

コミュニケーション能力をいかに測るか  
The Australian Second Language Proficiency  
Ratings (ASLPR) を通しての考察

智 原 哲 郎

Some Considerations on Communicative Language Testing:  
Introduction of the Australian Second Language Proficiency  
Ratings (ASLPR)

Tetsuro Chihara

抄 録

言語テストリサーチは、Spolsky (1978) の区分による「前科学時代」、「心理測定学・構造言語学時代」、「総合的・社会言語学時代」を経て、1990年代にはコミュニケーション言語能力をより明確に定義、測定する方向に進んできた。本稿では、オーストラリアにおいて言語運用能力を測定するものとして開発、使用されている Australian Second Language Proficiency Ratings (ASLPR) を紹介し、また、その信頼性・妥当性・実用性の検討を通して第二言語／外国語学習者のコミュニケーション言語能力をいかに測るかを考察した。

キーワード：言語テスト、コミュニケーション言語能力、逆流的影響

(1996年9月6日 受理)

Abstract

According to Spolsky (1978), the history of language testing research can be broken into three major periods: the “prescientific” period, the “psychometric-structuralist” period, and the “integrative-sociolinguistic” period. In the 1990s, the trend is toward defining “communicative language ability” more clearly and assessing such proficiency more precisely. The Australian Second Language Proficiency Ratings (ASLPR) is one of direct tests which is said to assess general language proficiency. In this paper, the ASLPR is introduced and the implications for communicative language test construction and use are discussed.

**Key words:** language testing, communicative language ability, washback effects

(Received September 6, 1996)

## 1 はじめに

言語テスト理論と教授理論は互いに密接な関連があり、またそれらはその時代の言語理論によって変化する。或る言語理論で外国語能力の概念が規定されると、教授法や言語テスト理論研究もその概念を反映するものとなる。Spolsky (1978) は現代に至るまでの言語テスト理論研究の変遷を、「前科学時代」(Prescientific period)、「心理測定学・構造言語学時代」(Psychometric structuralist period)、「総合的・社会言語学時代」(Integrative-sociolinguistic period) の3つに区分してまとめているが、これによって言語テスト理論がいかにその時代の言語理論に反映されていたかを見ることができよう。

1950年代前半までの前科学時代においては、いわゆる「伝統文法」(Traditional grammar) を中心とする教授法が盛んであり、話しことばよりもむしろ書きことばに重点をおいていた。従って、テストは文法・訳読(grammar-translation) によるものが主であり、また科学的な言語テスト リサーチが発達していなかったため教師の主観的な評価に頼っていた。

1960年代後半までの心理測定学・構造言語学時代では、当時の構造言語学 (Structuralist linguistics) と行動主義心理学 (Behaviorist psychology)<sup>(1)</sup> の影響を受けて「部分的測定法」(Discrete-point testing) が主流となった (Lado 1961)。言語を読む・書く・聴く・話すの4技能に分け、さらにそれぞれを発音・語い・文法などの構成要素に分割してそれぞれの部分を測定し、それらを合計すれば全体の言語能力が表わせるとした。まさに「言語を使うこと」と「言語の構造を知ること」とが同一視された時期であった。また、テストの信頼性 (Reliability)、妥当性 (Validity)、実用性 (Practicality) を重視し、科学的統計処理によって客観的に言語能力を測定しようとした。この点では、この時期に科学的な言語テストリサーチ法が確立されたと言えよう。

確かに、部分的測定法では言語についての知識を客観的に測定することは可能であった。しかしながら、新たな疑問は、果たして「コミュニケーションの道具としての言語」を運用する能力を分析的に捉えることが妥当であるかということであった。Oller (1976, 1979) は「言語能力は統合的な相互作用能力であり、部分の和はその総和より大である。従って、言語能力を部分に分けて測定することは妥当ではない」として「総合的測定法」(Integrative testing) を提唱した。これは言語を細分化することなく、内在する言語運用能力<sup>(2)</sup>を測定しようとしたものであり、この測定法には言語の余剰性を利用したノイズテスト (Noise test)、ディクテーション (Dictation)、クローズテスト (Cloze test) などが主たるものであり、それらに関する数多くのリサーチがなされた (Gradman & Spolsky 1975; Oller & Conrad 1971; Oller & Streiff 1975; Oller 1970, 1972, 1976, 1979)。Oller は、本来「読む」能力を測るとされるクローズテストと「聴く」能力を要求するディクテーションとの相関が .90 に近いことを根拠として、すべての言語<sup>(3)</sup>に内在する共通の言語運用能力が存在するとした。彼はこの言語能力の仮説を「単一特性仮説＝言語能力はその構成要素ごとに分割できない単一な能力である」(Unitary trait hypothesis) とし

た (Oller 1979)。

しかしながら、その後、数多くの言語能力の本質に関するリサーチがなされた結果、Oller の仮説は誇張されすぎており、「言語能力には一般能力と個別能力が存在し、言語の四技能はそれぞれ異なった特性を持つ」ことが明らかにされた (Farhady 1982; Bachman & Palmer 1981, 1982; Upshur & Homburg 1983)。かくして言語能力の仮説は単一特性仮説から「多特性仮説＝言語能力は多面性であり、複数の異質の能力からなる」(Multi-trait hypothesis) へと移行したのである。さらにテストにおいても、これまでのテストが測定しようとしている言語能力は間接的<sup>(3)</sup>で、実際の言語使用能力を測定していないとの批判を受け、1980年代から1990年代にかけての言語テストリサーチは、総合的・社会言語学時代を経て、四技能におけるコミュニケーション言語能力 (Communicative language ability) をより明確に定義し、測定する方向に進んできた。

Chomsky (1965) のいう「言語能力」(Language competence) だけではコミュニケーションには不十分であり、「言語が実際に使用される『場』での言語運用能力や知識が大切なのだ」とするのがコミュニケーション言語能力の構成概念である。Canale & Swain (1980) では、コミュニケーション能力 (Communicative competence) は「文法能力」(Grammatical competence)、「社会言語学的能力」(Sociolinguistic competence)、「方略的能力」(Strategic competence) から成るとし、後に Canale (1983) は社会言語学的能力をさらに「談話能力」(Discoursal competence) と「社会言語学的能力」に分類した。これらの能力は以下のように定義されている。

- (1) 文法能力→語彙・語形・統語・意味・音声などの規則についての知識、すなわち言語の構造についての理解と運用。
- (2) 談話能力→文法能力を互いに補って完全なものとする。例えば、文を連結させて一連の談話にする能力、一連の発話からまとまった意味あるものにする能力など。
- (3) 社会言語学的能力→言語使用の際の社会的・文化的規則や談話規則についての知識。例えば、会話の始め方や終わり方についての知識、話しかける相手と状況によってどの呼びかけ形式を用いるべきなど。
- (4) 方略的能力→上記の規則についての知識が不十分なためコミュニケーションが円滑に行われなくなった時にその埋め合わせとして使用される言語、非言語的コミュニケーション要素の運用。例えば、文法や語彙の知識不足を補うのに身振りやまねの使用など。

その後、Backman (1990) もこれを基にしてコミュニケーション能力の詳細な枠組を定義している<sup>(4)</sup>。しかしながら、これらの定義をテストの面から見れば、実際の言語使用に関わる社会言語学的能力や方略的能力を測定するテストの開発はかなり困難であるように思われる<sup>(5)</sup>。

このように、実際にコミュニケーション能力を測定するテストとなると、越えるべきハードルが数多くあるが、言語テストの流れは「言語知識」と「言語使用」の両方を測定

する方向に確実に向かっている。Brown (1994) はまとめとして次のようにコミュニケーション テストが持つべき基準を挙げている。

- 1) コミュニケーション能力の構成概念を持つこと
- 2) 内容が真正 (Authentic)<sup>(6)</sup> であること
- 3) 言語能力を直接測定 (Direct testing) すること
- 4) 多方法でテストすること

要するに、あるべきテストとは、複数のテスト方法で、実生活に即したコンテキストと言語で学習者の言語運用能力を直接測定できるもの、ということになる。そこで本稿では、オーストラリアにおいて、言語運用能力を測定するものとされている Australian Second Language Proficiency Ratings (ASLPR) を紹介し、それを通して第二言語/外国語学習者のコミュニケーション言語能力をどう測るかを考察してみる。

## 2 The Australian Second Language Proficiency Ratings (ASLPR) とは

### 2.1 概要

オーストラリアは毎年多くの移民を受け入れており、彼らが移住した直後に移民教育を行っている。彼等が移住して生活していく上でまず必要になるのは語学力 (= 英語力) だが、その語学教育プログラムの開発に着手する際に必要となったのは、そのプログラムのゴールと移住者が駆使できる第二言語能力とを関連させていくために彼らの第二言語能力発達のはっきりした道筋を示すことだった。さらに、その語学プログラムは実践的語学能力の訓練、すなわちコミュニケーション スキルの訓練を目指していたので、当然のことながら、その能力を適切に測れる「物差し」が必要となったのである (Ingram, 1979)。

この物差しとして1978年より着手・開発された ASLPR は、現在 ESL (English as a Second Language) 学習者のクラス配置や英語能力評価、クィーンズランド州における外国人教員応募者の英語能力の評価、移住地域別による移住者の英語能力レベルの評価、専門教育・職業機関への外国人応募者の英語能力の評価などに使用されている。

ASLPR は Speaking・Listening・Reading・Writing の4つの部分から成り、Speaking→Listening→Reading→Writing の順に、Writing 以外はすべて試験官との1対1の口頭インタビュー形式で行われる。所用時間は受験者の学力により多少異なるが約2時間程度である。試験中での使用言語は英語である。

Speaking・Listening・Reading・Writing 能力は個別に評価され、0-5の9段階の能力レベル (0, 0+, 1, 1+, 2, 3, 4, 5)<sup>(7)</sup> で表示される。大まかに言えば、レベル0がわずかな単語を知っている程度で英語使用能力はない、レベル2がいわゆる「サバイバル英語能力」、レベル3で専門分野での仕事が可能、レベル5は英語母国語話者に匹敵する<sup>(8)</sup>。

試験官は“General Proficiency Version for the Australian Second Language Proficiency Ratings” と呼ばれる評価スケールに従って受験者のレベルを決定する。四技能別の9段階のレベルごとに“General description of language behavior”、“Examples of

language behavior”、“Comment”の3つの欄があり、それぞれの欄には、例えば「受験者はx,Y,Zのことができ、その例としてA,B,Cがある」というように詳細に記載されている。

使用される題材は、あらかじめ試験官によってすべてのレベルのものが用意されている。試験官は、まずSpeakingの段階で決定した受験者のレベルを基にして、次のListening用の題材を決定する。仮に、受験者のSpeakingのレベルが2であるとする、Listeningでは最初にレベル2の題材を与えてみる。それを受験者がクリアすれば、更にレベル3、4の題材を、そうでなければレベル1の題材を与えるといった段階を経て受験者の能力を評価する。

## 2.2 内容

### 2.2.1 Speaking

試験官は、受験者が出来る限り自然に振舞えるように緊張をほぐしながら、受験者の生活・仕事・趣味・将来のプラン、国際・社会・経済問題などについて質問していく。探索→分析→結論という会話の流れの中で能力レベルを決定する。所要時間は15-20分程度。

#### 例

試験官：Hello. How are you? My name is Helen Davies. Today, I am going to have an interview with you. Could you tell me your name?...

### 2.2.2 Listening

試験官がテープを流し、その後その内容の概要を言わせたり、質問に答えさせる。留守番電話のメッセージやニュース、天気予報なども含まれる。所要時間は15分程度。

#### 例1 Level 1 [Answering phone]

試験官：Now I'm going to test your listening. You may need to write a few words.  
It's 9 o'clock in the evening, Friday. Your mother suddenly becomes sick while watching TV and you telephone a doctor's office. This is what you hear on the telephone. I'll play it once, and then ask what you would do.

TAPE：[Dr. Johnson's office is now closed. The office hours are Monday to Friday, 9 a.m. to 6 p.m. and Saturday, 8 a.m. to 12 noon. For emergency please call the following number: 477-0270, 477-0270]

試験官：It's 9 o'clock at night. What do you do?

#### 例2 Level 2 [news]

試験官：You will listen to news. I'll play it once.

TAPE：[The Queensland government is urging Canberra to drop visa requirements for international tourists visiting Australia. The Queensland Tourism Minister Bob Givies says that Australia is missing out on 40,000 Japanese visitors each year because of delay associated with obtaining visas. On the other hand, New Zealand doesn't require visas.]

試験官：Tell me in your own word what the topic is about.

You heard the words “Australia”, “New Zealand” and “Japanese”. What about them?

Why does the Minister want to do this?

### 2.2.3 Reading

日常生活に密接したもの、例えば新聞広告、パンフレット、新聞の社説、雑誌の記事などの内容理解に至るまで広範囲にわたっている。読む時間が与えられた後、内容の大意を言わせたり質問に答えさせる。所要時間は30分程度。

例

試験官：I want you to read this article and then tell me about it. I'll start with very general questions. Use this photocopy so that you can underline parts if you want to. Please begin.

### 2.2.4 Writing

出題は3問。日常生活に関連したもの。申し込用紙に記入させたり、同僚や友人に手紙やメモを書いたり、苦情の手紙や自分の見解を書かせたりする。所要時間は50分程度。

例 Level 2

[You received a letter from a friend who you met while you were traveling in Sydney. She wants to visit your country in the near future. Write a letter to her and suggest a few things.]

## 3 考察

ASLPR の最大の特徴は言語運用能力を直接測定しようとすることであろう。Writing 以外のスキルは試験官との1対1の口頭インタビュー形式で評価され、言語に関する知識ではなく言語の実際使用能力を測るとしている。また、従来のテストに見られるように受験者の学力の差異にかかわらず同一の問題が与えられるのではなく、ASLPR では学力に応じて適切な題材を与えることにより個人の言語運用能力を適切に評価できるとしている。「良いテスト」の条件である信頼性、妥当性、実用性の面から ASLPR を検証してみる。

### 3.1 ASLPR の信頼性

テストの信頼性は測定・評価には最も重要な条件であるが、通常、インタビューでは二つの信頼性が検討される。それらは、複数の採点者間の採点基準がどの程度一致しているかを見る採点者間信頼性 (Inter-rater reliability) と、採点者が同一能力を同じ採点基準で複数回評価した場合、どの程度の安定性が見られるかの採点者内信頼性 (Intra-rater reliability) である。ASLPR では、16名の成人 ESL 学習者のパフォーマンスがビデオに収録され (\*Writing はペーパーによる)、それらを21名のネイティブスピーカー (オーストラリア人英語教師) と15名のノンネイティブスピーカー (中国人英語教師)、それに2名の ASLPR 開発者が評価した (Ingram 1979)。

採点者間信頼性では、オーストラリア人英語教師と ASLPR 開発者間、中国人英語教師と ASLPR 開発者間との採点にそれぞれ高い相関が報告されている（オーストラリア人英語教師と ASLPR 開発者 = Speaking: .91 ~ .99; Listening: .89 ~ .98; Reading: .91 ~ .99; Writing: .93 ~ .997, 中国人英語教師と ASLPR 開発者 = Speaking: .88 ~ .98; Listening: .88 ~ .98; Reading: .88 ~ .98; Writing: .88 ~ .98)。さらに、上記のオーストラリア人英語教師に一年後に再評価させた結果の採点者内信頼性も高い相関が出ている（Speaking: .93 ~ .99; Listening: .91 ~ .96; Reading: .87 ~ .98; Writing: .92 ~ .98）。この結果を踏まえて、ASLPR はノンネイティブスピーカーでさえ高い信頼度で ASLPR を使用することができるとしている。

この高い信頼性を支えているものは評価スケールの詳細な説明と試験官の訓練であろう。20時間に及ぶワークショップにおいて、ビデオに収録されたインタビューを評価スケールに示されている基準で各試験官が評価した後、納得いくまで議論し、確認しあうことで試験官の評価の安定性を維持している。

しかしながら、さらに検討を要すると思われるのが使用題材の安定性であろうと思われる。定まったテスト項目がないという点で ASLPR は正確な意味でのテストではなく、評価スケールである。インタビューでは試験官自身が与える題材を決定し、独自の質問をする。となれば、試験場 X の試験官 A がレベル  $\alpha$  に適当と思われる題材と、試験場 Y の試験官 B のそれは果たして同質であろうか。また、選んだ題材の内容によって受験者のパフォーマンスも違って来るかも知れない。この分野でのリサーチが必要であろう。

### 3.2 ASLPR の妥当性

テストの妥当性には、表面的妥当性 [=テストは表面上、言語能力を測定しているように見えるか] (Face validity)、内容的妥当性 [=テストは掲げた言語能力評価目標をどの程度反映しているか] (Content validity)、構成概念妥当性 [=テストの構成概念は規定された言語能力の概念と一致しているか] (Construct validity)、併存的妥当性 [=テストは外部基準とどの程度相関があるか] (Concurrent Validity) があり、それぞれの面からテストを検討しなければならない。

ASLPR は、オーストラリアの様々な機関で広範囲に使用されており（表面的妥当性）、実生活に即したコンテキストと言語使用で学習者の言語運用能力を直接測定し（構成概念妥当性）、また、それが測ろうとするもの、受験者に示される問題の属性、受験者の応答の習性などが能力レベルに応じて規定され、それによって語学プログラムのゴールと学習者の第二国語能力の発達を関連づけている（内容的妥当性）。

ASLPR は CELT (Comprehensive English Language Test)<sup>9)</sup> を使い併存的妥当性を検討している。オーストラリアで学ぶ「第二言語としての英語」学習者を対象とした調査では ASLPR と CELT 間に高い相関が報告されているが ( $r = .8 \sim .96$ )、中国で学ぶ「外国語としての英語」学習者の場合は全般的に低い ( $r = .01 \sim .59; -.25 \sim .62$ )。この相関の違いを Ingram (1979) は「学習者が教室で学んだ言語知識を日常のコミュニケーション

活動に応用できる『第二言語』環境では、彼らのコミュニケーション能力と言語知識とは同じようなレベルになるだろう。他方、『外国語』環境、特に従来の文法中心で学ぶところでは、学習者の持つコミュニケーション能力と言語知識とは大きな違いが出てくるのであろう。」と説明している。

この相関は非常に興味深い。妥当性の検討において、テストが X の能力を測るとしながら、実は Y の能力を測っていないかという問題がある。例えば、コミュニケーション能力を複数方法で測った場合、そのテスト間の相関は高いものであるはずである。逆に、コミュニケーション能力と言語知識が独立した特性のものであるならば、それぞれを測るテスト間には高い相関はないということである。このことから CELT のような主に言語知識を測定するとされる間接テストと言語使用能力を直接測定する ASLPR とでは本来異なった特性を測定しているので、両者間には高い相関は期待できないはずであるが、オーストラリアのような第二言語環境では、学習者のコミュニケーション能力と言語知識は同じようなレベルになるのであろう。

### 3.3 ASLPR の実用性

テストの実用性は、実施容易性 (Administrability)、採点容易性 (Scoriability)、経済性 (Economy) などから検討される。ASLPR では実施には受験者一人あたり約 2 時間要するし、採点には試験官の訓練が必要になるので、多人数受験の場合は ASLPR の実用性は低いように思われる。Bachman (1990) はテストの実用性に関して次のように述べている。

One of the main reasons oral interviews and compositions are not more widely used, for example, is that they are very time-consuming, both to administer and to score. Thus, despite the fact that most of us would agree that these tests types can involve authentic language use, considerations of efficiency often take precedence over those of validity and authenticity. (298)

このように実際には実用性が先行する傾向にあるが、最優先は「テストは測定したいものを実際に測定しているか」である。コミュニケーション能力を測りたいとすれば、テストも言語の実際使用に則したものでなくてはならないのである。これがなされてこそ「テストの実用性」がうんぬんされるのではないだろうか。多人数の場合の一つの解決策として、まず実用性の高いテストで ASLPR が実施されるのに可能な一定の人数をスクリーニングすることが考えられよう。また、コンピュータテクノロジーの発達で言語テストの領域も従来不可能であったことを可能ならしめていることから、近い将来、実用性の問題の解決にはかなりの進展が期待できるであろう。

## 4 おわりに

オーストラリアで使用されている Australian Second Language Proficiency Ratings (ASLPR) を紹介し、それを通して第二言語/外国語学習者のコミュニケーション言語能



力をどう測るかを考察してきた。ASLPRのような測定法はわが国における今後の言語テストの在り方に示唆を与えてくれるものであろう。

最後に、教授法との関連において大事だと思われるのはテストの「逆流的影響」(Wash-back effects)であろう。通常、テストは学習の効果や成果を見るためのものであるが、逆流的影響とは、この関係が逆転した、テストが授業や学習に与える影響のことである。この影響はテスト如何で良くもなり悪くもなりうる。仮に、テストが文法力を問う多肢選択型のものであれば、学習者はそのテストの形式に合った勉強をするようになり、結果的には言語についての知識のみの習得に終わってしまうようになる。一般に、いわゆる「受験英語」だけではコミュニケーション能力がつかないと言われるゆえんであろう。

一方、カリキュラムにコミュニケーション テストを取り入れると、そのテスト内容のため、学生はコミュニケーション能力を高めるように動機づけがなされ、また、授業の方向もコミュニケーション能力の育成に向くようになるだろう。もちろん、テストのための授業ではあってはならないが、テストが授業のツールとして機能することによって目指す言語能力の習得につながっていくことが可能となるであろう。このように、コミュニケーション テストの導入は英語教育にも大きな意味を持つてくると言えよう。

#### 注

- (1) 行動は本質的にその部分の集合としてとらえることができるとするもの。
- (2) Oller はこれを「予測文法」(Expectancy grammar) と呼んでいる。
- (3) 多肢選択法の文法テストのような部分的測定法では言語知識の測定が言語能力測定であるという点で間接的である。また、ディクテーションやクローズテストのような総合的測定法では、学習者の各言語スキルを統合させながら、それぞれのテスト項目を言語の出来事と結び付けているという点で半間接的である。いずれのタイプも言語能力を直接測定していない。これに対してインタビュー テストや作文テストは実際の言語使用に焦点を当てる、すなわち、われわれが実際のコミュニケーションを行う際に行使する言語能力を直接見るので直接的であると言える。
- (4) 言語能力を「構成能力」(Organizational competence) と「実用能力」(Pragmatic competence) に分け、さらにそれぞれを「文法能力」(Grammatical competence)、「テキスト能力」(Textual competence) と「発語内能力」(Illocutionary competence)、「社会言語学的能力」(Sociolinguistic competence) に分けている。詳しくは Backman (1990), 84-100 を参照。
- (5) Farhady (1983) は社会言語学的能力を測るものとして 'Functional language testing' を紹介している。これは多肢選択テストで、選択肢には 1) 社会的にも言語的にも適切な応答、2) 社会的に不適切だが言語的に適切な応答、3) 社会的には適切だが言語的に不適切な応答、4) 社会的にも言語的にも不適切な応答が含まれる。しかしながら、方略的能力に関しては依然として未開の分野であろう。
- (6) 言語の実際使用に即した状況のテスト内容でなければならないとする。
- (7) 評価スケールには記述されていないが、更に (2<sup>+</sup>, 3<sup>+</sup>, 4<sup>+</sup>) のレベル設定が可能とされているので、実際のレベルは12段階となっている。
- (8) Ingram (1991) の TOEFLスコアとの換算表によると、ASLPR のレベル 2 が TOEFL で 400, 2<sup>+</sup> が 450 ~ 500, 2<sup>+</sup> ~ 3 が 500 ~ 550, 3<sup>+</sup> が 580, 4 が 600 としている。
- (9) 中・上級 ESL 学習者の英語能力測定に幅広く使用されているテスト。「聴き取り」「文法」「語彙」の3部門で構成されている。

参考文献

- Bachman, L. F. 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L. F. and Palmer, A. 1981. 'The construction validation of the FSI oral interview'. *Language Learning* 31, 67-86.
- Bachman, L. F. and Palmer, A. 1982 'The construction validation of some components of communicative proficiency'. *TESOL Quarterly* 16, 449-465.
- Brown, H. D. 1994. *Principles of Language Learning and Teaching*. Third edition. Englewood Cliffs, NJ: Prentice-Hall.
- Canal, M. 1983. 'From communicative competence to communicative language pedagogy'. In Richards, J. C. and Schmidt, R. (eds.), *Language and Communication*. London: Longman Group Ltd.
- Canal, M. and Swain, M. 1980. 'Theoretical bases of communicative approaches to second language teaching and testing'. *Applied Linguistics* 1, 1-47.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.
- Farhady, H. 1982. 'Measures of language proficiency from the learner's perspectives'. *TESOL Quarterly* 16, 43-59.
- Farhady, H. 1983. 'New directions for ESL proficiency testing'. In Oller, J. W. Jr. (ed.), *Issues in Language Testing Research*, Rowley, MA: Newbury House Publishers, 253-269.
- Gradman, H. L. and Spolsky, B. 1975. 'Reduced redundancy testing: a progress report'. In Jones, R. L. and Spolsky, B. (eds.), *Testing Language Proficiency*. Arlington, Virginia: Center for Applied Linguistics, 59-70.
- Ingram, D. E. 1979. *Introduction to the Australian Second Language Proficiency Ratings (ASLPR)*. Brisbane, QLD: Language Testing and Curriculum Centre, Griffith University.
- Ingram, D. E. 1984. 'Assessing proficiency: An overview on some aspects of testing'. In Hyltenstam, K. & Pieneman, M. (eds.), *Modeling and assessing second language acquisition*, Clevedon, Avon: Multilingual Matters.
- Lado, R. 1961. *Language Testing: The Construction and Use of Foreign Language Tests*. London: Longman Group Limited.
- Oller, J. W. Jr. 1970. 'Dictation as a device for testing foreign language proficiency'. *English Language Teaching* 25, 254-259.
- Oller, J. W. Jr. 1972. 'Scoring methods and difficulty levels for cloze tests of proficiency in ESL'. *Modern Language Journal* 56, 151-158.
- Oller, J. W. Jr. 1976. 'Language testing'. In Wardhaugh, R. & Brown, H. D. (eds.), *Survey of Applied Linguistics*. Ann Arbor: University of Michigan Press, 275-300.
- Oller, J. W. Jr. 1979. *Language Tests at School: A Pragmatic Approach*. London: Longman.
- Oller, J. W. Jr. and Conrad, C. 1971. 'The cloze procedure and ESL proficiency'. *Language Learning* 21, 183-196.
- Oller, J. W. Jr. and Streiff, V. 1975. 'Dictation: A test of grammar based on expectancies'. *English Language Teaching* 30, 25-36.
- Spolsky, B. 1978. 'Introduction: Linguistics and language testers'. In *Advances in Language Testing: Series 2, Approaches to Language Testing*. Arlington, Virginia: Center for Applied Linguistics, v-x.
- Upshur, J. A. and Homburg, T. J. 1983. 'Some relations among language tests as successive ability levels'. In Oller, J. W. Jr. (ed.), *Issues in Language Testing Research*, Rowley, MA:

Newbury House Publishers, 188–202.

Wylie, E., Ingram, D. E., and Hudson, C. 1995. *Australian Second Language Proficiency Ratings (ASLPR)*. National Languages and Literacy Institute of Australian Ltd.