



2016年8月6日 VELC研究会 第5回研究会

VELC Test の概要と妥当性について

水本篤（関西大学）

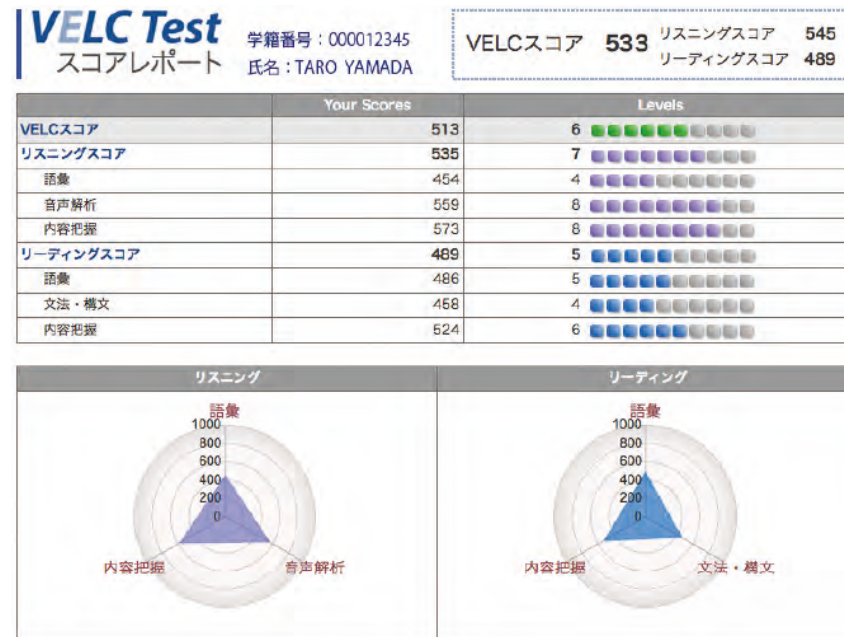
熊澤孝昭（関東学院大学）

本発表の内容について

- VELC Testの概要について
- テストの妥当性とは
 - 従来のテスト妥当性の概念
 - 近年の妥当性の概念について
- VELC Testの妥当性について

VELC Test とは

Visualizing English Language Competency
日本人大学生の英語力を可視化する



VELC Test とは

- 2010年より静先生、望月先生中心に開発
- 2013年より公式受験を大学生対象に行っている
- 特定のコースに依拠しない熟達度テスト
- 主にプレイスメントテストとして使われている
 - 日本人大学生に合った難易度・内容である
 - 迅速に結果が通知される（翌日中）， 安価（800円/1人）
 - TOEICスコア予想などがeポートフォリオ上で詳しくフィードバックされる
 - プレイスメントテストとして妥当性・信頼性が検証されている（静・望月, 2014）

VELC Testの構成

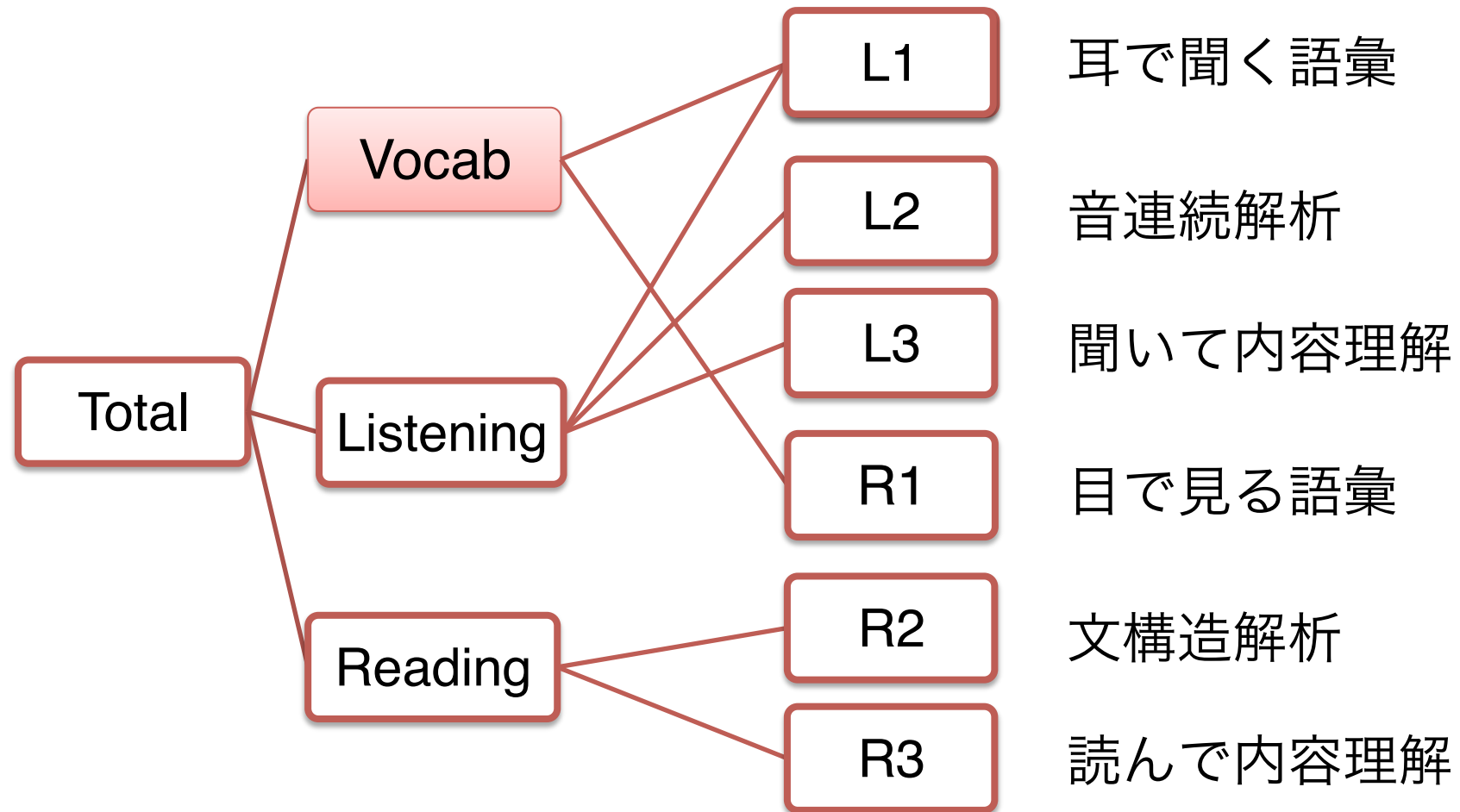
セクション	出題意図	問題例（青字は音声）	問題数
リスニング パート1	聞いて語彙を理解できるかを測る	歩く a) write, b) happen, c) walk, d) love	20項目
リスニング パート2	連続した音声を正しく単語単位で聞き取れるかを測る	If fact, () () () (*) () () () . a) all, b) though, c) must, d) almost In fact, my sister is almost as tall as me.	20項目
リスニング パート3	聴解力をもとに次にくる単語を予測する力を測る	In Japan today, more than 50% of high school students go on to ??? a) school, b) university, c) books, d) jobs	20項目
リーディング パート1	高校・大学レベルの語彙を知っているかを測る	経験 a) society, b) experience, c) notice, d) language	20項目
リーディング パート2	文法力および読解力を測る	Today, people a) can use the Internet b) find it easy to c) communicate with d) each other. [who]はa)～d)のうちどこに入るか	20項目
リーディング パート3	行間を読むことを含む読解力を測る	In Japan, high school students often _____, but university students usually do not. a) eat lunch, b) wear uniforms, c) work part-time, d) study English	20項目

VELC Testの特徴

- 英語を構成する小さな要素（ひとつの単語，ひとつの音声の変化，ひとつの文法項目など）の理解不足がコミュニケーションの失敗を引き起こすため，それらの要素を測定しフィードバックする。
- （正解が本文の一部となっている）問題形式がシンプルであるがゆえの信頼性の高い測定。測定したい構成概念をより正確に測定するつまり，妥当性が高い。

VELCスコアの階層性

(静・望月, 2014)



従来のテスト妥当性の概念

従来型のテスト妥当性の定義

- Validity is defined as the degree to which a test measures what it claims or purports to be measuring (Chapelle, 2012)
- テストが測ろうとしている概念を測っているか

従来型のテスト妥当性の観点

- 目標基準関連妥当性 (criterion-related validity) : すでに妥当性が高いと思われるテストと妥当化をするテストとの相関関係を算出する (妥当性係数)
- 内容的妥当性 (content validity) : 測りたい概念を項目が測っていると思われるか専門家が判断する
 - 到達度テストの内容が授業の内容を忠実に反映している場合、そのテストは内容妥当性があるといえる
- 構成概念妥当性 (construct validity) : 相関関係や因子分析などを用いて統計的にテストが測っている概念を検証する

従来型のテスト妥当性の問題点

- 単純すぎる
 - いずれかの観点について言及するのみでよい。つまり、妥当性係数を提示するのみで妥当性が高いといえた
- テストが妥当なのか
 - テストはただの紙であって、テスト自体が妥当とはいえない
- テストの影響について
 - テストは何らかの判断をするため使用され、受験者に影響があるが、その影響について考慮されていない

近年の妥当性の概念

近年の妥当性の定義

- 得点は何を表していると解釈するか、そしてその解釈をどのように使うか、この**解釈および使用方法の適切性**に関わる概念を妥当性という（静，2002，p. 31）
- したがって、テスト自体は妥当ではなく、テスト得点の解釈および使用方法が妥当かどうか検証することを**妥当化**という

論証型妥当性枠組みとは

- 英語プレイスメントテストの場合、テスト得点を用いて以下のような解釈をしたうえ判断をする：
 - このプレイスメントテストにある問題は良問ばかりであると解釈する
 - このテスト得点をもとに、この受験者の英語力はレベル1の上級で学ぶに必要な水準に達していると判断する
- このようにテスト得点をいかに解釈し、使用するかを明示して、結果をもとに論拠を提示する妥当化の枠組みを論証型妥当性枠組み（[argument-based validity framework] Kane, 2006）という

得点使用までの手順とそれを結ぶ推論 (Kane, 2006)

- 観測 (observation): プレイスメントテストの実施
 - 得点化 (scoring inference): 観測された英語力を得点化する
- 観測得点 (observed score): 受験者のプレイスメントテストの得点
 - 一般化 (generalization inference): 観測得点をもとにテストを何度か受験しても得点が一貫しているかを一般化する
- 一般化得点 (universe score): 一般化され、推定された得点
 - 外挿 (extrapolation inference): 一般化得点をもとに英語使用が想定される授業での学生の英語力を推測する
- 目標得点 (target score): 外挿された仮想の得点
 - 判断 (decision inference): 目標得点をもとに、ある学生が特定のレベルで学ぶに必要な水準に達しているかを判断する
- 使用 (use): 実際のクラス分けのための得点使用

手順と推論の図



VELCスコアの解釈と使用の妥当性について

VELCスコアの得点解釈と使用の論証

- VELC Testをプレイスメント目的で実施した場合のVELCスコアの解釈と使用の論証（argument）
 - VELC Test項目はプレイスメント用の項目として日本人大学生対象に適切に機能しているため、得点化が正確にできる
 - VELCスコアは一貫性があり、測定誤差がさほどないため、一般化が適切にできる
 - VELCスコアは受験者が英語を用いて何ができるかの指標になるため、外挿が適切にできる
 - VELCスコアはクラス分けの判断材料として適しているため、正しい判断が下せる

得点化についての論拠

- プレイスメントテストなのでVELCスコアはクラス分けの判断材料として用いられる。したがって、VELC Test項目が英語力が高いまたは低い受験者のVELCスコアを弁別していると機能しているといえる
- 項目弁別力：上位・下位群の正答率を引くことで求められる（.20以上だと、弁別しているといえる）
- VELC Testは120項目（ $k = 120, N = 4407$ ）からなるが、弁別力が低い項目は、リスニングでは17項目、リーディングでは2項目のみであった

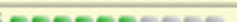
一般化についての論拠

- VELCスコアの測定誤差が少なく、一貫して得点を算出することができれば、一般化を正しくできる

セクション	項目数	平均	標準偏差	信頼性 α 値	標準測定誤差
聴解語彙問題	20	13.37	3.23	.70	1.77
音声解析問題	20	9.87	3.72	.74	1.90
内容把握問題	20	10.27	3.16	.59	2.07
語彙問題	20	15.19	3.49	.81	1.52
文法構文力問題	20	14.04	3.77	.78	1.77
内容把握問題	20	11.51	4.06	.78	1.90
合計	120	72.24	17.00	.93	4.50

外挿についての論拠（１）

- eポートフォリオ：
日本人大学生が
「実際には英語で
何ができるのか」
を言葉で表した状
況別Can Doレベル
を提示したもの

あなたのスキル別正答率(%)			▲ 閉じる
知識・スキル	あなたの正答率	全国標準正答率	
高校レベルの語彙を文脈の助けなしに理解・想起する力	87% 	★ 73	
大学基礎レベルの語彙を文脈の助けなしに理解・想起する力	72% 	★ 59	
大学応用レベルの語彙を文脈の助けなしに理解・想起する力	57% 	45	
はっきり発音される名詞、動詞、形容詞などを聞き取る力	70% 	★ 55	
弱く発音される前置詞、冠詞、助動詞などを聞き取る力	64% 	43	
語尾がはっきり発音されない語を聞き取る力	55% 	★ 47	
あいまいな音で発音される母音を含む名詞、動詞、形容詞などを聞き取る力	80% 	★ 56	
語と語がつながって（リンキング）発音されたものを聞き取る力	56% 	48	
/l/と/r/, /b/と/v/など、日本語にない音の区別を聞き分ける力	80% 	★ 49	
比較的長めの主語・主部を正しく理解する力	58% 	★ 51	
比較的長めの目的語を正しく理解する能力	59% 	★ 51	
前置詞で始まる長めの表現を理解する力	52% 	★ 53	
時や条件、目的などを表す表現（副詞節）を理解する力	64% 	★ 52	
関係代名詞・関係副詞を含む文を理解する力	65% 	53	
名詞を後ろから修飾・説明する構造を理解する能力	62% 	★ 54	
代名詞が指すものを理解する力	45% 	★ 52	
and, or で結ばれて長くなる表現を理解する力	59% 	★ 52	
文の途中に、語句が挿入されている構造を理解する力	64% 	★ 53	
関係詞の非制限用法・分詞構文などを理解する力	60% 	★ 56	
文と文の関係を理解する力	47% 	★ 52	

長めの前置詞句（前置詞で始まるフレーズ）に慣れましょう。
 時や条件、目的などを表す表現（副詞節）に慣れましょう。
 and, or が、どこまでの語句を連結しているかに注意を払いましょう。
 文の途中に、語句が挿入されている構造に慣れましょう。
 関係詞の非制限用法・分詞構文などで、文末に情報を付け加える構造に慣れましょう。
 文と文の関係をしめすつなぎことばに注意を払いながら、論理の展開を把握する努力をしましょう。

外挿についての論拠（2）

- VELCスコアとCan Do項目の難易度を比較することで、英語を用いてなにができるかの指標にすることができれば、外挿が正しくできる

VELC得点	項目
340	中学レベルの教材を読み上げた録音を聞いて...
500	高校レベルの教材を読み上げた録音を聞いて...
380	日本人の先生が授業中に簡単な英語で指示するのを聞いて...
560	ネイティブスピーカーの先生の英語の授業を受けて....
615	海外旅行中に食事・買い物のために店に入り、あなたの言ったことに店員さんが答えるのを聞いて....
595	テンポのゆっくりした英語の歌（バラードなど）を聞いて....
610	海外旅行中に空港や駅などで出発便やプラット英語のアナウンスを聞いて...
750	CNNなどの海外で作成されたニュース番組を見て....
750	英語の映画を字幕なしで見て....
810	ネイティブスピーカー同士が自然に話しているのを聞いて....
495	海外の公共交通機関の駅での英語の掲示（切符の買い方・値段など）を見て
350	中学レベルの教材として書かれた英文を読んで...
455	高校レベルの教材として書かれた英文を読んで....
545	学習者用に編纂された英英辞典（Longmanなど）の定義文を読んで....
550	海外レストランの英語メニューで料理名と説明を読んで
505	自分あてに書かれた簡単な英語のe-メールの文を読んで
590	日本で出版された英語の新聞で国内ニュースを読んで....
575	学習者用に書きなおされた短い小説を読んで...
670	海外で出版された英語の新聞でニュースを読んで.

判断についての論拠

- 日本の大学では大概の場合、3レベルにカリキュラム編成するので、VELCスコアが各大学の受験生を3レベルに分離することができれば適切なクラス分けの判断に用いることができる

大学	受験者数	平均値	標準偏差	ラッシュ受験者信頼性	ラッシュ受験者分離指数
A	297	85.42	12.26	.88	2.69
B	211	46.38	13.07	.87	2.64
C	228	52.94	14.35	.89	2.78
D	1011	77.93	15.62	.91	3.25
E	272	66.70	11.42	.87	2.55
F	1235	73.34	13.40	.88	2.76
G	1153	80.18	15.25	.92	3.36

まとめ

- VELC Test項目はプレイスメント用の項目として日本人大学生対象に適切に機能しているため、得点化が正確にできる
- VELCスコアは一貫性があり、測定誤差がさほどないため、一般化が適切にできる
- VELCスコアは受験者が英語を用いて何ができるかの指標になるため、外挿が適切にできる
- VELCスコアはクラス分けの判断材料として適しているため、正しい判断が下せる

詳しくは

- Validity argument for the VELC Test® score interpretations and uses. *Language Testing in Asia*, 6(2), 1-18. doi: 10.1186/s40468-015-0023-3

参考文献

- 静哲人. (2002). 「英語テスト作成の達人マニュアル」 大修館書店.
- 静哲人・望月正道. (2014). 「日本人大学生のための標準プレイスメント・テスト開発と妥当性の検証」 *JACET Journal* 58, 121-141.
- Chapelle, C. A. (2012). Conceptions of validity. In G. Fulcher & F. Davidson (Eds.), *Routledge Handbook of Language Testing* (pp. 21-33). New York: Routledge, Taylor & Francis Group.
- Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 17-64). Westport, CT: American Council on Education and Praeger.

謝辞

- 分析データを提供していただいた大学、およびVELC研究会事務局に感謝いたします。
- 静哲人先生・望月正道先生の両先生に、内容についてのご助言をいただきました。お礼を申し上げます。

VELCテストに関するお問い合わせ

ホームページ

<http://www.velctest.org/>

VELC 研究会事務局
株式会社金星堂