

# 項目応答理論に基づくテストの得点 — J-CATの得点換算・解釈・利用法について —

今 井 新 悟  
伊 東 祐 郎  
中 村 洋 一  
菊 地 賢 一  
赤 木 彌 生  
中 園 博 美  
本 田 明 子  
平 村 健 勝

## 要旨

J-CAT (Japanese computerized adaptive test)は受験者の能力に最も適切な難易度の問題が出題され、4セクションで100点満点、合計400点満点の得点で結果を返す。この得点のもととなるのが、項目応答理論による能力推定値であり、それは概ね-3.5から+3.5の範囲になることがデータから分かっている。この値を換算式にあてはめ、100点満点の得点に線形変換する。換算式の係数を決めるためのシミュレーションを行い、どのように得点換算が行われるか、及びその得点の解釈と利用法について述べる。

## キーワード

J-CAT, 適応型テスト, 日本語, コンピュータテスト, 項目応答理論

### 1 はじめに

J-CAT (Japanese computerized adaptive test)は、WEB上で実施する日本語能力測定のための適応型テストである。受験者の能力に応じて出題される問題アイテムの難易度が計算され、受験者の能力に最も適切な難易度の問題が、アイテムプールと呼ばれる問題アイテムのデータベースから選ばれ、出題される仕組みを持つ。その他J-CATの概要については、今井 (2006) やJ-CATプロジェクトのホームページを参照していただきたい。なお、J-CAT2008年度版が<http://www.j-cat.org/>で公開されており、個人登録で受験が可能である。

### 2 得点換算

J-CATでは、項目応答理論を用いて、各問題(アイテム)の項目困難度(難易度)と項目識別力の値を算出している。困難度と識別力の値が付加された問題(アイテム)が、アイテムプールに収納されており、受験者が解答するごとにその受験者の能力が回答パターン、困難度、識別力によって推定され、能力レベルに最も近い困難度値を持つアイテムがアイテムプールから選ばれて出題される。これを繰り返し、能力推定が安定してきて、測定誤差が一定の範囲に入ったと判断されたとき、あるいはテスト管理者が定めた出題数に達した時点での値を、その受験者の最終能力推定値として確定する。項目応答理論により算出

された難易度と識別力は受験者集団やアイテムの困難度に左右されない不変的な指標となる。そしてそこから推定されてくる能力値もまた、受験者集団やアイテムの難易度に左右されない値である。よって、異なる時期に受験しても、また、異なるアイテムを受験してもその能力値の比較ができる。

能力推定値はアイテムを作成後、事前のプレテストを実施して、その解答パターンから推定された困難度および識別力のパラメータの値に依拠している。よって、「項目応答理論の推定値が受験者集団の影響を受けない」という意味について以下の補足しておく。これは、アイテムプールができたあとでのことであり、アイテムの困難度と識別力の推定の際には、受験者集団の性質の影響を受け、その値はプレテストを受験した受験者集団の能力に依拠する。

であるから、他の統計推定と同様、アイテムのパラメータ推定の際のプレテストにおいては、標本抽出は母集団を反映しているように注意して実施する。J-CATは日本語学習者の日本語能力を測定するテストであるから、母集団は日本語学習者全員と想定される。よって、プレテストでは、日本語学習者のうち初級レベルにあるものから上級レベルにある者までを受験者としている。そうすることによって、標本としてのプレテスト受験者集団を、実際にJ-CATを受験する日本語学習者の母集団に近づけることができる。

もちろん、プレテストの受験者集団とJ-CAT受験者集団の整合性を図っても、日本語学習者の母集団の能力値の平均も分散も本来は不明ではある。厳密には、他の統計的推定の性質と同様、標本としてのプレテスト受験者集団の能力の分布が母集団の分布の完全な縮図になっているとは言い切れないものの、サンプリングが無作為に行われ、それが母集団の分布を反映していると仮定されるならば、困難度、識別力、そして能力推定値は、母集

団分布での値を推定しているとしてよい。プレテストにより推定された困難度は原点を持ち、それを平均的な能力の者が受験すると正答する確率は50%ということになる。よって、逆に、困難度が原点、つまり0のアイテムを確率50%で正答する受験者の能力は平均「的」といえる。困難度同様正負に伸びる比例尺度である能力推定値の原点も母集団の平均に「近い」ということになる。

なお、それぞれのパラメータの尺度が等間隔になっていることは保証されることから、たとえ受験者母集団の能力の平均・分散が未知であっても、能力推定値そしてJ-CATの結果として示される得点の比較は常に可能である。古典的テスト理論においても点数は間隔尺度であるが、天井効果や床面効果を考慮すると、その尺度は能力の違いを適切に等間隔で反映しているとはいえない。一方、得点の天井に近い者にはボーナス得点を加算し、得点の床面に近いものにはペナルティとして得点を差し引くような、別の言い方をすれば得点にウェイトをかけるような操作をしているのが項目応答理論での能力推定値である。このように、異なる受験者集団間での得点の比較が可能となることから、「項目応答理論の推定値が受験者集団の影響を受けない」と言えるのである。

能力値は、項目応答理論で算出され、J-CATにおいては、概ね-3.5から+3.5の範囲となるのがこれまでのデータから分かっている。しかし、この値のままでは、一般の受験者およびテスト利用者はその意味を理解することはできないので、一般になじみのある100点満点の値に変換することが望ましいと考えた。項目応答理論で求まる能力値は0を原点とする比例尺度である。よって、それを一次変換しても間隔は変わらないので、偏差値と同じようにして得点換算ができる。J-CATでは、以下の換算式を用いて得点を算出している。

得点=最終能力値×15+50

これで、聴解、語彙、文法、読解の4セクションの得点が、概ね0点から100点の間に入る。そして総得点は0点から400点の間に入る。

この得点換算式の係数15の値については、以下に述べるシミュレーションの結果から導きだされたものである。この係数は任意のものであるが、それが小さすぎると換算される得点の幅が狭く、得点に差がつかず、出題されたアイテムに正答を繰り返しても高得点が取れないということになる。また、係数が大きすぎると、設定したい得点幅（ここでは、0点から100点）を超えてしまう例が多く出てくることになり、不都合が生じる。そのような不都合が最も少なくなるような係数を探るため、困難度と識別力の実データを用いてのシミュレーションが必要になる。

### 3 シミュレーションと得点調整

先に、能力値が概ね-3.5から+3.5になると述べたが、項目応答理論では、能力値の値は正または負の無限大である。よって、すべての回答パターンのケースにおいてこの範囲に入るかどうかの検証が必要となる。そこで、全問正解と全問不正解の解答パターンをシミュレーションした。

このシミュレーションでの条件は次のように設定した。

- (1) アイテムプールはJ-CATの実際のものを使用する。
- (2) 能力推定値の誤差が一定値よりも小さくなった場合でもテストを終了しない。
- (3) 回答アイテム数が20に達したらテストを終了する。（ただし、能力推定を3段階で行っており、逐次推定部分には最大13アイテムを使っている。）

1番目の条件として、シミュレーションの依拠するデータとして、実際のアイテムの値（パラメータ）を使用することとした。今後、アイテムプール内へのアイテムの追加や削除が行われて、アイテムについているパラメータが異なることになれば、シミュレーション結果も変化することが予測される。よって、アイテムプールに変化がある場合に、このシミュレーションを再度実施し、最終推定能力値の変化を調べることが必要となる。特に最大値と最小値について、その変化に注意する必要がある。

2番目の条件は、全問正解あるいは全問不正解の極端な状況での能力推定を行うために設定した。実際のJ-CATでは、測定誤差が十分小さくなったと判断される範囲に入るとしてテスト終了のストップルールのひとつとしているが、このシミュレーションでは最終推定能力の最高値と最低値を予測したいため、測定誤差をストップルールとせず、テストが継続されるように設定した。その結果としての能力推定値は実際のJ-CATで想定されうる両極端の値であり、実際の最終推定能力値はこの最高値と最低値の間の値に必ず収まる。

3番目の条件は、現行のJ-CATの運用時の設定に則したものである。J-CATでは、もうひとつのストップルールとして出題数を各セクション最大20に設定している。回答数が20に満たないうちに、測定誤差が一定の範囲に入れば、回答数は20よりも少なくなり、測定誤差がその範囲に入らずに回答数が20に達した場合にはそこでテストを打ち切る。この回答数の上限をあまり少なく設定すれば誤差が大きいままにテストが終了してしまう。逆に、上限を大きくすれば、いずれ測定範囲に入るものの、テスト時間がとても長くなってしまふという弊害が生じる。能力推定時の誤差変動をシミュレーションした結果から、誤差は測定の始めの段階で大きく下がり、そ

の後下がり幅は小さくなりながら、誤差0を目指して漸近的に減少を続けることが分かっている。また、シミュレーションから、20アイテムを回答した時点で、概ね、設定誤差値に近づくことも分かったため、これを回答数の上限とした。これについても、今後、アイテムプール内のアイテムの数の変化があれば、ここでの最終能力推定値のシミュレーションを再度行い、その結果に注意することになる。

シミュレーションの結果は以下ようになった。得点の小数点以下は四捨五入して丸めである。

聴解

全問正解時  
最終推定能力値=2.809 →得点=92  
全問不正解時  
最終推定能力値=-3.453 →得点=-2

語彙

全問正解時  
最終推定能力値=2.821 →得点=92  
全問不正解時  
最終推定能力値=-3.176 →得点=2

文法

全問正解時  
最終推定能力値=3.326 →得点=100  
全問不正解時  
最終推定能力値=-3.239 →得点=1

読解

全問正解時  
最終推定能力値=3.031 →得点=95  
全問不正解時  
最終推定能力値=-3.071 →得点=4

最終推定能力値から各セクションとも得点が、概ね0点から100点の範囲に入ることが明らかになった。唯一聴解で全問不正解とな

った場合に得点が負の値を取っている。負の値になった場合は切り上げて得点を0点として調整することにした。これにより、最低得点は、聴解が0点、語彙が2点、文法が1点、読解が4点となる。最高得点は、聴解92点、語彙92点、文法100点、読解95点となる。そして、総得点は最低点が7点、最高点が379点となる。これは、理論的に取りうる得点範囲であり、実際のテストでは最低点は、テストが4肢選択形式であることから、偶然に正答を選ぶ確率が25%あり、理論値よりも一般に高くなることが予測される。実際のJ-CATの2008年度後期の最低得点は64点、最高得点は実験のための日本人受験者も含めて365点となった。

4 解釈

J-CATの結果(得点)は上述の通り、項目応答理論での能力推定値を元に換算されたものである。プレテストの標本抽出において日本語学習者母集団が反映されるように配慮していることから、得点が各セクションで50点、総合で200点は日本語学習者母集団全体の平均に近く、それは「いわゆる」中級レベルに相当するだろう。ただし、この点については下に述べることに注意する必要がある。

J-CATでは、特に個人受験者・利用者のために得点の解釈の目安を示している。

100点未満	初級
100~150	中級前半
150~200	中級
200~250	中級後半
250~300	上級前半
300以上	母語話者相当

それぞれの分割点やレベル名はこれまでJ-CATをプレースメントテストとして実施し、クラス分けに利用してきたY大学留学生セン

ターにおける調整の結果である。よってこれはY大学留学センターにおける解釈である。日本語教科書で用いられている初・中・上級のレベル名もまた任意のものであるが、それとここに示したレベル名はほぼ一致するものと考えている。しかし、ここで強調しておきたいのは、得点の解釈はあくまでも利用者の目的に応じて行われるべきであるということである。教育機関が異なれば、レベル名や分割点は当然異なる。特に日本国内と海外でのレベルは差が顕著である。よって、ここに示すレベル名も分割点も任意のものであり、目安として利用されるべきもので、絶対的なものではない。特に、J-CATの使用者が教育機関団体である場合には、ここに示した基準に捉われることなく、J-CATの得点をその機関のレベル分けに合わせて再解釈して使用するべきである。

J-CATの得点に対するレベルの区切り方が適切ではないという指摘を受けたことがある。これも上記で述べた通り、分割点の設定をその教育機関にレベル分けに則して設定しないことによる誤解である。

現行のJ-CATが保証しているのは得点の不変性であり、レベル分け自体や解釈ではない。例えば、TOEFLやTOEICの点数をもって、単位認定の条件としたり、入学の要件としたりする大学もあるが、何点以上でその要件を満たすとするかの設定は各教育機関に任せられており、一律の基準はない。TOEFLやTOEICの実施者であるEducational Testing Serviceが提供するの是不変的な指標としての得点だけであり、その解釈と利用方法は各教育機関に委ねられている。これと全く同じことがJ-CATの得点とその解釈にも当てはまる。将来的にはJ-CATにおいて言語能力記述文 (Can-Do Statements) を併記することによって、解釈と利用のための詳しい情報が提供できるように整えていきたい。

## 5 利用法

### 5.1 ケーススタディ (1) : S大学

S大学では、各学期開始時において、受講生のレベルチェックを目的として、インターネット上でJ-CATを利用している。大学の日本語教育の規模が小さく開講クラス数が少ないという事情から、各クラスを受講する留学生の身分も様々で、日本語能力にもどうしても幅が出る。また、近年では交流協定校からの短期交換留学生が増えているが、たとえ学習歴の長さが同じ学生でも、各国の日本語教育内容が様々であるため、能力が大きく異なっている場合も少なくない。以上のような状況にあるため、学期開始時にJ-CATを実施し、その結果から受講生個人とクラス全体のレベルを把握し、その後のクラス運営に役立てている。また、大半の学生が前期、後期（あるいは後期、前期）と続けてクラスを受講するが、そうすると約半年経過して再びJ-CATを受けることになる。そこでアダプティブテストの特徴的な点が見えてくるのだが、日本語力が伸びている学生は、テスト開始からしばらくすると前回とは異なる問題に挑戦することになる。レベルチェックを目的としてJ-CATを利用する観点から、この点は紙のテストと異なるよい特徴だと感じる。ペーパーテストであれば、前回と同じ内容と分量の問題を解かなければならないところを、能力が上がっていればより難しい問題が与えられるのであるから、学生自身が自分の能力向上を実感することにもつながっているようである。受験終了と同時に、得点が即座にわかる点も、教員にとって大変便利である。30人程度でもこの便利さを実感するのであるから、これが100人以上にでもなれば、コンピュータテストであるからこそこの利点は、さらに強く感じられるのではないだろうか。また、得点は技能別にグラフでも提示されることから、学生は自分自身の能力のバランスを視覚的に

把握することもできる。(参考資料(2)参照)

また、S大学では毎年12月に実施されている日本語能力試験の受験を奨励・支援しているが、受験申し込み期間中、まだ来日していない学生のために、申し込み手続きの代行も行っている。その際、受験級を決めてもらわなければならないのだが、受験希望者の中には日本語能力試験がどのようなものであるかを知らず、適当な受験級を自分では判断できない学生がいる。このような場合もJ-CATを海外で受験してもらい、結果を参考にレベルを推定し、適当と思われる受験級を助言することになっている。現時点では、まだ日本語能力試験とJ-CATの得点の相関は明確になっていないが、これまでに受験した学生のテスト結果と照らし合わせれば、おおよそのレベルが判断できるため、このような使い方も可能となっている。

S大学でのJ-CAT実施には、約20~30人が一斉に学内の情報処理室で受験する形をとっている。大学からSINETを介し、J-CATのサーバにアクセスして受験しているため、動作の不具合がほとんど見られなく、これまでに大きなトラブルはなかったが、受験途中で画面が「リロード中」となり、そのまま動かなくなったと申し出た学生が数名いた。しかし、その画面はすぐにクローズすることができ、ひとつ前に戻って再度読み込むことで、問題なく継続することができた。また、受験にかかる時間は、約40分~80分で幅があり、高い得点をとる学生ほど長くなっている。このことは、レベルの低い学生が不要に能力以上の難しい問題を解かされることがなく、学生の受験負担を適当なものにしているということを示しており、この点もアダプティブテストの利点であるといえよう。

## 5.2 ケーススタディ (2) : A大学

A大学では、試験運用を始めたところであ

る。当初、高速のSINETで学外とつながっていない環境でインターネットからJ-CATサーバへアクセスして実施を試みたが、同時受験者が多くなると回線の容量を越えるためか、画面表示のトラブルが生じるようになった。そこで、ノートパソコンを簡易サーバとして使うことで、学内LANで実施することとした。この方法で、通信の不具合は解消された。この方法で8日間の受験期間を設定し、74名が受験した。受験者から、以下のフィードバックが得られた。

- ・問題の難易度の差が大きい。難しい問題のあと、非常に簡単な問題がでてくると、前の問題が不正解だったのだとすぐわかる。
- ・読解の文字が小さい。
- ・読解の指示文と問題本文の区別がつきにくかった。
- ・解答時間が短すぎる。

これらは、インターフェースデザインに関わるものである。今後さらにフィードバックを集め、検証を重ねて、デザインの改良に役立てることができるだろう。

これまでの受験者の成績と教室活動の観察による評価は相関しているというように感じている。この点については、今後統計的な裏付けを行って、検証していく予定である。

## 5.3 ケーススタディ (3) : Y大学

### 5.3.1 日本語プレースメントテストとして

Y大学ではJ-CATを日本語授業のクラス分けのための日本語プレースメントテストとして利用している。2004年のJ-CAT開発と同時に運用を開始し、すでに5年間、プレースメントテストとして利用している。Y大学では、毎年4月と10月の2回、新留学生が到着する。そのほとんどの留学生は留学生センター開講の日本語授業を受講する。このため、新留学生到着時の4月と10月の年2回、2ヶ所のキャンパスのコンピュータ室でプレースメントテストを実施している。プレースメントテス

ト実施前に、新留学生学内オリエンテーションを行い、日本語時間割、日本語授業のレベル、プレースメントテストなどについて説明を行っている。受験時間は、学習者のレベルによっても異なるが、40分から80分程度で終了する。プレースメントテスト結果はテスト終了後、コンピュータ画面上に瞬時に表示される。各自の点数をJ-CATレベル表で確認した上で、レベルに合った授業を受けることになっている。なお、日本語未学習者は、プレースメントテストを受験する必要がないことを伝え、日本語入門レベルのクラスを受講するようにアドバイスを行っている。また、来日が遅れるなどの理由から、プレースメントテスト実施日に受験できなかった留学生も、時間の都合のつくときに随時受験することができる。このように、通年を通して来日する留学生にいつでも受験してもらうことができ、日本語能力測定が即座にできることから、留学生個々の日本語能力に合った日本語学習が可能となっている。

J-CATはWEB上での受験となるため、Y大学の留学生が学内のコンピュータでインターネットにアクセスするためには、原則として大学のコンピュータを利用するためのIDとパスワードを取得していなければならない。しかし、来日したばかりの新留学生はIDを取得していない上、日本語のコンピュータ操作にも慣れていない場合が多い。このため、プレースメントテストの実施時には、臨時の共通IDを発行し、留学生センター教員らがコンピュータ操作のサポートに当たっている。また、遅れて到着する者に対してはIDを取得後にJ-CATを受験させている。なお、J-CATの受験にはJ-CAT用パスワードも必要である。プレースメントテストでは、団体受験用パスワードを一斉に配布し受験させる。

Y大学では、日本語を前期・後期継続して受講する留学生は、J-CATを学期ごとに受講しなければならない。しかし、半年ごとのJ-

CATの成績では、伸び率が顕著ではない場合もある。例えば、日本語能力試験では、各級の学習時間数は表1の通りであるが、Y大学の日本語の授業時間数は、各レベル最大週3コマで68時間程度である。よって、授業時間数だけから言えば、半年で級をまたいで伸びることは難しい。また、J-CATは日本語授業のシラバスに即し作成した到達度テストではなく、言語運用能力テストである。よって、授業の到達度テストとして使うべき性格のものではない。J-CATをプレースメントテストとして利用する場合も、このようなJ-CATの特性を踏まえ、適切にクラス分けを行う必要がある。J-CATでは、テスト開始前、先ず留学生の日本語学習歴、日本語能力試験などの日本語テストの受験歴などを記入しているが、J-CATの結果と受験者の学習歴から予想される結果が大幅に異なっている場合があれば、再受験してもらうなどして、そのずれの原因を検証するなどの対応もしている。J-CATをプレースメントテストとして利用するのは、あくまでも日本語を受講してもらうためのレベルチェックとしてであり、日本語学習支援を行うための日本語能力測定ツールとして活用しているのである。

**表1 日本語能力試験級別学習時間数**

級	学習時間数
1級	900時間程度学習したレベル
2級	600時間程度を学習し、中級日本語コースを修了したレベル
3級	300時間程度の学習し、初級日本語コースを修了したレベル
4級	150時間程度の学習し、初級日本語コースを修了したレベル

『日本語能力試験出題基準』より作成

### 5.3.2. 奨学金受給者選考のための日本語能力測定として

Y大学には留学生のための奨学金は複数あり、そのうちのひとつの奨学金選考に、J-CAT日本語テストの成績を利用し、受給者を選考している。応募する留学生は、4月と10月に実施されるJ-CATを2回受験し、その得点を留学生支援室に提出する。目的は、日本語学習を奨励するものであり、2004年度J-CAT開発をスタートさせて以来、利用している。

### 5.3.3 渡日本前留学希望者の日本語能力の判定として

留学生受け入れに際し、日本語能力を証明する資料がない場合など、指導教員が留学生の日本語能力を判断することが難しいケースもある。このような場合、来日前にJ-CATを受験することで、指導教員が受け入れ留学生の日本語能力を把握することによって、日本語学習支援や指導内容を考慮するなどに役立てることができる。指導教員がJ-CATの点数を参考に学習支援に利用するケースが増えつつある。

現在、協定校からの交換学生（特別聴講生）が増えつつあるが、日本語能力は初級レベルから上級レベルまで幅広い。日本語能力の低い留学生には、渡日前日本語学習支援を行なうことも考えられるが、その際にもJ-CATで日本語能力をチェックし、現在開発中のJ-CAT語彙デジタル練習問題やEラーニング教材などを用い日本語学習支援を行っていくことができる。特に、日本留学を希望しているが、日本語を学ぶ機会が少ない（あるいは、ない）中国などアジアの理工系留学生を対象とした渡日前日本語学習に活用できると考え、中国などの協定校と協議を行っている。J-CATは今後、留学生30万人受け入れ計画によって増加すると考えられる留学生の日本語能力測定を留学生の負担を軽減できる方法で

渡日前に実施することができる画期的な試みと言えよう。

### 5.3.4 受験履歴・成績証明書

J-CATの受験履歴はコンピュータ上にデータとして保存されており、過去に受験した留学生の成績も確認できる。研究生が大学院入学に際し、日本語能力を証明する資料を学部から求められることがあり、J-CAT受験証明書がほしいという留学生からの依頼がある。そのような場合、受験履歴、日本語受講歴などを確認し証明書を出している。

日本語能力試験は、現在年1回しか行われていないため、来日後すぐに日本語能力を測る手段がないのが現状である。その場合、J-CATであれば、申し込み手続きなどの手間もかからず、どこでもいつでもインターネットに接続できるコンピュータさえあれば即受験でき、留学生の日本語能力を証明することが可能である。

### 5.3.5 実績統計 (2008年度)

#### 5.3.5.1 受験者数

Y大学では、毎回100名前後の留学生が受験している。日本語プレースメントテストは、全学で広報し、受験を呼びかけているが、4月と10月に新たに到着する新留学生には、新留学生学内オリエンテーションとJ-CATの日時を知らせ、全員に受験を薦めている。2008年度Y大学留学生数は309名である。受験者数は、前期109名、後期99名、合計208名であった。学部別受験者数(表3, 4)は、従来通り、留学生数の多い経済学部、工学部の受験者数が多いが、近年、日本語専攻などの文系の留学生が増え始めたことによって、教育、人文学部の受験者数も増えている。身分別受験者数(表5)からも分かるように、中国、韓国などアジアの協定校からの交換学生が増え始め、2008年度後期は、研究生、大学院生を抜いて最も多い受験者数となった。国別受



験者数（表6）は、Y大学でも圧倒的に中国が多い。ほとんどの交換学生も中国、韓国の協定校からであることも中国の留学生が増えている理由である。また、学内使用だけではなく、国内、海外の協定校やプレテスト協力校からの受験も開始しており、今後受験者数はさらに増えていくものと期待される。なお、表2の「海外」受験者は協定校およびプレテスト協力校に対して団体用のパスワードを発行してそれで受験した者の数である。それ以外に表2には示していないが、個人で登録し、個々にパスワードの発行を受けて実施する個人受験が2008年度は100人程度いた。

表2 2008年度受験者数（人）

機関名	前期	後期	1月	計
Y大学	109	99	0	208
S大学	28	23	21	72
国内その他	12	12	12	12
海外	11	11	11	11
総計		303		303

表3 Y大学2008年度前期学部別受験者数

	学部生	大学院生	研究生	特別聴講生	研究者	未記入	合計
人文科学研究科	0	2	0	7	0		8
経済学部	2	10	2	17	0		31
東アジア研究科	0	9	0	0	0		9
教育学部	0	9	6	3	0		18
農学部	0	2	2	0	0		4
工学部	11	1	0	9	0		22
理工学研究科	0	11	1	1	0		13
未記入・不明					2	2	4
合計	13	44	11	37	2	2	109

表4 Y大学2008年度後期学部別受験者数

	学部生	大学院生	研究生	特別聴講生	研究者	合計
人文科学研究科	0	1	4	9	1	15
経済学部	1	8	4	16	2	31
東アジア研究科	0	1	0	0	0	1
教育学部	0	5	8	7	1	21
理学部	1	1	1	0	0	3
農学部	0	3	1	0	0	4
工学部	4	0	3	10	1	18
理工学研究科	0	4	0	0	0	4
医学系研究科	0	0	1	0	1	2
合計	6	23	22	42	6	99

表5 Y大学2008年度身分別受験者数

	前期	後期	合計
学部生	13	6	19
大学院生	44	23	67
研究生	11	22	33
特別聴講生	37	42	79
研究者	2	6	8
未記入・不明	2	0	2
合計	109	99	208

表6 Y大学2008年度国別受験者数

前期		後期	
国	受験者数	国	受験者数
中国	57	中国	60
韓国	25	韓国	18
マレーシア	9	イギリス	4
インドネシア	4	インドネシア	3
アメリカ	3	アメリカ	2
オーストラリア	1	ラオス	2
ドイツ	1	ベトナム	2
ウクライナ	1	カナダ	1
ベトナム	2	ドイツ	1
タイ	4	ウクライナ	1
ニカラグア	2	ペルー	1
未記入	4	ニカラグア	1
		ブラジル	3
		ミャンマー	1
		マレーシア	1
合計	109		99

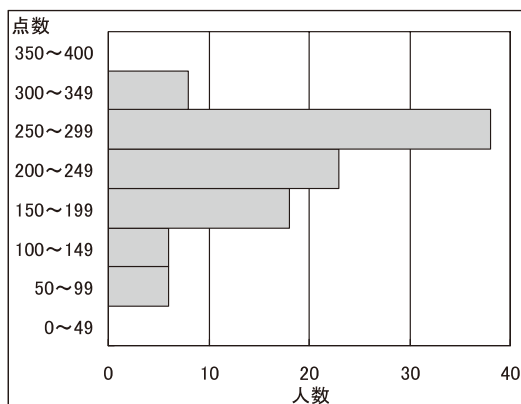


図1 総合

表7 Y大学受験者数(前期+後期)の推移

年度	2005	2006	2007	2008	合計
受験者数	197	174	231	208	810

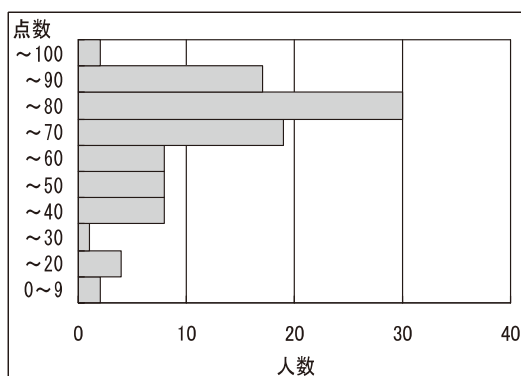


図2 聴解

### 5.3.5.2 得点分布

能力推定法と得点換算法を2008年後期から若干変更したため、それまでの得点との比較はできないので、2008年後期のY大学の受験者99人の成績に限ってみる。

聴解の平均点が一番高く、次に読解、そして文法、語彙の順番になっている。

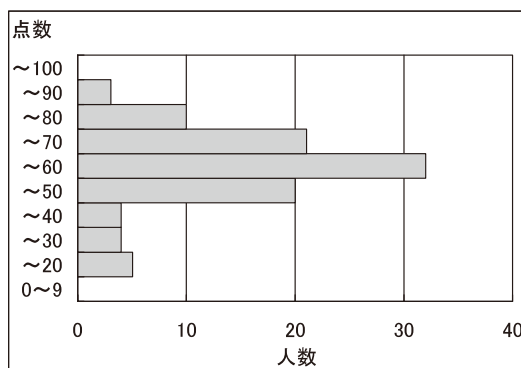


図3 語彙

表8 基礎統計量 (n=99)

	聴解	語彙	文法	読解	総合
平均	63	53	54	57	227
標準偏差	20	16	17	19	63
最高点	92	87	88	87	341
最低点	4	10	11	13	64

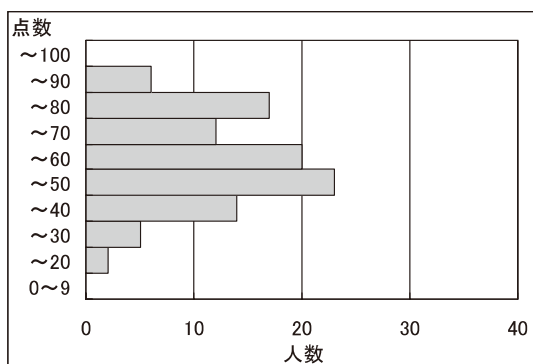


図4 文法

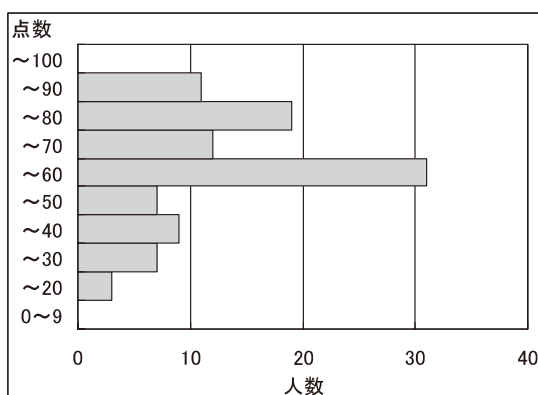


図5 読解

これまでの数回分のデータを見ると、例えば、日本滞在歴が長くなるにつれて、聴解の能力が高まる傾向があるようである。また、前期と後期で成績の伸びの著しいグループなどがあるようだが、このような時系列変化は今後さらにデータの蓄積を待って分析を行いたい。

## 6 さいごに

J-CAT日本語テストは、WEB上でいつでも世界中どこからでも受験することのできるテストであることから、表7で示したように、Y大学では過去4年間で延べ800名超の留学生が受験し、その成績を日本語授業のクラス

分けや奨学金受給者選考に役立ててきた。2008年度からは、本格的に他機関での運用検証や公開を始め、国内をはじめロシアなど海外からのアクセスも始まっている。今後、他機関からの受験者が増加していくものと期待される。

J-CATは、2006年から本格的にアイテム作成を始めた。そして、アイテムライター募集、アイテムの作成、世界規模でのプレテストの実施という膨大な労力のかかる作業を行ってきた。J-CATは、項目応答理論に基づいたアダプティブテストであるため、信頼性が高いテストと言えるが、アイテムの作成、アイテムの精選に労力と時間がかかることは否めない。それでも精度の高いアイテムプール構築のためのアイテム作成・プレテストを今後も継続していきたい。

(山口大学留学生センター 准教授)

(東京外国語大学留学生日本語教育センター 教授)

(常磐大学国際学部 准教授)

(東邦大学理学部 准教授)

(山口大学・留学生センター 講師)

(島根大学外国語教育センター 准教授)

(立命館アジア太平洋大学言語インスティテュート 准教授)

(東京工業大学 大学院生)

### 【参考文献】

今井新悟, 2006, 「コンピュータを使った適応型日本語絶対評価システム: J-CAT 2005 Version」『大学教育』第3号, 133-143.

赤木彌生・中園博美・今井新悟, 2009, 「コンピュータを使った日本語試験アイテムの開発-J-CAT文字・語彙アイテムを例に-」日本語教育学会第10回中国地区研究集會予稿集, 24-33.

J-CAT (Japanese computerized adaptive test)  
<http://www.j-cat.org/>

日本語能力試験 <http://www.jlpt.jp>

McNamara, Tim, 2000, *Language Testing*, Oxford

University Press (伊東祐郎・三枝令子・島田めぐみ・野口裕之監訳, 2004, 『言語テストニング概論』スリーエーネットワーク  
 中村洋一, 2006, 「言語テストにおけるテスト理論とデータ分析」『世界の言語テスト』くろしお出版, 15-24.

中村洋一 (大友賢二監修), 2002, 『テストで言語能力は測れるか—言語テストデータ分析入門』桐原書店.  
 大友賢二, 1996, 『項目応答理論入門』大修館書店

【参考資料】

(1) J-CAT登録画面



**J-CAT** Japanese Computerized Adaptive Test

日本語 [にほんご](#) English [中文](#)

---

**J-CAT 日本語テスト へようこそ**



[登録の流れ \(個人受験\) ▶](#)  
[登録の流れ \(団体受験\) ▶](#)

---

**J-CAT 日本語テスト とは**

J-CAT 日本語テストはコンピュータが項目応答理論に基づき、各受験者の能力に応じた問題を出题する日本語能力の絶対評価テストです。

---

**テストの構成**

J-CATには聴解、語彙、文法、読解の4つのセクションがあります。出題される問題は受験者の解答に応じて変わります。出題数は、各セクション 15問 ~ 20問かそれ以上になります。受験時間は 45分 ~ 90分ぐらいになります。

---

**システム要件**

J-CATを受験するには、以下のシステム要件を満たす必要があります。

- Web ブラウザ: Internet Explorer 6.0 以上または Mozilla Firefox 1.5 以上
- モニタ: 1024×768px 以上の解像度
- 音声を再生できること
- 最新版の Adobe Flash Player
- Javascript を有効にすること



**個人で受験する**

ログインするには登録が必要です。

**登録**

**ログイン**

**団体で受験する**

団体受験用のパスワードが必要です。

**登録**

**J-CAT のレベル**

[J-CAT レベルの詳細](#)

**管理**

[管理用ページ](#)

**お問い合わせ**

[j-cat@yamaguchi-u.ac.jp](mailto:j-cat@yamaguchi-u.ac.jp)

**リンク**

[J-CAT Project Webサイト](#)

---

(C) J-CAT Project.  
 Powered by CAT-Test 2.4. Icons: [famfamfam.com](http://famfamfam.com)

## (2) 成績表

成績証 Score report		J-CAT® Japanese Language Test <a href="http://www.j-cat.org/">http://www.j-cat.org/</a>	
あなたの J-CAT スコア は以下の通りです Your J-CAT scores are as follows.			
名前 / name	Takagishi, Hiromasa	聴解 / Listening	16
生年月日 / birth date	1988/08/21	語彙 / Vocabulary	34
所属 / affiliation	Yamaguchi University	文法 / Grammar	46
登録日 / register date	2008/04/08	読解 / Reading	20
受験日 / exam date	2008/04/08	合計 / Total	116
(watermark: 7c6b9f15cae0732bfbacc8c9da1f12b04) J-CAT Project.			

Interpretation of J-CAT score		
J-CAT score	Proficiency Level	JLPT 日本語能力試験
-100	Basic 初級	Level 4
100-150	Pre-Intermediate 中級前半	Level 3
150-200	Intermediate 中級	
200-250	Intermediate-High 中級後半	Level 2
250-300	Pre-Advanced 上級前半	Level 1
300-350	Advanced 上級	
350-	Near Native 日本語母語相当	

- "Novice" learners can use fixed expressions such as greetings.
- "Basic" learners can exchange basic ideas.
- "Intermediate" learners can manage daily communication.
- "Advanced" learners can manage academic and professional communication.

J-CAT Project.