成されることとなった。

### 再現性の検討

ATLAN がもとの紙版の検査結果をどの程度忠実に再現しているかを確認するために、紙版の語彙検査10版それぞれに関して、調査協力者のデータから約10%、292名を無作為に抽出した。次に、それぞれの版で用いられた項目のみで構成されたATLANを用いて、抽出された協力者の回答パターンに基づいた回答

でATLANによる能力値測定シミュレーションを行った。版ごとに、紙版の結果に基づいて算出された協力者の能力値とATLANによる能力値の相関を求めたところ、.83～.98の相関が得られた。また、各版は複数の学年の協力者に実施しているので、学年の要因を除いた偏相関を求めたところ、.76～.98の相関が得られた。10の版のうち、7つで.90以上の相関であり、ATLAN語彙検査では紙版の結果が十分に再現されることが確認された。

漢字検査に関しても同様の検討を行った。紙版15版(旧版8版, 新版7版)のそれぞれについて、調査協力者のデータから約10%, 合計424名を抽出し、抽出された協力者の回答パターンに基づいて算出された協力者の能力値とATLANによる能力値の相関を求めたところ、.68～.98, 学年の影響を除いた偏相関で.70～.98であった。15の版のうち、7つで.90以上の偏相関が得られ、ATLANの再現性は漢字検査においても確認された。

### 妥当性の検討1：絵画語い発達検査

(上野・撫尾・飯長, 1991)との関連

幼児期についてATLANの妥当性を検討し、その特徴を明らかにするために、ATLAN語彙検査と絵画語い発達検査(上野他, 1991)を実施し、比較する。

#### 方法

**調査協力者** 大阪府内の私立幼稚園年中児(31名, 男児15名, 女児16名, 平均年齢:5歳4か月), 年長児(30名, 男児15名, 女児15名, 平均年齢:6歳5か月)

**課題** 絵画語い発達検査(上野他, 1991)およびATLAN語彙検査を個別に実施した。

#### 結果と考察

絵画語い発達検査の年齢群ごとの平均の語彙年齢と標準偏差は、年中児:62.5か月(標準偏差:9.36か月), 年長児:82.5か月(標準偏差:19.91か月)であった。また、ATLANは年中児:-3.67(標準偏差:0.67), 年長児:-2.87(標準偏差:0.46)であった。また、それぞれの年齢群ごとに絵画語い検査の成績(語彙年齢)とATLANの得点の相関を求めたところ、年中児が.39, 年長児が.47と、それぞれ中程度の相関が見られた( $p<.001$ )。

絵画語い発達検査で終了基準まで実施した平均の問題数は、年中児:45.06(標準偏差:11.58), 年長児:53.67(標準偏差:13.14)であったのに対し、ATLANでは、年中児:13.06(標準偏差:2.61), 年長児:7.97(標準偏差:2.08)であり、ATLANの方が少ない問題数で終了していた。この点において、ATLAN語彙検査は絵画語

い発達検査よりも簡便に実施可能であることが示された。また、ATLANで用いられた項目は延べ644問、使用された問題は33項目であり、61人中すべて同一の問題が使用されたのは3組だけであった。ATLANが受検者の回答に応じて問題を提示する適応型の検査として機能していることが確かめられた。

### 妥当性の検討2：語彙・漢字課題

(天野・黒須, 1992)との関連

小学校段階におけるATLANの妥当性を検討するために、天野・黒須(1992)が国語・算数の学力の全国調査を行った際に用いた語彙課題と漢字課題、およびATLANの語彙・漢字検査を実施し、関連を検討する。

#### 方法

**調査協力者** 大阪府内の公立小学校4年生46名, 5年生46名, 6年生36名が本調査に参加した。

**課題** ATLAN語彙検査・漢字検査、および天野・黒須(1992)の語彙課題、漢字課題を実施した。ATLANは、小学校内のパソコンルームで集団実施し、語彙課題と漢字課題はATLANとは別に、学級単位で集団実施した。

#### 結果と考察

最初に、学年ごとに各課題の平均と標準偏差を算出した。また、ATLANに関しては、併せて収束するまでに要した平均の問題数と正答数を求めた(Table 4)。Table 4から、漢字・語彙のいずれにおいてもATLANは天野・黒須の課題に比べ、少ない問題数で収束することがわかる。また、ATLANの2つの課題では、漢字検査の方が語彙検査の半分程度の問題数で収束した。このことは、漢字検査の項目の方が全体に識別力が高いことを反映した結果であると考えられる。

次に、各課題の成績について、学年の要因を除いた偏相関を求めた(Table 5)。いずれの課題の間でも中程度の相関が見られた。紙版の漢字・語彙課題の間でも同様の相関が見られており、漢字と語彙の能力は類似性が高い。

ATLAN語彙検査で使用された延べの問題数は2,387, 使用された問題は74項目であり、125の問題パターンが出現していた。すべて同一の問題だったのは3組だけであった。また、ATLAN漢字検査では、延べ1,261問、使用されたのは79項目, 出現した問題パターンは105であり、すべて同一の問題で受検したのは23組であった。受検者の回答に応じて適切に問題が選択されていたと考えられる。

**Table 4** ATLANと紙版語彙・漢字課題の平均と標準偏差

	4年生	5年生	6年生
ATLAN語彙			
問題数	14.33 (4.95)	18.70 (6.89)	24.11 (7.80)
正答数	7.09 (4.39)	9.83 (5.37)	14.06 (6.53)
成績	-1.24 (0.79)	-0.79 (0.85)	-0.11 (0.82)
ATLAN漢字			
問題数	8.87 (0.54)	9.72 (0.91)	11.28 (6.51)
正答数	4.26 (1.00)	3.98 (0.65)	4.69 (5.95)
成績	-0.93 (0.44)	-0.57 (0.46)	-0.23 (0.60)
語彙課題			
問題数		36	
成績 <sup>注)</sup>	61.82 (14.48)	63.13 (19.42)	75.20 (12.60)
漢字課題			
問題数		120	
成績	79.48 (10.60)	90.85 (13.45)	113.75 (6.37)

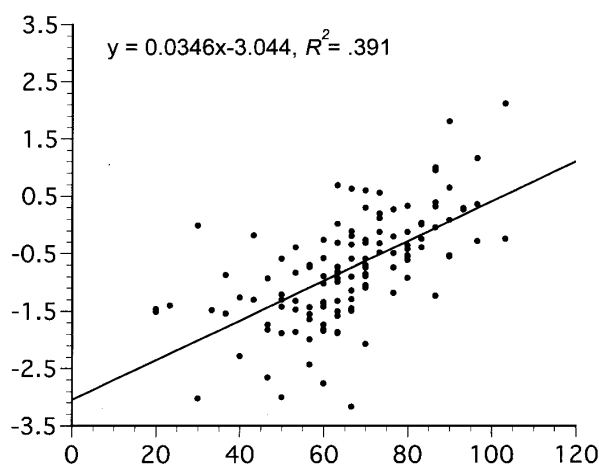
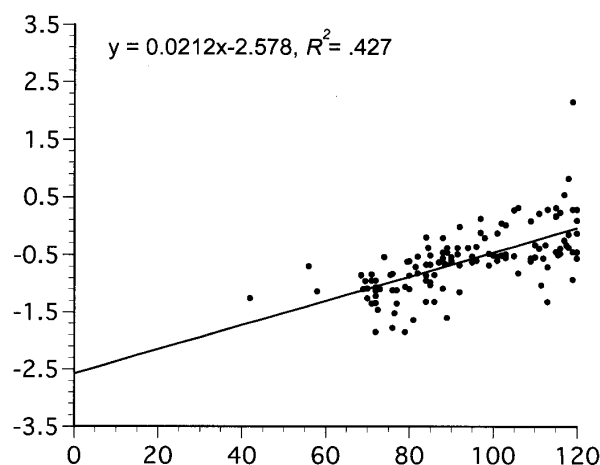
注) 天野・黒須 (1992) に従い、120 点満点に換算した。

**Table 5** ATLANと紙版語彙・漢字課題の学年の要因を除いた偏相関

	ATLAN語彙	ATLAN漢字	語彙課題	漢字課題
ATLAN語彙		.566**	.573**	.510**
ATLAN漢字			.544**	.494**
語彙課題				.636**
漢字課題				

\*\* $p < .01$

また、語彙・漢字の分布の特徴を見るために、ATLAN 語彙検査と語彙課題、ATLAN 漢字検査と漢字課題について散布図を描いた (Figure 6, Figure 7)。2つの散布図を比較すると、漢字の方が散らばり具合は少なく、また漢字課題では得点が満点近くの者がいることがわかる。天野・黒須 (1992) は語彙・漢字課題ともに各学年の国語教科書に基づいて作成しているが、習得状況は語彙と漢字で異なっているように思われる。漢字はばらつきが少なく、ATLAN でも漢字検査各項

**Figure 6** ATLAN 語彙検査と語彙課題 (天野・黒須, 1992) の散布図**Figure 7** ATLAN 漢字検査と漢字課題 (天野・黒須, 1992) の散布図

目の識別力が高いことを踏まえるならば、現在の学童期の子供も達は、語彙に比べ、漢字を均質な形で、したがっておそらくは学年配当に基づく学校教育の中で習得していると考えられる。一方、漢字と異なり、語彙の場合は授業や教科書のみで十全に習得されるわけではなく、具体的には日常の読書を通じて徐々に正確度を増す形で獲得している (高橋, 2006) ことを反映した結果であると考えられる。

## ATLAN 語彙検査と読書、読み能力との 関連性の検討

高橋 (1999, 2001) によれば、読解の能力を強く規定するのは語彙である。また、学童期の語彙は読解の能力を基盤とする読書によって高まると考えられている (高橋, 2006)。そこで、天野・黒須 (1992) の読解課題と ATLAN 語彙検査に加え、読書量を尋ねる質問紙を実施し、ATLAN 語彙検査の得点が読解課題の成績を予測しうるものなのか、また読書量が語彙検査の成績をどの程度予測するものなのかを検討する。

### 方法

**調査協力者** 和歌山県内の小学校に通学する 4 学年の児童 (3・4・5・6 年生) を対象とした。それぞれの人数は以下の通りであった：3 年生 (71 名, 男子 43 名, 女子 28 名), 4 年生 (82 名, 男子 38 名, 女子 44 名), 5 年生 (61 名, 男子 31 名, 女子 30 名), 6 年生 (89 名, 男子 41 名, 女子 48 名)。

**課題** ATLAN 語彙検査および天野・黒須 (1992) の読解課題、および読書に関する質問紙を実施した。ATLAN は、小学校内のパソコンルームで集団で実施した。4 年生以上は学級単位で実施したが、操作の説明の必要から、3 年生は 5 名程度に分けて実施した。担任教師と学生の補助者が検査の実施に当たった。

読解課題と読書の質問紙は学級単位で集団で実施した。読解課題 (天野・黒須, 1992) は、小学校 1 年生から 6 年生までの国語教科書の中から説明文、物語文、あるいは詩を各学年につき 2 種類ずつ選択し、それをもとに問題文を作成したものであった。実施手続きは天野・黒須 (1992) と同様である。また、読書の質問紙

は、読書量に関する質問項目 2 項目 (読書量 (冊数)：1 か月平均の読んだ本の冊数を「全く読まない～10 冊以上読む」の 11 段階で評定、読書量 (時間)：週平均の読書時間を「全く読まない～8 時間以上読んだ」の 8 段階で評定)、および読書に対する態度に関する質問項目からなっていた。ただし、本研究では読書量に関する 2 項目についてのみ分析する。

### 結果と考察

最初にそれぞれの学年ごとに読解課題、ATLAN 語彙検査、読書量 (冊数)、読書量 (時間) の平均と標準偏差を求めた (Table 6 参照)。次に、学年を込みにした各課題・年齢間の相関を求めた (Table 7 参照)。Table 7 の相関は、欠損値に関してはペアごとに除外して計算したものである。先行研究 (高橋, 2001；高橋, 2006 も参照のこと) と課題間の相関から、読解能力は語彙によって説明されること、語彙は読書量と年齢によって説明されることが予測されたので、この 2 段階の (重) 回帰分析を行った。後者に関しては、時間と冊数のどちらが語彙に影響するのか明らかではなかったので、両変数を用い、年齢と合わせ、3 つの変数を説明変数とするステップワイズ法による重回帰分析を行った。その結果、読解能力は語彙によって、また語彙は年齢と読書量 (時間) によって説明されることが示された (それぞれの重決定係数は .383, .430,  $p < .001$ )。

そこで、このモデル全体の適合度を評価するために、共分散構造分析を行った (Figure 8 参照)。適合度の各指標の値もよく、説明率も高いので、このモデルは適切なものであると考えられた。この結果から、ATLAN 語彙検査が学童期の語彙を測定する上で妥当性を持つこと、また語彙検査の成績は年齢と読書時間で説明さ

**Table 6** 読解課題、ATLAN 語彙検査、読書量の平均と標準偏差

	3 年生	4 年生	5 年生	6 年生
読解課題 <sup>a)</sup>	56.77 (19.14)	73.64 (17.75)	71.58 (19.57)	78.53 (15.48)
ATLAN 語彙	-1.98 (0.86)	-1.13 (0.86)	-0.37 (0.86)	-0.12 (0.87)
読書量 (時間)	2.64 (1.42)	2.81 (1.41)	2.59 (1.22)	3.02 (1.45)
読書量 (冊数)	6.69 (3.13)	5.14 (3.35)	5.15 (3.26)	4.11 (2.71)

注) 120 点満点

**Table 7** 課題間および年齢との相関

	読解 課題	ATLAN 語彙	読書量 (時間)	読書量 (冊数)	年齢
読解課題		.619**	.193**	.032	.350**
ATLAN 語彙			.208**	-.097	.641**
読書量 (時間)				.348**	.109
読書量 (冊数)					-.257**
年齢					

\*\* $p < .01$



れることが示された。

### 全体的考察

本研究では、主に学童期の子ども達を対象とした、項目反応理論に基づく適応型言語能力検査を開発してきた。その際、幼児期までとは異なる、学童期に熟達する言語能力である読み能力に焦点を当て、それと密接に関連する語彙と漢字の2種類の検査を作成した。それぞれについて、困難度・識別力という2つの項目パラメータを持った項目群からなる項目プールを用意し、スタンドアロンのパソコン上、およびネットワークを介してWeb上で動作する適応型言語能力検査(ATLAN)を開発した。

項目プールを作成する際に使用した紙版と同一の項目のみを用いたATLANを用意し、項目プール作成時に実施した紙版検査の受検者の解答パターンを利用してシミュレーションを行ったところ、十分な再現性が得られ、ATLANが能力推定の面で問題がないことが示された。また、既存の語彙、漢字課題との間にも中程度の相関が見られ、検査として妥当性があることが示された。さらに、紙版の課題に比べ少ない問題数でありながら、受検者ごとに異なる問題、問題パターンで出題され、受検者の解答に応じて出題する問題を変える適応型の検査として機能していることが確かめられた。最後にATLAN語彙検査を用いて読書量、読解の能力との関係を検討し、それぞれの関係を説明する適切なモデルが得られた。この点から、実用的にも問題ないことが確かめられた。

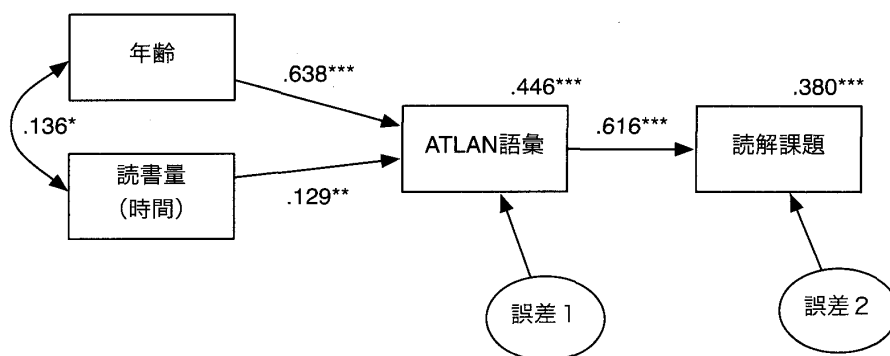
ATLANの実用に向けた次のステップはどのような形でこれを公開するか、ということであるが、現時点ではネットワーク上で無償公開することを考えている。ただし、受検者である子ども達の個人情報扱うことにもなるので、利用者に関しては一定の制限を設

けることが必要であると考えている。また、検査の改善に必要なので、検査結果はサーバにも保存されることに同意した上での利用ということにしたい。詳細については次のアドレスを参照してほしい：<<http://www.psy.osaka-kyoiku.ac.jp/atlan.html>>

最後に今後の課題について整理しておきたい。第1の課題は、問題数についてである。適応型検査の場合、広範囲の能力に関して精度の高い測定を行うためには、項目プールに識別力の高い項目を特定の困難度には偏らないように追加していく必要がある。そうした点から見ると、現状は十分な問題数が確保されたとは言いがたい。第2に、本論文では語彙と漢字についてのみ報告したが、われわれの研究プロジェクトでは、この他に文法(大伴, 2005)、語用(大井・高橋, 2005)についても適応型検査の可能性について検討し、試行的にはWeb上で動作するようになっている。こうした検査の実用の可能性に関しては、今後も検討を進める必要があると考えている。

### 引用文献

- 天野 清・黒須俊夫 (1992). 小学生の国語・算数の学力 秋山書店
- Bock, R. D., Thissen, D., & Zimowski, M. F. (1997). IRT estimation of domain scores. *Journal of Educational Measurement*, **34**, 197-211.
- Frith, U. (1997). Brain, mind and behaviour in dyslexia. In C. Hulme & M. Snowling (Eds.), *Dyslexia : Biology, cognition and intervention* (pp. 1-19). London : Whurr.
- Ho, C. S-H., Chan, D. W-O., Lee, S-H., Tsang, S.-M., & Luan, V. H. (2004). Cognitive profiling and preliminary subtyping in Chinese develop-



\*\*\*  $p < .001$ ,  $p < .01$

$\chi^2(2) = 2.77$ , ns GFI = .995, AGFI = .973, RMSEA = .039

Figure 8 読解と語彙能力の間のパス図

- mental dyslexia. *Cognition*, **91**, 43-75.
- Ho, C. S. H., Chan, D. W. O., Tsang, S. M., & Lee, S. H. (2002). The cognitive profile and multiple-deficit hypothesis in Chinese developmental dyslexia. *Developmental Psychology*, **38**, 543-553.
- 岩立志津夫・小椋たみ子 (2005). よくわかる言語発達 ミネルヴァ書房
- 小林春美・佐々木正人 (編) (2008). 新・子どもたちの言語獲得 大修館書店
- 国立国語研究所 (1994). 常用漢字の習得と指導 東京書籍
- 国立国語研究所 (2001). 教育基本語彙の基本的研究: 教育基本語彙データベースの作成 明治書院
- 西村辨作 (編) (2001). ことばの障害入門 大修館書店
- 野口裕之 (1985). 心理・教育テストとパーソナルコンピュータ 心理学評論, **28**, 423-440. (Noguchi, H. (1985). Personal computer and educational & psychological testing. *Japanese Psychological Review*, **28**, 423-440.)
- 小椋たみ子 (編) (2006). 特集子どもの言語獲得 心理学評論, **46**. (Ogura, T. (Ed.) (2006). Special issue: Child language. *Japanese Psychological Review*, **46**.)
- 大井 学・高橋 登 (2005). 語用の能力の発達とその評価 高橋 登・大井 学・大伴 潔・中村知靖・小松孝至 パーソナルコンピュータを用いた適応型言語能力検査の作成 平成 14-16 年度科学研究費 (基盤研究(B)) 報告書
- 大伴 潔 (2005). 統語的・談話的側面の評価とその発達—学齢前半期を対象とした検討— 高橋 登・大井 学・大伴 潔・中村知靖・小松孝至 パーソナルコンピュータを用いた適応型言語能力検査の作成 平成 14-16 年度科学研究費 (基盤研究(B)) 報告書
- 芝 祐順 (1978). 語彙理解尺度作成の試み 東京大学教育学部紀要, **17**, 47-58. (Shiba, S. (1978). Construction of a scale for acquisition of word meanings. *Bulletin of the Faculty of Education, the University of Tokyo*, **17**, 47-58.)
- 芝 祐順・武井澄江・荻野美佐子 (1985). 幼児の語彙理解尺度作成の試み 東京大学教育学部紀要, **24**, 47-60. (Shiba, S., Takei, S., & Ogino, M. (1985). Construction of a scale for acquisition of word meaning in early childhood. *Bulletin of the Faculty of Education the University of Tokyo*, **24**, 47-60.)
- 平 直樹・小野 博・前川真一・林部英雄・米山千佳子 (1995). 高校生程度の日本語能力テストの開発: 語彙理解テスト・漢字読み取りテストの尺度化 教育心理学研究, **43**, 68-73. (Taira, N., Ono, H., Maekawa, S., Hayashibe, H., & Yoneyama, C. (1995). Development of Japanese language tests for senior high schoolers' level: Vocabulary test and Chinese character reading test. *Japanese Journal of Educational Psychology*, **43**, 68-73.)
- 高橋 登 (1999). 子どもの読み能力の獲得過程 風間書房
- 高橋 登 (2001). 学童期における読解能力の発達過程—1-5年生の縦断的な分析— 教育心理学研究, **49**, 1-10. (Takahashi, N. (2001). Developmental changes in reading ability: A longitudinal analysis of Japanese children from first to fifth grade. *Japanese Journal of Educational Psychology*, **49**, 1-10.)
- 高橋 登 (2006). 学童期の語彙能力 コミュニケーション障害学, **23**, 118-125. (Takahashi, N. (2006). Vocabulary development in elementary school children. *Japanese Journal of Communication Disorders*, **23**, 118-125.)
- 高橋 登・大井 学・大伴 潔・中村知靖・小松孝至 (2005). パーソナルコンピュータを用いた適応型言語能力検査の作成 平成 14-16 年度科学研究費 (基盤研究(B)) 報告書
- 上野一彦・撫尾知信・飯長喜一郎 (1991). 絵画語い発達検査 1991 年修正版 日本文化科学社
- 宇野 彰・金子真人・春原則子・松田博史・加藤元一郎・笠原麻里 (2002). 発達性読み書き障害—神経心理学および認知神経心理学的分析— 失語症研究, **22**, 130-136. (Uno, A., Kaneko, M., Haruhara, N., Matsuda, H., Kato, G., & Kasa-hara, M. (2002). Developmental dyslexia: Neuropsychological and cognitive-neuropsychological analysis. *Higher Brain Function Research*, **22**, 130-136.)

## 謝 辞

本研究にご協力いただきました園児, 児童, 生徒の

皆さん，また先生方に深く感謝します。本研究は科学 の助成を受けた。  
研究費（基盤研究(B)課題番号 17330141 研究代表者：高橋登）

(2008.5.9 受稿, 11.15 受理)

## *Development and Evaluation of the Adaptive Tests for Language Abilities (ATLAN)*

NOBORU TAKAHASHI (OSAKA KYOIKU UNIVERSITY) AND TOMOYASU NAKAMURA (KYUSHU UNIVERSITY)  
JAPANESE JOURNAL OF EDUCATIONAL PSYCHOLOGY, 2009, 57, 201–211.

The present research used item response theory to develop the Adaptive Tests for Language Abilities (ATLAN) for elementary school children. These tests, which are available on-line and also for stand-alone computers, are comprised of 2 subtests : vocabulary and kanji reading, both of which are componential language skills needed for Japanese reading comprehension. Children from preschool to junior high school ( $N=3,067$ ) participated in determining the parameters of each item in the vocabulary subtest, whereas 4,463 children from elementary to junior high school participated in the development of the kanji reading subtest. Because the results of a simulation using the Adaptive Tests for Language Abilities adequately replicated the original data, the parameter estimation was judged to be reliable. The tests also have validity, because the results from each subtest moderately correlated with existing vocabulary and kanji tests. Advantages of the Adaptive Tests for Language Abilities are that they need fewer items than existing tests for measuring these abilities, and that different patterns of items are shown for different children.

Key Words : Adaptive Tests for Language Abilities (ATLAN), testing vocabulary, testing kanji reading, item response theory, test development and evaluation