

研究テーマ： 日本中国語検定試験の信頼性と妥当性に関する研究	
研究代表者： 人間文化学部国際文化学科 教授 侯 仁鋒	連絡先： hourf@pu-hiroshima.ac.jp
共同研究者： 教授 丸山 浩明	
<p><b>【研究概要】</b></p> <p>本研究は、日本で中国語学習者の能力を測る試験として実施されている中国語検定試験（以下中検と省略）の信頼性と妥当性を明らかにすることを目的とする。本研究は、より公平で正確な尺度としての出題の質をアップし、より正しく中国語教育の方向性に導くことにつながり、また、言語テスト理論に基づくその実践にも貢献できるものと考えられる。</p>	

**【研究内容・成果】**

## 1. 研究の学術的背景

### 1. 1 ニーズが高い

グローバル化の時代の到来にしたがって、外国語の教育はますます繁盛になる一方であると同時に、言語テストもいろいろな目的に応じて多種多様に実施されている。中国語教育も日本ではブームになっており、中検をはじめ、実用中国語技能検定試験（C.TEST）、漢語水平試験（HSK）などが実施されている。このうち、中検は、英語検定と同様に実施実績が長く、受験者が多いということで、社会に認められている存在である。

### 1. 2 研究が必要かつ可能

言語テストもほかのテストと同じように、まず尺度として公平で正確なものでなければならぬのである。公平で正確なものにするために、そのテストを多角的視点から常に分析し研究する必要がある。よく知られている TOEFL も TOEIC も、その信頼性、妥当性、項目の成否について研究がさかんになされている。また、世界各地で実施されている日本語能力試験（JLPT）も、内部（出題者）による報告や外部による研究がよく見られる。このように言語テストについて、その正確性を検討する理論も方法もあり、可能になっている。これによって、テストの公平で正確な尺度が保証されることになる。

良いテストとされる条件の一つは、そのテストの透明性である。即ち、試験問題も試験実施後のデータも公開することが求められる。中検の問題は公開されているし、試験後のデータも一部公開されているので、研究が可能である。

大型な試験も、特にマークシートの回答式の問題について、コンピュータとソフトの利用により、仔細なデータの統計もできて、ゆえに各項目を細かく検討することが可能であり、必要不可欠な時代になっている。

### 1. 3 中検に関する先行研究は皆無

中検に関する研究は、インターネットや「GeNii 学術コンテンツ・ポータル」にかけて調べてみたが、個別の項目についての研究がわずかにあるだけで、全面的かつ系統的な研究は見つからず、研究がほとんどなされていないのが現状である。

## 2. 研究対象と内容

---

## 2. 1 研究対象

中検は、準4級、4級、3級、2級、準1級、1級の六つのレベルに分けて試験が行われる。本研究は今回3級を研究対象とする。中国語学習の最も基礎的な内容を測定する級であるし、また、3級を受験する受験者が最も多いからである。

## 2. 2 研究内容

本研究は今回、平成19(2007)年度、20(2008)年度2年度計6回の試験を研究内容とする。2年度6回の分量は、ある程度試験の内容及びその代表性(妥当性)、安定性や一貫性(信頼性)などが観察されると考えられるからである。

## 3. 研究成果

### 3. 1 試析日本漢語検定考試3級聴力試題(日本中国語検定試験3級のリスニングについての考察)『中国語教育』第9号(2011. 3) P89~100

テストの内部一致性と聴解、聴解試験のあり方の視点から、中国語検定試験3級のリスニングを考察してみた。

日本中国語検定協会が公表した過去2年分(07、08)の試験結果データによると、リスニングと筆記の平均点は、64回(0.2点)と62回(2.5点)、61回(3.7点)、63回(5.5点)では、差が小さくテストの内部一致性において信頼性が高いと見られるが、65回(10.3点)、66回、(9.3点)では、差が大きくて信頼性が低いと言わざるをえない。

聴解と聴解試験のあり方から検討を加えた結果、リスニングは、自然な中国語、適切な設問、設問を提示してから課題文を聞く方法などが、信頼性を維持する要素として多く見られる。一方、①テストとしてのテキストが適当ではない。②課題文が長すぎる。③選択肢が長すぎる。④設問と選択肢が合わないなどの問題点が散見される。それがテストの信頼性を低下させる素因になっているのではないかと思われる。

### 3. 2 試析日本漢語検定考試3級筆記試題(日本中国語検定試験3級の筆記問題についての考察)『中国語教育』第10号(2012. 3) P155~175

本論文は、2007年・2008年に実施された中国語検定試験3級の筆記試験計6回分を対象として、何をどのように測っているかの観点から検討したものである。試験後に公表された平均点や合格率のデータ及び本学の学生が行った結果を参考に、言語テストの理論に基づいて特徴を3点にまとめて記す。

①知識と技能の両方を測ろうとする教育観、試験観を反映している。

②発音(ピンインと声調)、文法(空欄補充と語順整序)、読解、和文中訳の出題により、言語習得の達成度を全面的に測定し、試験の妥当性を保証しようとしている。

③客観問題を主にして、主観課題も加えることにより、信頼性を確保しながら妥当性を高めようとしている。

しかし、以下に挙げる点からはさらに工夫が必要であると判断される。

①問題数が少ない。②項目間に重複が見られる。③形式に厳格さが欠ける部分がある。④読解の空欄補充がセンテンスレベルにとどまる。⑤実施回ごとの難易度の差や変化が大きい。