

看護師国家試験における外国人候補者のテストパフォーマンス —テスト提示条件とテスト問題の影響を 日本人看護師・日本人大学生との比較から探る—

堀場裕紀江

深谷 計子

要旨

本研究は、外国人候補者が看護師国家試験を受ける際にどのような要因がテストパフォーマンスに影響するかを調べるために、母国で看護師の資格を持つインドネシア人候補者（L2候補者）に、3種類のテスト提示条件（時間延長、ルビ付き、母語）で試験問題（必修・一般・状況設定問題から成る）を解いてもらい、そのテスト結果を分析した。また、日本語母語話者で内容知識を持つ看護師（L1看護師）と持たない大学生（L1学生）のテスト結果とも比較した。分析の結果、L2候補者の成績は、テスト提示条件間で有意な差がなく、どの条件間にも信頼できる中程度の相関が見られた。また、問題の種類と分野については、状況設定問題と一部の一般問題で成績がかなり低く、さらに詳しく分析したところ、内容知識と日本語力の両方の個人差がテスト結果に影響していることが分かった。

キーワード：外国人看護師、看護師国家試験、テスト提示条件、
テスト問題、母語

1. はじめに

2008年以降、経済連携協定 (EPA) の締結により毎年インドネシアやフィリピンから約200名の外国人看護師候補者が来日し、2010年4月の時点で370人に達した。日本政府側から課された就労条件は3年以内に看護師国家試験に合格することで、合格できなかった場合は帰国を余儀なくされる。しかし、2009年2月の試験では合格者は0で、2年目の2010年は254名が受験したが、合格者は3名（インドネシア人2名とフィリピン人1名）だった。すなわち、外国人受験

者の合格率は1.2%だった。これは日本人受験者の約90%という合格率と比べてきわめて低い数値である。来日した外国人看護師候補者が母国の大学や専門学校で看護を学び実践経験を持つことから、合格の難しさは「言語の壁」によるものと推測され、この問題が有識者やその他の関係者の間で活発に議論されるようになった。

しかし、外国人候補者が看護師国家試験を受ける際に直面する困難は言語(日本語)の問題だけであろうか。言語の問題だけであるとすれば、そのハンディを取り除いた、あるいは、その負担を軽減した形で受験することができれば、外国人候補者は既に持っている看護知識を反映したテスト結果を出して合格ラインに達することができるのではないか。例えば、試験時間の延長、問題文の漢字に振り仮名をつける、母語で受験する、等である。このような提案は様々な憶測とともになされているが、その効果について調べた研究はまだ報告されていない。そこで、外国人候補者が看護師国家試験を受ける際に直面する困難にどのような要因が関わっているかを調べるために、テスト方法の影響に焦点をあてて模索的な研究を行った¹⁾。本稿では、まず看護師国家試験について簡単に説明し、関連する第2言語(L2)読解の先行研究について概観する。その後、本研究の結果について報告する。

2. 背景

2.1. 看護師国家試験

看護師国家試験は、看護師として就業するために必要な専門知識と能力が備わっているかを判断するための資格試験で、毎年1回行われる標準テストである。テスト方法としては、短い文を読んで多肢選択により応答する形式の筆記テストである。出題分野としては、「人体の構造と機能」「疾病の成り立ちと回復の促進」「社会保障制度と生活者の健康」「基礎看護学」「在宅看護論」「成人看護学」「老年看護学」「小児看護学」「母性看護学」「精神看護学」がある。問題の種類としては、必修、一般、状況設定の3種類がある。必修問題と一般問題は医療・看護に関する特定の知識について問うもので、状況設定問題は医療・看護に関する知識を特定の状況設定のもとで使用できるかについて問うものである。表1に問題例を示す。合格基準は必修問題の80%、一般および状況設定

問題の60～70%とされている。問題数は2010年度より変更されたが、2009年度までは必修問題30問、一般問題150問、状況設定問題60問であった。

この試験の問題で使われる日本語について日本語能力試験出題基準を参考に分析した齋藤 (2010) は、文法については日本語能力試験出題基準の2級までのものが使われているが、語彙については一般的な語彙だけでなく、専門分野の用語や一般日本語では使用されない特殊語や表現が多く、日本語能力試験出題基準の1級までのリストには含まれていない語彙が問題文の50%に用いられていると報告している。また、表1に示した例からも判断できるように、試験問題で使われる日本語には一般の日本語母語話者が読んでも理解するのが難しい漢字や語彙・表現が数多く使用されている。

表1 看護師国家試験の問題の例

問題の種類 (分野)	問題例
一般 (人体)	上腕を外転させる筋肉はどれか。 1. 大胸筋 2. 三角筋 3. 上腕二頭筋 4. 上腕三頭筋
状況設定 (成人)	次の文を読み11～13の問いに答えよ。 Aさん、62歳の男性。印刷工場で長年働いている。最近、肉眼的血尿が出現し泌尿器科を受診した。診察の結果、入院し精密検査を受けることになった。 11 硬膜外麻酔下で膀胱鏡による膀胱組織の生検が予定された。 最も確認する必要があるのはどれか。 1. 検査前の絶飲食 2. 検査中の下肢のしびれ 3. 検査直後の眠気 4. 検査後の血尿の増強 12 膀胱癌と診断され、膀胱全摘術と回腸導管造設による尿路変更術が行われることになった。術後の生活に関する説明で適切なのはどれか。 1. 「尿意は残ります」 2. 「尿の出口が二つになります」 3. 「自己導尿が必要となります」 4. 「パウチの装着が必要となります」 13 退院2週後の朝、Aさんから「昨夜から39.5℃の熱が続き、だるくて腰も痛い。食欲もないし、水も飲む気にならない」と泌尿器科外来に電話があった。
	確認する情報で優先度が最も高いのはどれか。 1. 咽頭痛の有無 2. 排ガスの有無 3. 排尿の量と性状 4. 脈拍数とリズム

2.2. 関連先行研究

L2読解に影響を与える最も重要な要因として、言語（テキストの言語とそれに関する読み手の知識）と内容（テキストの内容とそれに関する読み手の知識）がある。記述文を読んで理解するために言語に関する知識と能力が必要なのは言うまでもない。言語習熟度は、語の認識や単文の意味の理解だけでなく、テキスト全体の内容の理解に大いに影響を与える（加藤, 2002; 菊池, 1997; Horiba, 1993）。これらの先行研究によると、言語習熟度が低い読み手の場合、言語処理、とくに語認識、節や文レベルの意味の理解など、下位レベルの処理が効率よく行われず、文と文を結びつけたり異なる文から抽出した命題を統合するといった上位レベルの処理が十分に行われない。その結果、読み時間が長くかかり、テキスト内容に関する質問の正答率や内容再生率が低いことが明らかにされている。

また、言語習熟度が低い読み手は、L2の文やテキストの処理に、母語の知識や処理ストラテジーを使うことが多く、その影響はL2の習得にも現れる（Koda, 2005）。例えば、視覚提示語の意味理解に関する代表的な理論「二重ルートモデル」によると、語の意味の理解には、語の視覚認識に続いて、視覚情報から音韻情報へと変換されて意味にアクセスするルートと、視覚情報から音韻情報を介さずに意味にアクセスするルートの2つのルートがある。その両方の情報処理が競合的に行われて語彙認識が達成されるが、どちらのルートによってより頻繁に意味へのアクセスが起こるかということは特定言語の表記と語彙の関係によって異なっている（Perfetti, Liu & Tan, 2005）。日本語の表記は仮名と漢字という2種類の文字体系によって行われるが、仮名で表記された語の認識は、視覚情報から音韻情報へと変換されて意味にアクセスされるのに対し、漢字で表記された語は、視覚情報から音韻情報を経て意味にアクセスされるルートと、視覚情報から音韻情報を介さずに意味にアクセスされるルートがある。一方、英語やフランス語、インドネシア語などは表音文字のアルファベットを使用しており、一般に視覚情報から音韻情報へと変換されて意味にアクセスされると考えられる。例えば、L2日本語の漢字の処理について調べた Mori (1998) は、疑似漢字を提示された時、英語話者が音韻情報に依存した処理を

行うのに対し、中国語話者と韓国語話者は視覚情報に依存した文字処理を行ったという結果をえている。

母語の影響は文やテキストのレベルの処理にも現れる。日本語は SOV という語順で文が構成されるが、語順は比較的柔軟で、項と項の関係を表す格助詞が重要である。また、文と文の繋がりを示す照応関係は省略によることが多い。一方、英語や中国語、インドネシア語は SVO という語順で文が構成され、項と項の関係を表す格助詞のようなものをもたず、代名詞などで照応関係を示すことが多い。インドネシア語は、時制やアスペクトを示す動詞の語形変化をもたず、複合語を作る接辞が多く用いられる。L2日本語の文処理ストラテジーを調べた Koda (1993) は、語順は母語において語順が重要なグループ（英語と中国語）には影響したが、母語において語順がそれほど重要ではないグループ（韓国語）には影響しなかったと報告している。また、L2日本語の語彙知識の質的特徴と読解の関係について韓国語話者と中国語話者を比較した Horiba (2010) は、母語と日本語の言語的相違によってテキスト処理と語彙・統語などの下位処理の関係が異なることを示唆する結果を報告している。

読解に関わるもう一つの重要な要因として、テキストの内容に関する知識が挙げられる (Kintsch, 1998)。テキスト文で扱われているトピックや内容に関する知識がない場合、あるいは、持っているても活用することができない場合、言語に関する知識があってもテキスト内容を正確に理解することはできない。L2日本語を対象にした研究は数少ないが、工学系の中国人学習者を対象にした研究 (山田, 1995) では、文法知識と読解の間に相関関係はみられなかったが、テキスト内容に関する背景知識と読解の間には相関関係がみられたとしている。また、アメリカ人学習者の物語文読解を日本語母語話者および英語母語話者の読解と比較した研究 (Horiba, 1996) では、上級学習者は、どちらの母語話者グループよりも多くの読み時間がかかるが、読解中の推論生成については英語母語話者よりも日本語母語話者に類似していたと報告している。これは、読み手がテキスト内容に関連のある背景知識を使って、下位レベルの言語処理の効率の悪さを補おうとするストラテジーが現れたものと考えられる (Stanovich, 1980)。

本研究では、これらの関連先行研究からの知見をもとに、インドネシア語を

母語とする候補者（以下、L2候補者）がL2日本語で看護師国家試験を受ける際に言語知識による影響と内容知識による影響がどのように関わるのかという問題に焦点をあて、筆記試験におけるテストパフォーマンスについて調べた。具体的には、テスト結果に影響を与える要因として、テスト方法に関する要素、使用されるインプット（言語）と解答時間の長さを取り上げ（c.f., Backman, 1990）、インプットを日本語で提示した2条件（すなわち、解答時間を1.5倍に延長した条件、すべての漢字にルビをつけた条件）とインプットをインドネシア語で提示した母語条件を設定し、そのテスト結果を比較した。また、言語知識と内容知識の両方を持っている日本人看護師と、言語知識を持っているが内容知識を持っていない日本人大学生についても調べ、これらを参考にL2候補者のテスト結果を分析した。

3. 本研究

3.1. 課題

本研究の課題として、以下の2つの質問を設定した。

質問1 L2候補者は、テスト提示条件によって、看護テストの成績が異なるか。

(1) 母語で受けた場合（L1条件）は、日本語で受けた場合（L2条件）と比べて、テスト成績がよいか。

(2) 日本語で受けた場合、漢字にルビをつけたルビ付き条件は、漢字にルビをつけていない時間延長条件と比べて、テスト成績がよいか。

質問2 L2候補者は、日本語を母語とする看護師（以下、L1看護師）、および、大学生（以下、L1学生）と比べて、テスト成績がどのように異なるか。

3.2. 参加協力者

調査協力者はインドネシア語を母語とする看護師候補者28名（女性21名、男性7名、年齢20代～30代）である。これは同時期に来日した104名の27%にあたる。全員が母国で看護学校または看護大学を卒業し、2～10年前に免許を取得し2～9年の看護実践経験を持つ。来日前の日本語学習経験はなく、来日後1年4、5カ月であった。参考にした日本母語話者は看護師10名（全員女性、30代～40代、6～14年の看護実践経験をもつ大学院生または看護教員）と大学

生10名(全員女性、18才～20代、看護大学入学後1週間以内の学部1年生)であった。

3.3. 材料

調査では看護師国家試験を参考にした看護テストを作成して使用した。このテストは、過去5年間の看護師国家試験の問題の中から種類と内容と言語的特徴を考慮して選んだ問題を使って作成した。問題の選定は、日本語教育の専門家が看護大学教員の意見を参考に、「系統看護学講座」編集室(2009)より候補となる問題を選び、厚生労働省の開示した正答を参考にして問題を絞り、インドネシア人候補者が背景知識を欠くと思われる領域(「社会保障制度」、「在宅看護」、「老年・小児・母性看護」)や内容を扱った問題を除き、最終的に必修問題6問、一般問題24問(人体3問、疾病3問、基礎6問、成人9問、精神3問)、状況問題9問(成人9問)を採用した。対象となる看護知識および言語に関する難易度がほぼ同等の3バージョン(A・B・C)を用意した。問題数は協力者の参加時間の制約を考慮して決定し、問題の種類による割合は実際の試験での割合に合わせた。A・B・Cのバージョンそれぞれにつき、日本語版の他に、すべての漢字にルビをつけたルビ付き版、および、インドネシア語に翻訳した母語版を作成した。母語版は、社団法人国際厚生事業団(JICWELS)の許可を得てそのe-learningサイトに提示されている翻訳版をそのまま使用した。

3.4. テスト提示条件

L2候補者は、L2時間延長、L2ルビ付き、L1の3種類の提示条件でテストを受けた。L2時間延長条件では日本語版テストの時間を1.5倍に延長して解答した。L2ルビ付き条件では、日本語版テストの問題で使用されている漢字に全てひらがなでルビがつけられたテストを、時間延長なしで解答した。協力者は各条件でA・B・Cのうちの1バージョン、全部で3条件で3バージョンのテストを受けた。テストのバージョンと提示条件の組み合わせは協力者間でカウンターバランスをとった。L1看護師とL1学生は、日本語版の3バージョンのテストを時間延長なしで、A・B・Cの順に解答した。

3.5. 手順

初めに、調査者が協力者に調査の内容と手順を口頭と母語による文書で説明し、同意書を得た後、協力者の背景情報（性別・年齢・看護歴・専門分野など）を質問紙に記入してもらった。次に、L2候補者はL2時間延長条件（30分）、L2ルビ付き条件（20分）、L1条件（20分）の順でテストを計70分で受けた。L1看護師およびL1学生は日本語版の3バージョンのテストをそれぞれ20分ずつ、計60分で受けた。L2候補者は、それぞれが勤務している医療機関の研修室等で数名のグループ単位でテストを受け、調査者の他に当該機関関係者が立ち会った。L1看護師およびL1学生は、所属している教育機関の部屋で、数名のグループ単位でテストを受けた。

3.6. 分析方法

看護師国家試験の配点にならい、必修問題と一般問題の計10題は各1点、状況設定問題3題は各2点とし、各バージョン16点満点とした。分析は、まずテストバージョンによる影響がないことを確認した。次に、各協力者のテスト全体得点、および、問題の種類ごとの得点を提示条件ごとに算出し、提示条件間の比較およびグループ間の比較をt検定により行った。また、L2候補者の提示条件による得点の変動を記述的に分析した。さらに出題分野の影響について調べるために、各問題の正答率を提示条件ごとに算出し、提示条件間の比較を記述的に行った。

4. 結果とその考察

はじめに、テストバージョンによる問題の正答率の結果を表2に示す。グループごとに対応のあるt検定（ $\alpha = .006$ ）を行ったところ、いずれもバージョン間で正答率に有意な差はなかった（L2候補者：Paired-t = .81/ 1.11/ 1.92, n.s.; L1看護師：Paired-t = .42/ 1.36/ 1.95, n.s.; L1学生：Paired-t = .73/ .89/ .17, n.s.）。よって、テスト問題は難易度の点からバージョン間で同等に配分されていると考えられる。

表2 グループのテストバージョンによる問題の正答率

グループ	テストバージョン					
	A		B		C	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
L2候補者	.48	.24	.53	.21	.56	.24
L1看護師	.65	.26	.80	.15	.77	.21
L1学生	.33	.23	.42	.18	.35	.28

4.1. L2候補者のテスト提示条件における結果

次に、L2候補者の提示条件におけるテスト結果を分析した。表3に問題の種類別の得点の平均と標準偏差を示す。全体得点を見ると、L1条件 (Range = 3-16, M = 8.4, SD = 3.2)、L2ルビ付き条件 (Range = 1-14, M = 8.3, SD = 3.3)、L2時間延長条件 (Range = 3-15, M = 8.4, SD = 3.2) と、どの提示条件においてもかなりの幅があり、平均および標準偏差が提示条件間で似通っている。L2候補者の全体得点が提示条件によって異なるかどうか調べるために、Paired-t 検定 ($\alpha = .017$) により分析したところ、L1条件は、L2ルビ付き、L2時間延長と比べて、有意な差がなく (L2ルビ付き: Paired-t = .11, n.s., L2時間延長: Paired-t = .18, n.s.)、L2ルビ付きとL2時間延長の間にも有意な差はなかった (Paired-t = .05, n.s.)。問題の種類別の比較においても、必修・一般 (Paired-t = 0/ 1.94/ 1.97, n.s.)、状況設定 (Paired-t = .17/ 1.47/ 1.49, n.s.) とともに、提示条件間で有意な差は見られなかった。

表3 L2候補者のテスト提示条件による問題の種類別の得点

テスト提示条件	問題の種類					
	全体 (16点満点)		必修・一般 (10点満点)		状況設定 (6点満点)	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
L2時間延長	8.4	3.2	5.4	1.8	3.0	1.8
L2ルビ付き	8.3	3.3	4.7	1.8	3.6	1.9
L1	8.4	3.2	5.4	1.8	3.1	2.1

提示条件による相関関係を調べたところ、いずれの条件間でも信頼できる中程度の相関がみられ、L1条件とL2ルビ付き条件との間の相関 ($r = .51, p <$

.01) は、L1条件と L2時間延長条件の相関 ($r = .41, p < .05$)、L2ルビ付きと L2時間延長との相関 ($r = .39, p < .05$) と比べてやや高かった。

以上の結果から、テスト提示条件は L2候補者のテスト得点の違いにあまり影響しなかったと考えられる。このような結果は、テスト全体得点についても問題種類別の得点についても観察された。すなわち、L2看護テストを母語（インドネシア語）で受けても、第2言語（日本語）で受けた場合と比べてよりよい成績が出なかったということは、テストで使用される言語（L1か L2か）がテストパフォーマンスに一貫した影響を与えなかったことを示唆する。また、日本語で看護テストを受ける時に、漢字にルビがついたテスト問題を（時間延長なしで）解答した場合、ルビなしで1.5倍のテスト時間で解答した場合と比べて同等の成績が得られたことから、漢字で表記された語彙の音韻化が語の意味の認識につながったか否か、それが問題文の理解を容易にしたか否かについても一貫した影響がなかったことを示唆する。

また、異なる提示条件間で信頼できる中程度の相関が見られたことから、ある特定の提示条件でよい成績をあげた L2候補者は別の提示条件においても比較的よい成績をあげており、ある提示条件下で成績のよくない L2候補者は他の提示条件下でも成績が芳しくないという傾向があることが判明した。本研究に参加した L2候補者は来日前の日本語学習経験がなく来日後1年4、5カ月であったことを考えると、看護知識における個人間の差異がテスト結果にかなり影響しているのではないかと考えられる。そこで、L2候補者の個人の特徴について詳しく調べた。

表4は、L1条件でのテスト全体得点をもとに上位からランク付けし、L2ルビ付き条件および L2時間延長条件でのテスト全体得点と、その L1条件での得点との違いを示したものである。必修・一般問題は1問1点、状況設定問題は1問2点という採点基準を参考に、L2ルビ付き条件、L2時間延長条件のテスト得点が L1条件のテスト得点と比べて3点以上得点の差があるケースの特徴を調べた。L1条件に比べて L2条件で低い場合は太字と*で、高い場合は太字と**で示した。

この表を全体的にみると、提示条件による得点の変動のしかたにかなりの個人差があることが伺える。提示条件によって得点が大きく異なる L2候補者(例

例えば、27番) もいれば、あまり変わらない L2 候補者 (例えば、2 番) もいる。また、L1 条件に比べて、L2 ルビ付き条件で 3 点以上下がっているケース (4 番、5 番、11 番、14 番、23 番) より、L2 時間延長条件で 3 点以上下がっているケース (1 番、4 番、5 番、7 番、8 番、12 番、13 番、16 番) の方が多く、中にはルビ付き条件と時間延長条件の両方で 3 点以上下がっている者 (4 番、5 番) もいる。しかし一方で、L1 条件に比べて L2 条件で逆に得点が 3 点以上上がっている候補者 (例えば、9 番、18 番、27 番) も存在し、このようなケースは L1 条件での得点が比較的低い方の候補者により多く見られる。

表4 L2 候補者のテスト提示条件によるテスト全体得点

L2 候補者 ID 番号	テスト提示条件 (16 点満点)			(A) と (B) の平均
	L1	L2 ルビ付き (L1 との差 (A))	L2 時間延長 (L1 との差 (B))	
1 番	16	14 (-2)	13 (-3) *	-2.5
2 番	12	12 (±0)	13 (+1)	+0.5
3 番	12	10 (-2)	12 (±0)	-1
4 番	12	9 (-3) *	5 (-7) *	-5 *
5 番	12	7 (-5) *	7 (-5) *	-5 *
6 番	11	11 (±0)	15 (+4) **	+2
7 番	11	9 (-2)	8 (-3) *	-2.5
8 番	11	9 (-2)	7 (-4) *	-3
9 番	10	13 (+3) **	14 (+4) **	+3.5 **
10 番	10	12 (+2)	8 (-2)	±0
11 番	10	7 (-3) *	12 (+2)	-0.5
12 番	10	11 (+1)	6 (-4) *	-1.5
13 番	10	8 (-2)	4 (-6) *	-4 *
14 番	9	4 (-5) *	10 (+1)	-2
15 番	8	6 (-2)	10 (+2)	±0
16 番	8	10 (+2)	5 (-3) *	-0.5
17 番	8	6 (-2)	7 (-1)	-1.5
18 番	6	9 (+3) **	11 (+5) **	+4 **
19 番	6	10 (+4) **	8 (+2)	+3 **
20 番	6	7 (+1)	9 (+3) **	+2
21 番	6	9 (+3) **	5 (-1)	+1
22 番	6	8 (+2)	4 (-2)	±0
23 番	6	2 (-4) *	8 (+2)	-1
24 番	5	6 (+1)	7 (+2)	+1.5
25 番	5	4 (-1)	3 (-2)	-1.5
26 番	4	6 (+2)	9 (+5) **	+3.5 **
27 番	3	13 (+10) **	8 (+5) **	+7.5 **
28 番	3	1 (-2)	6 (+3) **	+0.5

これらの結果について、以下のようなことが考えられる。テスト問題が母語で提示されたL1条件では、既存の看護知識を使って解答することができるためよい成績をあげられるが、L2日本語で提示された条件では問題文の日本語の処理が効率的に行われないため、看護知識を持っていてもそれを活用することができないということが結果に反映されている。また、L2ルビ付き条件の方がL2時間延長条件と比べて得点の下がっているケースが少ないという結果から、ルビによって漢字表記された語彙を音韻化しやすくなっていれば、その意味と結びつけやすく、既存の看護知識を使って応答するという行為が引き出されやすいということが現れていると考えられる。さらに、L2条件でL1条件とほぼ同等（あるいは、それ以上）の得点を出しているケースについては、来日以降の学習の効果が現れている可能性とも考えられる。

本研究に参加したL2候補者は母国で看護教育を受け看護師としての実務経験を持っているが、大学・専門学校等のプログラム修了によって資格を得ており、日本の看護師国家試験のような試験に合格して資格を得ている訳ではない。従って、L2候補者が受けた教育プログラムの違いによる影響や、L2候補者が実際に持っている看護知識、とりわけ、看護師国家試験で測定される明示的な分析された看護知識をどのくらい持っているかについてかなりの個人差があるのではないかと考えられる。また、L2候補者は来日前に日本語学習経験がなく、来日以降の1年半弱の間、医療機関で勤務のかたわら看護国家試験合格のために指導を受け学習に取り組んできたということから、その効果が看護テストの結果に一部現れた可能性がある²⁾。

4.2. L2候補者とL1看護師およびL1学生のテスト結果との比較

次に、L2候補者のテストパフォーマンスに関わる言語と内容に関する知識がどのようなものかを別の角度から調べるために、テスト結果をL1看護師およびL1学生のテスト結果と比較分析した。表5に各グループのL1条件におけるテスト得点を示す。L1条件での全体得点を見ると、L2候補者 (Range = 3-16, M = 8.4, SD = 3.2)、L1看護師 (Range = 6-16, M = 11.8, SD = 1.4)、L1学生 (Range = 3-10, M = 6.0, SD = 1.1) と、どのグループも得点にかなりのばらつきがあるが、L2候補者の平均はL1看護師とL1学生の間位置し、L2候補

者の標準偏差は L1 看護師および L1 学生と比べて大きい。

表5 グループの L1 条件における問題の種類別のテスト得点

グループ	問題の種類					
	全体 (16点満点)		必修・一般 (10点満点)		状況設定 (6点満点)	
	平均	標準偏差	平均	標準偏差	平均	標準偏差
L2候補者	8.4	3.2	5.4	1.8	3.1	2.1
L1看護師	11.8	1.4	7.4	1.0	4.4	.6
L1学生	6.0	1.1	3.5	.7	2.5	.8

L1条件でのテスト全体得点についてグループ間による違いがあるか否かを分散分析を用いて調べたところ、グループによる有意な効果 ($F(2,47) = 12.65, p < .0001$) が見られた。グループ間比較を Student-t 検定で行ったところ、L1看護師は L2候補者より有意に成績がよく、L2候補者は L1学生に比べて有意に成績がよいことが分かった。テスト問題の種類別に調べたところ、必修・一般では、テスト全体の結果と同様、グループによる有意な効果が見られ ($F(2,47) = 17.37, p < .0001$)、L1看護師は L2候補者より成績がよく、L2候補者は L1大学生に比べて成績がよいことが分かった。また、状況設定では、グループによる有意な効果が見られた ($F(2,47) = 3.45, p < .05$) が、L1看護師は L2候補者および L1学生に比べて成績がよく、L2候補者と L1学生との間には有意な差がないことが分かった。

よって、L2候補者は、一般に、母語で看護テストを受けた場合、日本人看護師と比べて低い成績を出しているが、日本人大学生と比べてより優れた成績を出していることが分かる。しかし、L2候補者のテストパフォーマンスは問題の種類によって異なっている。L2候補者は、必修・一般問題では日本人学生と比べて優れた成績を出しているが、状況設定問題での成績は日本人学生の成績とあまり差がなかった。これらの結果から、L2候補者は、看護テストで測定される看護分野の知識をある程度持っているが、状況設定問題で対象とされる看護状況で使うことが求められる知識や能力が不十分である、あるいは、

持っていても応答する際に使うことができないのではないかと考えられる。

必修・一般問題では医療・看護に関する特定の知識の有無が問われるが、それに対して、状況設定問題では医療・看護に関する知識や技能を特定の状況や場面のもとで使用できるか否かについて測定される。状況設定問題では、記述描写されている状況を理解するために医療や看護に関する専門知識だけでなく、人や事物や出来事など一般的な文化社会知識が必要で、その状況を把握した上で看護知識を使って応答することが要求される。本研究で使用したテスト問題は、インドネシア人協力者が背景知識を欠くと思われる領域や内容を扱った問題を含んでいないが、状況設定問題で扱われる一般的な看護状況であっても、それが日本の社会で起きる看護状況であるために、文で描写されている状況における人物（患者と家族および医療関係者）と一連の行動や出来事を十分に理解することが難しく、関連する看護知識を持っていても正答を得られないのではないかと。さらに、必修・一般問題で測定される看護知識の場合と比べると、日本における看護状況で必要とされる知識は、母国における看護状況で必要とされる知識と比べて異なる部分がより多いのではないかと。特定の国や地域で行われる看護状況には歴史・地理・政治経済的条件も含めた文化社会要因が関わっている³⁾。そのため、母国で得た看護教育と看護実践に基づく専門知識と技能を使っても、日本の看護状況で必要とされる知識と技能を持っていないければ、国家試験の状況設定問題で正答を得ることは難しいということが結果として出ているのではないかと。

続いて、L2候補者のテストパフォーマンスに問題の種類と分野による影響があるかをテスト結果をもとに分析した。L2候補者の問題正答率を、L1グループの正答率とともに、表6に示す。

表6 テスト問題の種類と分野ごとのグループの平均正答率 (%)

種類	分野	問題数	L2候補者			L1看護師	L1学生
			L1	L2ルビ付き	L2時間延長		
必修	必修	6	82	75	77	90	53
	計	6	Range= 82~82	Range= 71~79	Range= 74~79	Range= 80~100	Range= 37~70
一般	人体	3	26	17	29	40	20
	疾病	3	61	33	43	70	40
	基礎	6	64	47	58	80	27
	成人	9	38	42	44	72	32
	精神	3	42	50	64	77	30
	計	24	Range= 18~72	Range= 17~61	Range= 29~65	Range= 40~80	Range= 20~40
状況	成人	9	52	61	50	73	42
	計	9	Range= 46~64	Range= 57~65	Range= 47~53	Range= 67~80	Range= 37~53

問題の種類と分野によって出題数に差異があるため、単純な比較はできないが、L1条件における成績から判断すると、L2候補者は、必修問題に比べて、一般問題のいくつかの分野（「人体」「成人」「精神」）や状況設定問題で正答率が低く、正答率にばらつきが大きい分野（「成人」）もあることが観察できる。また、L2候補者の一般問題の正答率は、「人体」と「成人」の分野についてはいずれの提示条件でも低いが、その他の分野については提示条件によってパターンが異なっている。

L1条件における問題別の正答率についてグループ間の相関関係を調べたところ、L2候補者グループは、どちらのL1グループとも信頼できる正の相関が見られたが、L1看護師グループとの相関 ($r = .73, p < .005$) はL1学生グループとの相関 ($r = .58, p < .04$) に比べて高いことが分かった。これらの結果から、L2候補者のテストパフォーマンスには、テスト問題で扱われる看護の分野やトピックによる影響もあるが、テスト問題による影響の中には、日本語母語話者（L1看護師・L1学生）に見られる影響と共通する部分もあると考えられる。

最後に、本研究に参加したL2候補者にとって難しかった問題の例を表7に挙げる。これらのテスト問題がL2候補者のテストパフォーマンスにどのように影響したかについて本研究では十分な証拠はないが、上記の結果をもとに考えると、特定のテスト問題で扱われている特定の看護知識と、それを測定するために提示されたテスト問題の言語の特徴（表記・語彙・文法など）とテスト形式の特徴（多肢選択式）の両方が複雑に絡んでいるのではないかと考えられる。

表7 L2候補者の正答率が低かったテスト問題の例

種類(分野)	問題例
L1条件での正答率が極めて低かった問題	
一般(成人)	冠状動脈造影検査で穿刺に最も多く用いるのはどれか。 1. 総頸動脈 2. 橈骨動脈 3. 尺骨動脈 4. 鎖骨下動脈
一般(精神)	パニック発作でみられるのはどれか。 1. 便秘 2. 強い怒り 3. 強い予期不安 4. 間代性けいれん
L1条件での正答率に比べてL2ルビ付き条件の正答率が極めて低かった問題	
一般(基礎)	注射の準備で適切なのはどれか。 1. 注射針の刃面は注射器の目盛り面と反対側にする。 2. ガラスのアンプルはカット後にカット面を消毒する。 3. バイアルはゴム栓を下にして薬液を吸う。 4. 針についた薬液はアルコール綿で拭き取る。
一般(人体)	肺拡散能に影響を与えるのはどれか。 1. 肺胞表面積 2. 気道抵抗 3. 死腔換気量 4. 残気量

5. 結論

本研究では、外国人候補者（L2候補者）が看護師国家試験を受ける際に直面する問題、すなわち、内容知識と言語知識がテストパフォーマンスにどのように影響するかについて調べるために、L2候補者の3種類のテスト提示条件（時間延長、ルビ付き、母語）でのテスト成績をもとに比較分析した。また、L2

候補者のテスト結果を、日本語を母語とする L1看護師と L1学生のテスト結果を参考にして、問題の種類や分野による影響についても分析した。その結果、L2候補者のテスト成績は、L1条件は L2条件（時間延長・ルビ付き）と同程度で、提示条件間で信頼できる中程度の正の相関があった。また、L2候補者のテスト成績はテスト全体および必修問題では L1看護師に比べて低く、L1学生に比べて高かったが、状況設定問題では L1学生と有意な差がなかった。また、L2候補者の正答率は出題分野によってもかなり異なっていた。これらの結果から、L2候補者のテストパフォーマンスには看護に関する知識や能力（専門および一般）と言語に関する知識や読解力による影響があり、それ以外に、日本社会に関する一般知識による影響もあると考えられる。また、L2候補者のテストパフォーマンスにはかなりの個人間の差異があり、その個人差には看護に関する知識と言語に関する知識の両方が複雑に絡んでいるのではないかと考えられる。

最後に、本研究で対応できなかった点や今後の研究への課題について述べる。まず、本研究に参加した L2候補者について、来日前に母国（インドネシア）で受けた看護教育と医療機関での実践経験についての詳しい情報を得ることができなかった。インドネシアでは看護師国家試験は課されず、教育機関のカリキュラムを修了することで看護師の資格が与えられることから、実際にどのような看護知識が習得されているかの判断が難しい。また、関連 L2読解研究から明らかにされているように、L2読解における母語での経験に基づく読解力や受験ストラテジーの影響も考えられる。今後の研究では、対象とする L2候補者についてより多くの情報を収集した上で、L2候補者のテストパフォーマンスを調べる必要がある。また、先日提出された有識者会議のとりまとめによると、新しい看護師国家試験ではテスト問題文で使用される日本語の平易化や簡素化、英語の併記などが行われるという。今後の研究では、実際の看護師国家試験に出題される問題の特徴を多方面から考慮して、受験者のテストパフォーマンスがどのように影響されるかを調べる必要がある。

早急に対応を迫られている外国人看護師候補者の受け入れ体制については、日本語に関する知識と運用力だけでなく、看護に関わる専門分野の内容に関する知識とその応用力についても実態を把握した上で、個人差にも配慮して、そ

れに適した教育的措置（カリキュラムと教材の開発、学習指導）などを検討する必要がある。また、今後、外国人看護師を受け入れていくのであれば、アメリカやカナダの前例のように、医療現場で必要とされる専門知識と言語コミュニケーション能力がどのようなものを明示化してそれに合致した政策や教育的方策を展開していく必要があるだろう。

注

- 1) 本研究は文部科学省科学研究費補助金（基盤研究C 21520594）を受けて行われたものである。本研究に参加協力して下さった看護師候補者および関係者の方々に感謝いたします。また、テスト作成に協力して下さった齋藤隆氏に感謝します。
- 2) 本研究に参加したL2候補者は、本研究参加の約1ヶ月後に2度目の看護師国家試験を受け、そのうちの2名が合格した。この2名は本研究の看護テストで最も成績のよかったL2候補者で、母語（L1）条件だけでなく、日本語（L2）のルビ付き条件と時間延長条件でも高得点を出していた。かれらのテスト結果には日本語での看護師国家試験受験のための1年半にわたる学習の効果が反映していたと考えられる（五十嵐他, 2011）。
- 3) このような看護師国家試験で前提とされている一般世界知識をもっていないために、外国人候補者が問題を解こうとする際に様々な困難に遭遇するという事は、日本語指導の現場からも報告されている（池田他, 2011）。

参考文献

- 加藤由香里（2002）「学術論文の内容理解における視覚情報の活用」 日本語教育 114号, 11-19.
- 五十嵐博美・樋口博一・Y.F. フェルナンデス・R. アグスティナ・深谷計子（2011）「インドネシア人看護師候補者の国家試験合格への道：三之町病院の取り組み」 聖路加看護大学紀要 37号
- 池田敦史・深谷計子・堀場裕紀江（2011）「インドネシア人看護師候補者への日本語指導：ある病院での実践から」 聖路加看護大学紀要 37号
- 菊池民子（1997）「日本語の読解におけるテキスト構造の影響と読解事前指導の効果」 日本語教育 95号, 25-36.
- 系統看護学講座編集室（2009）『2010年版 系統別看護師国家試験問題一解答と解説』 医学書院

- 国際交流基金・日本国際教育支援会（1994）『日本語能力試験出題基準 [改訂版]』凡人社
- 齋藤隆（2010）「日本の看護師国家試験問題の言語的分析—日本語能力試験出題基準との比較を通して」 修士研究報告. 神田外語大学
- 山田みな子（1995）「読解過程にみられる既有知識の影響と文法能力の関係について」 日本語教育 86号, 26-38.
- Bachman, L. F. (1990) *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Horiba, Y. (1993) The role of causal reasoning and language competence in narrative comprehension. *Studies in Second Language Acquisition*, 15, 49-81.
- Horiba, Y. (1996) Comprehension processes in L2 reading: Language competence, textual coherence, and inferences. *Studies in Second Language Acquisition*, 18, 433-473.
- Horiba, Y. (2011) Word knowledge and its relation to text comprehension: A comparative study of Chinese- and Korean-speaking L2 learners and native speakers of Japanese. *The Modern Language Journal*.
- Kintsch, W. (1998) *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.
- Koda, K. (1993) Transferred L1 strategies and L2 syntactic structure in L2 sentence comprehension. *The Modern Language Journal*, 77, 490-500.
- Koda, K. (1996) L2 word recognition research: A critical review. *The Modern Language Journal*, 80, 450-460.
- Koda, K. (2005) *Insights into second language reading: A cross-linguistic approach*. New York: Cambridge University Press.
- Matsunaga, S. (1999) The role of kanji knowledge transfer in acquisition of Japanese as a foreign language. *Japanese-Language Education around the Globe*, 9, 87-100.
- Mori, Y. (1998) Effects of first language and phonological accessibility on kanji recognition. *The Modern Language Journal*, 82, 69-82.
- Mori, Y. (2003) The roles of context and word morphology in learning new kanji words. *The Modern Language Journal*, 87, 404-420.
- Perfetti, C. A., Liu, Y., & Tan, L. H. (2005) The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review*, 112, 43-59.
- Stanovich, K. E. (1980) Towards an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32-71.