

## 外国語としてのロシア語の検定試験の 口頭発表力テストについて

トルストグーゾフ アレキサンダー<sup>※</sup>

### はじめに

現在の日本の大学におけるロシア語教育の状況では、第二外国語科目としての『ロシア語』科目を活性化させることが重要である。ひとつの方法としては、教育課程へロシア語検定試験を導入して、カリキュラムと授業の内容を見直すことを挙げることができる。その場合、教育のために具体的な目標が定められ、知識と技能の客観的な測定措置が与えられると考えている。

しかし、各テストが特定の目的のために作成され、特定の状況の下で効率的に利用される。したがって、テストを教育課程に導入する前に、テストの利点と欠点、テストの適用性が研究の対処にならなければならない。

ロシア語能力検定試験(ТРКИ—TORFL)は、ロシア教育省によって開発され、ロシアの大学や大学院への外国人入学希望者に義務化されている。合格者は、資格認定書が与えられる。ロシア語検定試験の多レベルシステムはALTEというヨーロッパの言語試験官の制度に含まれている。

ロシア語能力検定試験は、6段階の試験があり、各レベルの試験は、いずれも文法・語彙力、読解力、作文力、聴解力、口頭発表力の5つの知識・技能領域から構成されている。初級(入門)レベル(Русский язык. Элементарный уровень - ТЭУ)と基礎レベル(Русский язык. Базовый уровень - ТБУ)は、資格認定のためのものではなく、ロシア語学習者の日常会話の習熟度や、基礎的なロシア語理解力を測定するためのものとして設定された。

日本において、ロシア語検定試験の、教育課

程への適用の試みは行われている。ロシア語検定試験は、専攻語としてロシア語の教育を行っている大阪大学外国語学部と札幌大学外国語学部では、すでに検定試験に準拠する試験として導入されている。現在の形のテストの非専攻ロシア語の教育への適用の可能性は認められている。臼山利信は、「中等教育や大学の教養教育におけるロシア語能力検定試験の資格習得目標としては、前者〔東京ロシア語学院が主催する検定試験 — T.A.〕の検定試験では4級ないし3級、後者〔ロシア教育省認定 — T.A.〕検定試験では入門レベルないし基礎レベルの合格とうことになると思います<sup>1)</sup>」と書いている。しかし、内容の高いレベルのため、テストの適用は難しい。北岡千夏は指摘している。「ロシア語能力検定試験で初級(入門レベル)と定められている文法項目すらすべてを1年間の授業で消化することができない<sup>2)</sup>」。武藤研介の意見も同じ。「ТРКИはバランスのよいロシア語能力の定着度を確認するために理想的なテスト形式であろう。しかし、ロシア語能力を測る上では理想的ではあるが、日本の大学の専門外「ロシア語」科目の受講学生が合格するには、ТРКИはハードルが高い<sup>3)</sup>」である。

たしかに、外国語としてのロシア語テストは、世界で一番長い言語テストのひとつであるという印象を与える。たとえば、ロシア語初級テストは、二日間、3時間35分(99年版)で行われる。ロシア語をできるだけ徹底的にテストするという、テスト作成者の狙いは理解できる。しかし、重要な意思決定(入学合否判定、クラス分けなど)を行っていない初級レベルの試験は、初級(入門)レベルの教育での、動機付けの役割を果

※青森公立大学准教授

たすべきか、ロシア語の勉強に興味を起こさせるものとして考えたほうが良いのではないか。テストのあまりの難しさは、可能な受験者の数を下げる。ロシア国外の教育者（在日本も含めて）の提案：現在の初級（入門）レベルの下レベルの、より簡単なテストの導入の案は、ロシア教育省に認められていない。

いづれにせよ、全体のテストあるいはそのある要素を教育課程へ導入することが検討に値する。本論文は、理論的な枠としてスピーキング能力の性格、言語テストの目的と種類、テストの評価について簡単に触れた後に、初級レベルのロシア語検定試験のスピーキング下位テストの妥当性と信頼性、波及効果、実用性を議論する。

## 1. スピーキング能力とコミュニケーション能力

### (1) 構成概念。

言語テストは言語能力の測定を目的とするが、言語能力は目に見えない。したがって、より正確には、何らかの言語理論に基づいて構成概念を構築し、それをテストという形に具体化あるいは操作化し、受験者の言語能力を引き出し、そしてテスト得点から測定対象の言語能力を推測することを目的とする。「言語能力」、「コミュニケーション能力」、「スピーキング能力」、「文法力」などは、すべて構成概念の例である。

構成概念は人間によって作られる特性である。すなわち、人間行動のパターンを容易にするために役立つものとして、概念化されるものである。テストは、何らかな構成概念に関しての受験者の位置を数値的に表現するための道具と言える。教育上のテストでは構成概念はシラバスという形で具体化される（静哲人、竹内理、吉澤清美共編著、2002）。

### (2) スピーキング能力。

言語を知っているということが何を意味するのかに関する理論的な枠組みは、言語教育や言語学習にとって非常に重要な課題である。言語の4技能の中で「話すこと」はそれを代表するものである。第二外国語の教育においても、興味をもたれている諸問題の研究にあたっても、習

得の過程や程度を測る尺度としてスピーキングが用いられている。

第二言語教育理論の変遷はスピーキングの指導法に様々な影響を与えている。伝統的な文法訳読式はスピーキングを全く無視したが、オーディオリンガル・メソッドはその反動として、モデルとなる文型を繰り返して発話することに焦点を置いた。その後、単なる発話ではなく自発的であることが大切だという認識が持たれ、教師や学習者間の意味のある相互交流の重要性が挙げられるようになった。その時から、スピーキング指導の主眼は、発話者の意思に基づいた発話を引き出すことに置かれるようになった。しかし、スピーキングを独立したひとつの技能としてではなく、もっと幅広い要因を含んだ言語の熟達度という視点から捉え、教師中心の教育から学習者中心の教育へと視点が移ったことも相まって、生産的スキルとしてのスピーキングが注目を浴びるようになった。このように、スピーキングの研究は第二外国語研究で扱うほとんどの課題と密接な関係を持っている（小池等、2004）。

### (3) コミュニケーション能力。

話すことで意思伝達するためには、聞き手に理解可能な語彙、発音、文法構造などの言語学的要因が欠かせない。しかし、これだけでは効果的なコミュニケーションは望めない。言語構造以外の言語要素が含まれているからである。

コミュニケーション能力を分析する過程において、研究者たちは、自分のアプローチに応じて、異なる数の内部構成の要素を区別し、同じ要素の規定に使う学術的な範疇も異なるケースも少なくない。しかし、一番重要な構成要素の決定は、カリキュラムの編成、テストの作成、テストの結果の評価に欠かせないものである。

コミュニケーション能力の具体的な構成要素に関しては、研究者間で意見が異なる。代表的な理論的な枠組みは、以下のものである。1. 文法能力（grammatical competence）— 語彙・発音・文法構造などの言語規則を習熟している程度のこと。2. 社会言語能力（sociolinguistic competence）— 社会的規則や約束事を把握した、

様々な機能で言語使用ができる程度のこと。3. 談話能力 (discourse competence) — 前後関係、場面、環境などの状況を把握して的確に伝達できる程度のこと。4. 方略的能力 (strategic competence) — 言語能力の不足や伝達上の支障を補い、より円滑にコミュニケーションをすすめられる程度のことである。しかし、これらの4つの構成要素がどのように相互作用して働いているのか、どのように働いてパフォーマンスへと繋がるのか研究者の間で説明が異なる。

## II. テストの目的と種類

言語テストの目的は、目に見えない言語能力の測定であるので抽象的な言語能力(構成概念)を引き出すことになる。しかし、通常その結果をもとにして何らかの意思決定、すなわち評価を行う。評価の目的は多岐にわたるが、基本は、熟達度テストと到達度テストである。熟達度テストは、テストを実施する時点における、学習者の相対的な言語知識、能力、スキルを測定するためのテストである。特定のコースやカリキュラムに基づかないテストである。熟達度テストは、現実生活で必要となる言語行動に照らして受験者がどの程度に熟達しているかを測定しようとする。このテストは入学者選抜の情報収集などに相応しい。標準化された熟達度テストとしては、TOEFL、IELTS、Cambridge examinations (CAE、CPE、FCE、PETなど)が世界的によく知られている。日本の英語検定なども熟達度テストに分類できる。

到達度テストは、一定の期間内に、コース、教科書、プログラムの学習科目を学習者がどれだけ習得したかを測定するためのテストである。学期末試験や、学年末試験は到達度テストと考えられる。到達度テストは受験者全員が与えられたレベルにまで到達することを前提としているので、限られた定員の選抜にむかない。熟達度テストと対比して使われるが、同一のテストでも目的によって、到達度テストにも、熟達度テストにもなる。外国語としてのロシア語検定試験は到達度テストである。

テストの目的は得点の解釈にも関係する。集

団基準準拠テスト(norm-referenced test)は、相対的な解釈を行うためのテストである。受験者のパフォーマンスが、他の受験者のそれとの相対的な優劣によって評価されるようなテストである。テストは、目標受験者集団の能力上の個人差をできるかぎり弁別するように困難度が設定される。

目標基準準拠テスト(criterion-referenced test)は、絶対的な解釈を行うためのテストである。あらかじめ明確に定義された知識や技能のレベルの基準(criterion)が設定しており、そのレベルに受験者が達しているかいないか、あるいは、複数のレベルのうちどのレベルに達しているかを測るテストである。テスト結果は数値ではなく、受験者が「できること」と「できないこと」に関する記述的な説明になることが多い。現実には、目標基準の設定が恣意的になりがちであるなど、問題も多い(静哲人、竹内理、吉澤清美共編著、2002)。外国語としてのロシア語検定試験は目標基準準拠テストである。

スピーキング能力が代表するコミュニケーション能力を測定する言語テストはコミュニケーション能力テストである。その第一の特徴として、コミュニケーション能力テストはパフォーマンステストである。それは、受験者が目標言語でコミュニケーションをしている際に評価が下されるテストを意味し、コミュニケーションの相手に情報を伝達する様式や相手から情報を受容し理解する受容の様式のどちらか一方あるいは両者が含まれる。第二に、コミュニケーション能力テストでは実際のコミュニケーションを測る状況で受験者が担うと予測される役割とタスクが重要視される。言語能力を言語使用能力と捉える考え方がより広く行われるようになったので、最近のトレンドとしては、テスト研究がテスト得点を統計的に分析する心理測定の枠を超えたということにある。

## III. テストの評価

研究者によっては、様々な言語能力の概念の構成要素が挙げられるので、テスト評価へのアプローチも様々である。筆者は、Bachman and

Palmerの枠組みを使用する。彼らが提案するように、言語テストを検討する際に考慮すべきもっとも重要なことは、その「有用性」である（バックマンL.F.、パーマーA.S.、2000）。そして、これは6つのテストの特質という観点からの定義である。その6つは、信頼性、構成概念妥当性、真正性、相互性、影響そして実用性である。これら6つのテストの特質はすべてテストの有用性に貢献しているので、それぞれを個々に独立して評価することはできない。さらに、これらの異なった特質の相対的な重要性はそれぞれのテスト状況によって異なるので、テストの有用性の評価はある特定の状況にのみ限定して行われるべきである。むしろ、テストの目的、目標言語使用領域やテスト受験者の特性、そして測定しないと思っている構成要素の定義方法などを考えて、適切なバランスが達成できるようにする必要がある。

「信頼性」は測定の一貫性と定義することができる。一貫性を欠く場合はそれはテスト得点の差異であり、測定したいと思っている構成概念以外の要因に起因しているものである。

「構成概念妥当性」は、テスト得点をもとに行う解釈の意味や適切性に関係している。テストの妥当性の検証については、その解釈は正当である証拠を提供することもある。

「真正性」は、ある言語テスト課題の特性と目標言語使用課題の特性の一致の度合いと定義される。

「相互性」とは、テスト受験者がテスト課題を達成する際に自己の言語能力（言語知識とメタ認知的方略）、話題の知識そして情意スキーマをどの程度取り込み、どのような取り入れ方をしているのかという度合いと形式の関数といえる。

「影響」は、テスト利用が社会、教育組織、そして個人に与える様々な影響という観点から広く定義することができる。

「実用性」は、テスト得点と密着した使用と関係する他の5つの特質と違って、テストの設計、開発、使用で要求される資源とこれらの活動のために利用できる資源の関係と定義することができる。

#### IV. 初級レベルロシア語検定試験の妥当性と信頼性

テストを評価する際に、もっとも重要な特質は妥当性と信頼性である。これは、測定しようとしている構成概念を、当該テストが実際に測定している程度である。あるいは、テスト得点の解釈が正当かどうかの程度である。テストが妥当性を持つためには、まず必要条件として得点に信頼性がなければならない。しかし、信頼性の高いテスト得点であっても、妥当性を欠く場合もある。妥当性を論じる際には、つねに、どのような構成概念のテストとして論じているのかを明らかにする必要がある。その意味で、かならず、「何のためのテストの妥当性か」の問いかけが必要である。伝統的な妥当性は、内容的妥当性、併存的妥当性、予測的妥当性、構成概念妥当性、などの、いくつかの異なる切り口から論じられてきた。しかし現在では、構成概念妥当性こそが妥当性のすべてであって、他のすべては下位概念として含まれる、という立場が一般的である（静哲人、竹内理、吉澤清美共編著、2002）。

テストの項目が測定する能力(構成概念)、テストの目的、項目の形式、想定する受験者層、時間配分、採点基準、採点方法などを細かく規定する。大規模のテストの場合にはそもそも細目自体が作られているのかなどが調査対象となる（チャールズ・オルダーソンなど、2010）。

外国語としてのロシア語の検定試験が測定する言語能力は『基準』と『プログラム』で記載されている。初級レベルの試験の具体的な内容は、1999年に出版された『外国語としてのロシア語の国家教育基準。初級レベル。案（Государственный образовательный стандарт по русскому языку как иностранному. Элементарный уровень. Проект）』（以下、99年版）によって決められる。ここでは、コミュニケーション活動領域、テーマの簡単な記載に続いて初級レベルで要求される言語構造能力の細かい記述がある。2001年に『外国語としてのロシア語の国家教育基準。初級レベル。案。改訂版

(Государственный образовательный стандарт по русскому языку как иностранному. Элементарный уровень. Проект. Издание второе, дополненное и расширенное)』(以下、01年版)が出た。同じ2001年に『外国語としてのロシア語の教育プログラム。初級レベル。基礎レベル。1資格レベル(Образовательная программа по русскому языку. Элементарный уровень. Базовый уровень. Первый сертификационный уровень)』が発表された。この年に出た別の『1資格レベルのプログラム』についての林田理恵の説明が、初級レベルの『プログラム』にも当てはまる：「プログラムは包括的なCERF<sup>6</sup>能力記述分のパラメータ・カテゴリーをロシア語学習に関して具体化したものである。そこでは…能力尺度に関する抽象的な記述のみならず、活動領域・テーマ・課題・場面ごとの具体的なコミュニケーション活動内容、さらにそれらのコミュニケーション活動に必要な言語表現形式、言語構造知識の全体が記述されている<sup>7)</sup>」。

通常、『プログラム』は『基準』の内容を説明する。ところで、2001年の『プログラム』は、言語構造能力を初級、基礎、1レベルのレベルごとに記述するが、コミュニケーション活動領域・テーマをレベルごとに分けず、まとまった形で記述する。すなわち、1レベルのコミュニケーション活動領域・テーマになる。したがって、初級レベルのものはその中に含まれているので、分けることはできない。そのため、理想的に考えれば、『基準』と『プログラム』の内容の対比の必要があるが、それはできない。そこで、『基準』の99年版と01年版の比較を行う。

初級レベル『基準』の99年版で必要とされるロシア語学習時間は、100から120学習時間(1学習時間=45分)、最低必要語彙数は760語である。01年版では学習時間は変わらないが、最低必要語彙数は780まで増えている。実際に、ロシア語の単語は、格などの変化をするので、学習者が習得しなければならない語彙の数はかなり増えている。

99年版の『基準』で必要とされる技能は、聴解力、解読力、作文力、口頭発表力、その後には大きいセクションの言語構造能力がある。テストに高い妥当性を与えるために重要なことは、テストの対象を明確にすることである。この意味で、外国語としてのロシア語のテストは、教育過程で現実的に教えらるすべての技能を測定するので、かなり高い外面妥当性(face validity)を持っている。テストは、教育標準の文法の80-85%の内容を含む<sup>8)</sup>ので、テストの内容妥当性も高い。技能別にいろいろなタイプのタスク(選択肢、穴埋式、フリーレスポンス)を含むので、受験者に、すべての言語知識や生産的、受容的なスキルを表すために十分なチャンスが与えられる。

99年版の『基準』は以下のロシアの日常生活に必要な初歩的なコミュニケーション能力を記述する。

#### コミュニケーション言語運用能力の内容<sup>9)</sup>。

コミュニケーション活動。場面とテーマ。

##### (1) 機能別コミュニケーション課題

- 呼びかけ、紹介、自己紹介、挨拶、別れ、問いかけ、感謝、謝罪、感謝と謝罪に対して答え、問い返し。
- 事実・出来事、人、物、人と物の存在・不在、物の質と所属、行動の時間と場所、理由についての質問、情報の伝達。
- 希望、願い、提案、招待、同意・意義、拒否。
- 自分の態度：人、物、事実、行動の評価。

##### (2) 状況ごとのコミュニケーション活動内容：

店、売店、レジ；郵便局；銀行、外貨両替所；レストラン、喫茶店、カフェ、食堂；図書館；授業；町、交通機関；病院、医者、元、薬局。

##### (3) コミュニケーション活動のテーマ：

自分について；履歴：幼年時代、教育、仕事、関心；私の友(知り合い、家族の一員)；家族；私の労働日；暇、休息、趣味。

このような要求が、外国人の実際のニーズをどの程度に満たしているかの研究はまだ行っていない。しかし、現段階でもある程度の修正は

必要があると考えている。与えられたコミュニケーション活動のテーマを見ると、まず、ロシアで教育を受ける学習者のために必要なものが含まれるという印象を受ける。

活動場所のリストには、ロシアに到着したばかりの外国人のための空港と国境にある鉄道ステーションが欠けている。また、ロシアにきた学習者の実際の生活の場であるホテルも提示されていない。これは、検定試験が、ロシアの大学へレギュラコースの勉強に来る学習者を主に想定していることを反映する。ロシア語のゼロのレベルを持ってロシアに来て、レギュラコースに入る前に、1年間準備学部でロシア語を習ってからフルコースを受ける学習者にとっては、このテーマは難しいかもしれない。ホテル、鉄道の駅、空港は1レベルのリストに含まれている。けれども、初級のレベルでももっとも簡単なコミュニケーション活動が可能であり、そして実際の必要がある。他方、現代では、リストに入っている郵便局でのコミュニケーションはそれほど重要ではない。より大事なものは、インターネットでコミュニケーションを行うことである。

外国人は、このレベルに必要なテーマの枠において以下のコミュニケーション活動を実行する。自分についての話、履歴、幼年時代、教育、仕事、関心、自分の友（知り合い、家族の一員）、家族、自分の労働日、暇、休憩、趣味。このすべてのテーマは大事であるが、このレベルでバイオグラフィについての話は絶対必要とさえ言えないであろう。全体として、コミュニケーション活動、社会的場面とコミュニケーションテーマのリストは、初級レベルが要求するコミュニケーション要素を十分に満たしている。

99年版の『基準』では、口頭発表力テストに合格するに必要とされる言語運用能力が以下のようになっている。

#### A. モノログ

- 所与のテーマと質問を手がかりにモノログ形式の発話。発話量は最低7文。テキスト量は：150-180単語。未知の言葉の数は：なし。
- 与えられたテキストについて内容要約を行い、

自分の意見を展開する。

テキストのタイプは：初歩レベルに相当する語彙・文法によって特別に作成されたシステムチェックなテキスト。

#### B. ダイアログ

- 与えられた状況に応じて相手の発話を理解する；
- ダイアログに参加して適切な応答を行う；
- 与えられた状況に応じて、適切なダイアログを導入する。

学習者の発話は、言語待遇の一般的な、社会的に要求されているノルマを含む現代ロシア語の規範に応じて形作られなければならない。

01年版の『基準』の口頭発表力に関しては、モノログの課題の量は変わらない：7文である、しかし受験者が理解しなければならないテキストのボリュームは150-180単語から150-200単語まで増やし、そして1%の未習熟の単語のノルマが導入された。ちなみに、未知語は他のテストにも含まれている。

教科書で習っている語彙が少ない上に、未習語は出さないとすると、本当の限られた単語しか使えなくなり、テキストのオーセンティシティ（真正性）は大きく下がってしまうことになる。現実の生活では、外国語の使用において、未知語がないという状況はあまり想定できないであろう。このため、コミュニケーション・テストでは、ある程度の比率の未知語は容認される傾向にある（根岸、2007）。

全体的に、初級レベルの『基準』は、コミュニケーション言語運用能力と言語構造能力の内容を十分に表している。しかし、6つレベルの外国語としてのロシア語検定試験の『基準』では、コミュニケーション言語運用能力と言語構造能力が詳しく記述されているが、他の能力（知識）の内容、それらのレベルによる変化は、まだ規定されていない。もちろん、テストでこれらの能力が測定されるが、その範囲とボリュームについてははっきりした規定がない。また、これらの能力の徹底的な分析は必要である。このような分析により、すべての能力の細かい記述は、テストシステム全体の妥当性のレベルを高めて、

テストの今後の発展に良い影響を与えるにちがいない。

国家教育基準は、外国語としてのロシア語の学習者のレベルを客観的に測定し、評価することのための基礎を与える。同時に、初級レベル専用のプログラムの作成、基準の内容の再検討と一定の修正が必要である。

その際に、自然的に、多面的なテスト制度のいろいろな要素にも修正が行われると考えられる。これは、テストの対象、テストタスクの形式、課題の内容、評価の規準、評価のパラメータなどである。レベルごとに言語運用能力に対して要求が変わるなら、タスク形式なども変わる。

#### 口頭発表力テスト

「口頭発表力」テストは、外国語としてのロシア検定試験の中で、与えられる点数で一番重要なテストである。総合点数の530点の中で、口頭発表力 — 130点、読解力 — 120点、語彙・文法力 — 100点、聴解力 — 100点、作文力 — 80点を占める。

「口頭発表力」テストは、公式的な資格を持つ試験官との、直接、1対1の面接の方式で行われる。テストは終わってから、試験官によって採点が行われる。一般に、スピーキングテストを実施しつつ、信頼性の高い採点を行うことは難しいとされる。それは、多くの場合、テストの実施に気をとられて、採点に十分な注意が向けられなくなってしまうからである（根岸、2007）。

この理由で、ロシア語検定テスト管理者は、採点者に対する信頼性を保つために、大きな注意を払う。録音された受験者の回答と採点者の評価はテスト本部に送られ、もう一人の試験官によってチェックされる。二人の採点者の評価に隔たりがあるなら、もう一人の採点者でチェックが行われる。他のテストの点数も本部で検査される。

#### テストタスク

テスト対象を直接測ることは重要である。テストは間接的になればなるほど、妥当性は下がる。スピーキング能力は話させてみてはじめてわかる。外国語としてのロシア語試験は、直接的な会話の形で行われるので、高い妥当性を持つ。

口頭発表力のタスク形式を作るときに考慮に入れなければならないのは、1) スピーキングはモノログ、ダイアログ、ポリログの形を持つこと；2) スピーキングは準備された性格と準備されていない性格を持つことなどである。

初級レベルの口頭発表力テストは、はっきりとモノログの部とダイアログの部に分かれている。99年版の「口頭発表力」テストで試験官はダイアログの相手と、受験者のモノログの発起者になる。テストの時間は45分間である。

パート1の設問1（問1-5）では、試験官による日常生活に関係する質問に対して、受験者が回答する。すなわち、与えられた状況に応じて相手の発話を理解して、ダイアログに参加して適切な応答を行う能力のテストである。すべての回答はフリーレスポンスの形式である。発話時間は4分間である。

1. Скажите, пожалуйста, сколько сейчас времени?  
(どうぞ、教えてください、今は何時ですか)
  2. Вы будете пить чай, кофе или сок?  
(あなたは何になさいますか、  
紅茶、コーヒそれともジュースですか)
  3. Где вы изучали русский язык?  
(あなたはどこでロシア語を勉強しましたか)
  4. Скажите, пожалуйста, где здесь станция метро?  
(どうぞ、教えてください、  
この近くのどこに地下鉄の駅がありますか)
  5. Почему вы изучаете русский язык?  
(なぜあなたはロシア語を勉強しますか)
- 設問2（問6-10）では、試験官によって与えられた状況に応じて受験者はダイアログを導入する。発話時間は5分間である。
6. Вы забыли, куда положили газету. Спросите у друга.  
(あなたは新聞を置いたところを忘れてしまいました。友達に聞いてください)
  7. Спросите друга о погоде завтра.  
(友達に明日の天気について質問をしてください)

8. В магазине вы хотите выбрать хорошую книгу (или сувенир).  
(店であなたは良い本

{それとお土産} を選びたい)

9. Вы хотите пригласить в гости подругу. Что вы ей скажете?

(あなたは女友達を客として招待したい。

彼女に何をいわなければならないか)

10. Расскажите ей, где вы живете.  
(彼女に、あなたが住んでいる

場所を教えてください)

以上の設問は準備なしでの発話能力を測るものである。日常生活の場面で用いられる表現の定着と自らの要求を伝える能力が確かめられる。ところで、このようなタイプの課題を検討する際に、学習者の回答は、いくつかの単語、いくつかのフレーズを含み、どれぐらいのスピードで行うかなどは重要な情報になるが、これらについてテストの要綱では説明されていない。

パート2の設問3 (問11) では、受験者は与えられたテキスト (テレビ番組表、180単語まで) について内容要約を行い、モノログの形で自分の意見を述べる。準備時間は15分間で、発話時間は5分間である。読解の際に辞書の使用が可能である。

Задание 3. Познакомьтесь с программой телепередач, которые Вы можете посмотреть в воскресенье.

Выберите передачи, которые Вы хотите посмотреть. Скажите:

- какие передачи, по-вашему, интересные;

- почему Вы хотите их посмотреть.

06.00— «Доброе утро»

09.00— «Новости»

09.30— «Здоровье». Сегодня в передаче Вы встретитесь с известными китайскими врачами и познакомитесь с восточной медициной. (以下略)

この課題では、自分の好きな番組と選択の理

由を伝達しなければならないので、自由度の高い表現やより広い選択肢の中から適切な表現を選ぶ。

パート3の設問4 (問12) では、受験者は所与のテーマと3つの質問を手がかりに7文までのモノログ形式の発話を行う。準備時間は9分間、発話時間は6分間である。

Задание 4. Расскажите:

- Какие передачи смотрят обычно по телевизору ваши родители, братья, сёстры, жена (муж), дети?
- Почему они смотрят эти передачи?
- Какие передачи Вы любили смотреть, когда учились в школе?

この設問では、受験者は家族のメンバーが好むテレビ番組、好みの理由、または、受験者が学校時代にどんな番組が好きであったかについて説明しなければならない。このテストは、実際の生活を反映する課題にたいしての、完全に自由スタイルの発話を要求するテストである。テストの実施時間は45分間であるので、いろいろな言語能力を測る総合テストのスピーキング下位テストとしては、異例に長い時間であって、口頭発表力が徹底的に測られる。テストは高い信頼性を持つ。ところで、妥当性の観点からみれば、問題がないと言えない。

99年版の口頭発表力テスト設問3の難点はモノログの発話にある。受験者には辞書使用可能な10分間の準備時間が考えられている。発話する前にテキストの内容を要約しなければならないので、回答するとき、口頭発表力だけでなく、解読力と記憶力も必要である。すなわち、測ろうとしている能力以外の能力がかかわる。これは、この設問と口頭発表力のテスト全体の内容妥当性(content validity)を下げる。

林田理恵は、05年版の一つ上の、基礎レベルの口頭発表力テストについて次のように指摘する。99年版のテストと比べて、「口頭発表力テストでは第3問の読解用テキストが量・内容ともに難易度が高くなっている点を除いては、さほど大きな変化点は見られない<sup>9)</sup>」。というのは、ロシア語検定試験の難易度の向上であるトレンド

を反映して、基礎レベルで口頭発表力テストの半分に近い時間を占める、辞書の利用を認める、難易度のかかなり高い設問3は、もっと難しくなった（単語数だけで見ると400ワードから600ワードまで）。

ところで、初級レベルの04年版の口頭発表力テストは大きく変わった。すなわち、設問1と設問2はそのまま残っているが、注目の設問3は、逆に、テストから姿を消した。同時に、設問4の量は大きくなった：所与のテーマと10の質問を手がかりに10－12文までのモノログ形式の発話を行うようになった。さらに、設問3（元の設問4）の3つの案が出されている。04年版で設問3が除かれた後、テストの妥当性はより高くなった。テストの時間も短縮して25分間になって、テストはより経済的になった。しかし、受験者にあまり

に自由をあたえてしまうのも、課題の中からひとつを選ぶのも、信頼性の観点から見ると問題があると言える。

#### タスクの採点

スピーキング・テストの採点方法には、「全体的採点」(holistic scoring)と「分析的採点」(analytic scoring)がある。「全体的採点」の特徴は、複数の観点がひとつのバンドの中に入れられている点である。これに対して、「分析的採点」では、複数の観点が立ち、その観点ごとに採点していく。このため、一般に、「全体的採点」に比べ高い信頼性が得やすいといわれている（根岸、2007）。

99年版の初級レベルの口頭発表力テストの『評価用チェックリスト』では、以下の能力や手段がチェックされる。

#### 評価用チェックリスト

表 1. 設問1・2 － 20点、各問 － 5点

1	コミュニケーション課題に適切に対応しているか	課題が達成されていない場合に採点しない
2	待遇表現の誤り	各 -0.5
3	コミュニケーションに障害のある語彙・文法・文体上の誤り	各 -1
4	コミュニケーションに障害にならない語彙・文法・文体上の誤り	各 -0.5
5	発音・イントネーションの重大な誤り	各設問単位で -2
6	発話の完結性と十分な展開	+1
7	発音・イントネーションの正確さ	+1
8	テーマ展開と言語材料使用の独自性	+1

表 2. 設問3 - 40点

1	再現したテキスト情報の適切さ	
2	再現したテキスト情報の正確さ	内容の逸脱について 各 -5
3	情報の論理的展開と整合性	各 -2
4	コミュニケーションに障害のある語彙・文法・文体上の誤り	各 -2
5	コミュニケーションの障害にならない語彙・文法・文体上の誤り	各 -0.5
6	発音・イントネーションの重大な誤り	設問全体で -2
7	発話の完結性と十分な展開	各 +5
8	発音・イントネーションの正確さ	各 +1

表 3. 設問4 - 50点

1	コミュニケーション課題に適切に対応しているか	課題が50%以上に達成されていない場合に -30点
2	与えられたテーマに対する適切さ	テーマからの逸脱度に応じて -10~30
3	発話量	規定量以下の場合 -10
4	情報の論理的展開と整合性	各 -1
5	コミュニケーションに障害のある語彙・文法・文体上の誤り	各 -2
6	コミュニケーションに障害にならない語彙・文法・文体上の誤り	各 -0.5
7	発音・イントネーションの重大な誤り	設問全体で -2
8	発話の完結性と十分な展開	+5
9	テーマ展開と言語材料使用の独自性	+5
10	発音・イントネーションの正確さ	+1

「評価用チェックリスト」の内容を見ると、言語運用能力を測るために3つの標準が利用されている。「コミュニケーションに障害のある語彙・文法・文体上の誤り」や「コミュニケーションの障害にならない語彙・文法・文体上の誤り」や「発音・イントネーションの重大な誤り」である。その他テストされるのは、課題に適切な対応、目的への達成度、再現した情報の正確さと情報の量、発話の論理性、発話のスピードなどの能力である。学習者はコミュニケーション能力を持っているというのは、学習対象である外国語を担う人との実際のコミュニケーションのときに、その国の言語待遇的習慣に応じて、相互理解と相互関係の課題を解決するという考え方である。次に、すべての能力は、『評価用チェックリスト』でどのように反映するかどうか見よう。

言語構造能力は、言語記号とその統合ルールによって言語的に正しい文を理解し、生産する能力である。この要素は、まず「語彙・文法テスト」で、または「口頭発表力」と「作文力」で測定される。「口頭発表力」テストで、「コミュニケーションに障害のある語彙・文法・文体上の誤り」の場合の減点(-2)は「コミュニケーションに障害にならない語彙・文法・文体上の誤り」の減点(-0.5)より多い。言語構造の能力の低いレベルは、「聴解力」と「読解力」のテストで情報知覚に負の影響を与えることができる。

コミュニケーション活動を行う際に、個々の文ではなく、纏まった発話を行う。そのため、次の重要な要素とは、個々の文を、いろいろな統合論的、語彙的手段を使って、整合性のある発話にあわせる能力、すなわち語用的能力である。語用的能力は、生産的な言語運用能力を測定するテストの場合に、テストの重要な対象となる。そのため、整合性の欠如の時に(-1)点が減点される。

コミュニケーション運用能力のもう一つの要素は、社会言語的な背景に相応しい言語を理解し、発話できる社会言語的な能力である。この要素は「口頭発表力」や「作文力」や「聴解力」のテストで測れる。「口頭発表力」テストでは、起こりやすい問題は、学習者が違う言語・構造

サブシステムの材料を使うことである。たとえば、正式な場での俗語の使用である。

コミュニケーション活動は、いつも一定の目的を持つ。情報の整理、コミュニケーションの伝達などを担当する言語運用能力である。その内容は『基準』で紹介されている。この能力は、「口頭発表力」や「作文力」や「聴解力」テストで測ることができる。この能力は、特に重要な意義を持つ。学習者は、いくら高い言語構造的なレベルで発話しても、テストで与えられたコミュニケーション課題に適切に対応しなければ、発話が採点されず、0点になる。

実際のコミュニケーション活動の重要な特徴は、相手の行動と発話の形と内容を予測できないことの、偶然性の要素を含むことだろう。そのため、テストの際に、創造的な側面を持つスキルがテストされる。コミュニケーション活動の相手は、常にお互いの行動を予測し、お互いに援助し、実際のコミュニケーション活動を判断し、必要な場合に言語戦術を変える(別の言葉や違う手段と構造を選ぶ)。ある場合には、コミュニケーション方略も変える。このようにコミュニケーション活動の際に言語知識の不足と外国語での言語運用と社会交際の経験の不足を補う能力である。この能力はテストで非常に高く評価される。「口頭発表力」テストの設問4では、テーマの展開と言語材料の独自性は最高の点数の(+5)点で採点される。

コミュニケーション活動は、社会、政治、イデオロギー、文化、民族性などのコンテキストの中で行われる。そのため、言語地域学能力も区別される。受験者は、コミュニケーション活動の際に、コミュニケーション相手の国の特徴、民族の歴史などを考慮しなければならない。自分の国にはないレアリティ(もの、現象、思想)を表す外国語の言葉と表現を知っていて、正確に利用しなければならない。現代の外国語としてのロシア語検定試験の制度ではこの能力は専門的にテストされていないが、将来、ある専門的な能力のテストの際に測れることがありうる。

言語運用能力を測定するときに、評価の対象

だけでなく、評価のパラメーターの件も検討しなければならない。「口頭発表力」テストでは、「コミュニケーションに障害のある語彙・文法・文体上の誤り」の場合の減点(-1, -2)と「コミュニケーションに障害にならない語彙・文法・文体上の誤り」の減点(-0.5)である。減点の限度は定めていない。ミスの特に多い場合、すなわち文法・語彙・意味理解能力のレベルが低い場合に、受験者は不合格にされる。

「評価用チェックリスト」は、コミュニケーション活動をテストするために必要な能力を測る。妥当性を下げる要素、すなわち、採点において、測ろうとしている能力以外のものを測るということは見られない。しかし、採点者信頼性を高めるには、まず、明確な基準を持つことが重要である。スピーキング能力のテストの対象のリスト、その対象のレベル分け、採点の基準などを含む試験官向けのマニュアルの作成は急務である。また、この採点基準を運用するためのトレーニングも重要である。ロシアにおいて、周年にテストを行う試験官にとっては、特別のトレーニングは必要ないが、ロシア国外の試験官、または教員に、複雑な基準による採点の能力の維持は、定期的な練習なしで、きわめて難しい。

## V. 波及効果

外国語としてのロシア語のテスト全体は、学習者と教育過程全体にポジティブな影響をあたえる。すべての技能を測定する総合テストであるから、合格するために全面的な準備（勉強）が必要である。教育過程でよく軽視されがちな技能の練習（たとえば、聞き取りと会話）の必要性が早い段階で浮かぶ。根岸雅史は、日本の外国語教育にたいするテストの変化の影響について次のように指摘する。「これまでに大学入試センター試験にリスニング・テストが含まれていなかったために、受験者がリスニングの学習をしてこなかったことや、TOEFL と TOEIC にこれまでライティングやスピーキングの能力を直接測るテスト問題が含まれていなかったために、受験者はこれらの技能に関する学習を行ってき

ていなかったと思われる。これらは、いずれも望ましい波及効果とはいえなかったであろう。今日では、大学入試センター試験にはリスニングが、TOEICではスピーキングとライティングがそれぞれ選択で、TOEFLにはスピーキングとライティングが必修で課されるようになった。これらの導入は、受験者の学習に大きな変化をもたらすものと思われる<sup>9)</sup>。

ロシアでは、検定試験の制度の導入は、外国語としてのロシア語の教育過程に大きな影響を与えた。各レベルの教育は、はっきりとした目標と内容を持つカリキュラムによって行われるようになった。最後の10年間に出版している大多数の総合教科書は、何レベルに向いているか明記される。個別の技能の教材と練習問題も、どのレベルとどの技能を練習することが明記される。これにより、教科書の選択と教材の組み合わせが楽になった。

日本では、教科書を書いている大学の教員は、自分の大学の教育環境と他の理由で、自分なりに教材を作る。沢山ある『ロシア語入門』、『初級ロシア語』、『標準のロシア語』などは、どんな基準でタイトルで示したレベルの内容を選んだか、レベルに必要な技能はどの割合で練習するかのことは説明されていない。筆者の知っているかぎり、初級レベルと基礎レベルの検定試験に向けて、5つの技能（知識）を練習させる教科書は、札幌大学の『テレモク』だけである。ロシア教育省認定のテストの、教育環境と教育思考の違う日本での普及率は低いので、沢山の教材の作成は期待できない。ところで、日本では、もう50年間以上、4級の固有のロシア語検定試験が行われている。この試験向けの教科書がないので、教員は、一般の教科書を利用しながら工夫をこらさなければならない<sup>10)</sup>。

## VI. テストの実用性

外国語としてのロシア語のテストが高いレベルの内容を持つだけでなく、テストの実施も困難であるため、テストの「実用性」は限られている。たとえば、試験官の確保が大問題である。テスト本部は、試験を実施したい機関と契約を

結んで、試験を行うために試験官を派遣する。これはテストのコストを上げる。現地の教員によつての実施の場合、試験官の資格の習得は、72時間の有料特別教育コース（テスト理論の教育とテスト実施の実習を含む資格習得コース）を要求するので、日本では、ごくわずかな教育施設に資格試験官がいるのみである。

テストの考案者たちは、できるだけ多くの能力を直接テストしたいので、試験官の、受験者の一人当たりのテスト時間が長い。ロシアでは、テストは、申し込みが入り次第いつでも行われるので、大学入学合否判定テストとコース卒業テストの受験者の多い時以外に問題は起こらないが、ロシア国外で多数の受験者がいた場合、テストの実施は難しくなる。

## VII. おわりに

全体的に、外国語としてのロシア語の初級レベルのテストは、コミュニケーション言語運用能力と言語構造能力の内容を十分に測定している。テストは、教育過程で現実的に教えられるすべての技能を測定するので、かなり高い外面妥当性を持っている。ロシア語の学習者のレベルを徹底的に測定するために、テストの作成者は、言語の基礎である文法の主な内容をチェックして内容妥当性を高めるため、コミュニケーションテストとしての真正性を犠牲にすることを覚悟した上で、伝統的な多肢選択肢の文法・語彙の下位テストを導入した。

ロシア語検定テストの口頭発表力テストは直接的な会話の形で行われるので、高い妥当性を持つ。採点は、「分析的採点」の方式で行われるので信頼性が高い。テストの管理者は、採点者に対する検査官間の信頼性を保つために、大きな注意を払う。99年版テストの妥当性の観点からみた難点だが、04年版で改善された。「評価用チェックリスト」では、妥当性を下げる要素、すなわち、採点において、測ろうとしている能力以外のものを測るということは見られない。04年版で、テストの時間も短縮して、テストはより経済的になった。しかし、受験者にあまりに自由を与えてしまうのも、課題の中からひとつを選ぶの

も、信頼性の観点から見ると問題が残る。

ところで、あるものの得点はその欠点になることも少なくないだろう。テスト全体のあまりの難しさと限られた実用性のため、ロシア国外での普及に大きな問題がある。

日本では、この問題の解決の方法として、いくつかの可能性がある。まず、非専攻ロシア語の教育の場合に、ロシア教育省認定の試験の、独自の価値を持つ各下位テストを必要に応じて利用する。特に、伝統的なテストの形を持つ、生徒に違和感の少ない文法・語彙テストを挙げることができる。次に、ロシア教育省認定の試験を拒否して、より簡単な日本固有の検定試験だけを適用することである。また、両テストをカリキュラムに入れて、段階的に適用する。さらに、一番理想的な方法は、一定のコースを終えてから、テストへの準備を条件づきのロシアでの短期留学を利用して、留学中に受験する。青森公立大学のすべての留学生はこの形でテストに合格した。いずれにしても、検定試験の教育過程への適用は、研究すべき大きな問題として残されている。

(2010年12月15日受付、2011年1月26日受理)

## 注

- 1) 白山利信「ロシア語検定試験」『日本私学教育研究所調査資料』243、2007、p.96
- 2) 北岡千夏「日本の大学におけるロシア語能力検定試験実施の試み」『外国語教育フォーラム』関西大学外国語教育研究機構、第2号、2003、p.119
- 3) 武藤研介「日本の大学におけるロシア語教育の現状に関する調査研究（Ⅶ）－ロシア語能力検定試験の導入の問題点」『京都外国語大学研究論叢』第70号、2007、p.85
- 4) Common European Framework of Reference for Languages（ヨーロッパ言語共通参照枠）
- 5) 林田理恵「ロシア教育省「ロシア語能力検定試験」の概要・現状と課題」－林田理恵研究代表者『到達度評価制度構築のための「国際基準」に準拠したロシア語総合試験開発』大阪大学大学院言語文化研究科、2010、p.3

- 6) Андрюшина Н.П. Разработка лексико-грамматических тестов по РКИ для начального этапа обучения// Вестник ЦМО МГУ, 1998, №1, с.27
- 7) 林田理恵の専門用語の利用は多いに参考になった
- 8) 林田理恵「ロシア教育省「ロシア語能力検定試験」の概要・現状と課題」 — 林田理恵研究代表者『到達度評価制度構築のための「国際基準」に準拠したロシア語総合試験開発』大阪大学大学院言語文化研究科、2010、p.25
- 9) 根岸雅史『コミュニケーション・テストングへの挑戦』三省堂、2007、p.4-5
- 10) 武藤研介「日本の大学におけるロシア語教育の現状に関する調査研究（Ⅶ） — ロシア語能力検定試験の導入の問題点」『京都外国語大学研究論叢』第70号、2007、p.85

## 参考文献

### 和文

- アーサー・ヒューズ（2003）『英語のテストはこう作る』研究社
- 白山利信（2007）「ロシア語検定試験」『日本私学教育研究所調査資料』243
- 北岡千夏（2003）「日本の大学におけるロシア語能力検定試験実施の試み」『外国語教育フォーラム』関西大学外国語教育研究機構、第2号
- 小池生夫編集主幹、寺内正典、木下耕児、成田真澄編集（2004）『第二言語習得研究の現在』大修館書店
- 小林潔（2006）「『外国語としてのロシア語』が目指すもの」『外国語教育論集』28
- 小林潔（2007）「ロシア語教育とヨーロッパ共通参照枠」 — 中澤英彦・小林潔編『ロシア語学と言語教育』東京外国語大学
- 静哲人、竹内理、吉澤清美共編著（2002）『外国語教育リサーチとテストングの基礎概念』関西大学出版部
- 『世界の言語テスト。Language Tests in the World』国立国語研究所、2006
- 授業づくりハンドブック — ロシア語（2008）『授業づくりハンドブック — ロシア語』ロシア語教育委員改編著、大阪大学出版会
- チャールズ・オルダーソン、キャロライン・クラッファム、ダイアン・ウォール（2010）『言語テストの作成と評価 — 新しい外国語教育のため』春風社

- トルストグーゾフ・アレキサンダー（2004）「外国語としてのロシア語の国家教育基準（初級レベル）について」『青森公立大学紀要』第10巻・第1号、2004年9月
- トルストグーゾフ・アレキサンダー（2008）「外国語としてのロシア語の国家教育基準（初級レベル）について（2）」『青森公立大学紀要』第13巻・第2号、2008年3月
- トルストグーゾフ・アレキサンダー（2009a）「外国語としてのロシア語の国家検定試験（初級レベル）について」『青森公立大学紀要』第14巻・第1号、2008年9月
- トルストグーゾフ・アレキサンダー（2009b）「外国語としてのロシア語の国家検定試験（初級レベル）について（2）」『青森公立大学紀要』第14巻・第2号、2009年3月
- トルストグーゾフ・アレキサンダー（2010）「外国語としてのロシア語の国家検定試験（初級レベル）のバージョンの比較」『青森公立大学紀要』第15巻・第2号、2010年3月
- 根岸雅史（2007）『コミュニケーション・テストングへの挑戦』三省堂
- 武藤研介（2007）「日本の大学におけるロシア語教育の現状に関する調査研究（Ⅶ） — ロシア語能力検定試験の導入の問題点」『京都外国語大学研究論叢』第70号、2007
- 林田理恵（2010）林田理恵研究代表者『到達度評価制度構築のための「国際基準」に準拠したロシア語総合試験開発』大阪大学大学院言語文化研究科
- バックマンL.F.、パーマーA.S.（2000）『＜実践＞言語テスト作成法』大修館書店

### 露文

- Андрюшина Н.П.(1998) Разработка лексико-грамматических тестов по РКИ для начального этапа обучения// Вестник ЦМО МГУ, 1998, №1, с.27
- Андрюшина Н.П. и др.(1999) Типовые тесты по русскому языку как иностранному. Элементарный уровень. Общее владение М.—Спб.; ЦМО МГУ— «Златоуст»
- Н.П.Андрюшина, Т.В.Козлова.(2000) Лексический минимум по русскому языку как иностранному. Элементарный уровень. Общее владение М.—Спб;

ЦМО МГУ– «Златоуст»

Антонова В.Е. и др.(2004) Типовые тесты по русскому языку как иностранному. Элементарный уровень. Общее владение М–Спб.; ЦМО МГУ– «Златоуст»

Балыхина Т.М. (2006) Словарь терминов и понятий тестологии. Русский язык - Курсы, М

Владимирова Е.Т. и др.(1999) Государственный образовательный стандарт по русскому языку как иностранному. Элементарный уровень. Проект М–Спб.

Образовательная программа по русскому языку как иностранному. Элементарный

уровень. Базовый уровень. Первый сертификационный уровень. М., 2001.

### 英文

Stoynoff Stephen and Chapelle A.Carol (2005) ESOL Tests and Testing, TESOL.

Davies A., Brown A., Elder C., Hill K., Lumley T., McNamara T. (1999) Dictionary of language testing, Volume 7, UCLES/CPU

McNamara T.(2000) Language Tersting. Oxford University Press

Underhill N. (1989) Testing Spoken Language. A handbook of oral testing techniques. Cambridge University Press

# About the Oral Test of Russian Federation State Test for Speakers of Russian as a Foreign Language (Beginner's Level)

Tolstoguzov ALEXANDER

## Abstract

This article is devoted to an analysis of the elementary level test of Russian, which is one of the official tests of the Russian state system of Russian as a Foreign Language (TORFL). It's a standardized test which determines whether the student has achieved the elementary level of competency in Russian as a foreign language. The test consists of five sections. Each section tests a different type of language proficiency.

This study aims to investigate the test of Russian entirely and the oral performance test especially, using Bachman and Palmer's scheme of test usefulness. This scheme includes reliability, construct validity, authenticity, interactiveness, impact and practicality.

The first version of the oral test of Russian as a foreign language, officially published in 1999, is compared with the new version, published in 2004. The new version has some new features both in types of tasks and their arrangement inside the test.

The results indicate that the whole test has a great degree of reliability and validity, but poor practicality because of its long (two-day) format and assessment difficulties.

After the length of the test was reduced from 45 minutes to 25 minutes, by removal of tasks, and after dictionary use was permitted, it became a more valid and authentic test.