

Development of a Test Instrument Measuring the Cognitive Academic Language Proficiency: Focus on Writing Skill in Filipino of College Students

Ester T. Rada, Ph.D.
San Beda College, Manila

This paper discusses the bases of developing a test instrument to measure the cognitive academic language proficiency focusing on the writing skill of Filipino college students. The study is composed of two parts: 1) development of a test instrument and 2) experts' validation of the instrument and administration of the test instrument to freshmen and sophomore students of the College of Arts and Sciences in San Beda College, Manila. The construction of the test instrument was founded on the basic principles stemmed from the following linguistic theories: Cummins' Basic Interpersonal Communicative Skill (BICS) and Cognitive Academic Language Proficiency (CALP), Chamot and O' Malley's Cognitive Academic Language Learning Approach (CALLA), Douglas and McNamara's Language for Specific Purposes, Palma's Curriculum Development System, and Process Writing Approach using analytic scale particularly the modified Diederich scale for the evaluation of the essay part of test. Based on the results of the study, 11 out of 214 respondents got a "good" mark, 201 got "average" mark and two (2) got "weak" mark. No one reached an "excellent" mark nor "poor" mark based on the scoring evaluation criteria made by the researcher. In the linguistic knowledge, most of the respondents found the "correct usage" and "vocabulary" items the most difficult. In the essay part, the respondents were marked "good" in the lower-order-thinking skill and "average" in the higher-order-thinking skill. The results were consistent with statistical tools used: item analysis, degree of difficulty, point biserial correlation and cronbach alpha of .77 (base 50). The items in the correct usage and vocabulary were the most difficult among the respondents. Statistical results though modest still conformed to acceptable standard in statistics especially for a pilot test.

Background and Framework of the Study

Jim Cummins (1979) introduces proficiency levels in language learning, i.e. Cognitive Academic Language Proficiency (CALP) and Basic Interpersonal Communicative Skill (BICS). This delineates differences in the level of language used in academic setting and in personal encounter or conversational language.

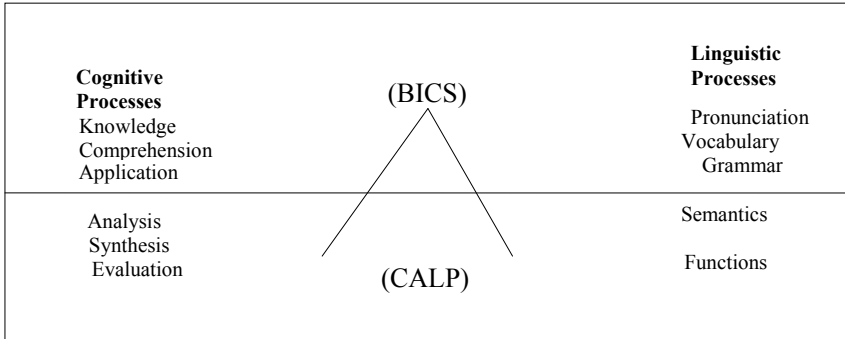


Figure 1 - Lower-Order and Higher-Order Thinking Skill (Marietta Otero, 2000), "Content-Based Instruction for the Development of Cognitive Academic Language Proficiency," *Philippine Journal for Language Teaching* vol. XXXIX, Mayo, 2000, pp. 32-39.)

Here is the diagram that illustrates the differences between CALP and BICS:

To further illustrate cognitive aspect in language learning, other frameworks were used such as Bloom's taxonomy in its cognitive domain, Halliday's Functions of Language and Cognitive Academic Language Learning Approach (CALLA), hence:

<u>Functions of Language</u>	<u>Cognitive Processes</u>		<u>Linguistic Processes</u>
Seek Information	Knowledge	C	Writing Elements
Inform Infer	Comprehension	A	Vocabulary

Relate	Application	L	Grammar
Compare	Analysis	P	Semantics
			Order
Classify		in	
Analyze			
Synthesize	Synthesis		Language Functions
Solve			
Justify	Evaluation	Filipino	
Evaluate			

*Figure 2 – Cognitive Academic Filipino Language Proficiency Levels:
Focus on Writing Skill*

Cummins posits that there are factors affecting language learning: context and cognitive. Academic language is more difficult to learn because it requires cognitive ability than social language. Academic language has specific goals such as to seek information, describe abstracts and cultivate abilities to understand concepts. There is a relation between cognitive and linguistic processes. The linguistic process involves writing elements, vocabulary and grammar and in-depth understanding of semantics and functions of language. Both lower-order and higher-order thinking skills are prerequisites to carry on language functions. For example, lower-order skill such as remembering or recall of concepts learned is needed in order to develop vocabulary skill. On the other hand, higher-order thinking skill requires analysis, synthesis and evaluation. There is a close relation between language functions and cognitive level (Chamot and O'Malley, 40-42).

The flow of the study is shown below. In the development of the instrument, other approaches were also used such as Language for Specific Purposes (Douglas & McNamara, 2000), Table of Specifications for the test, Curriculum Development System (Palma, 1992), Process Writing Approach and Diederich of the Analytic scale in the evaluation of the instrument.

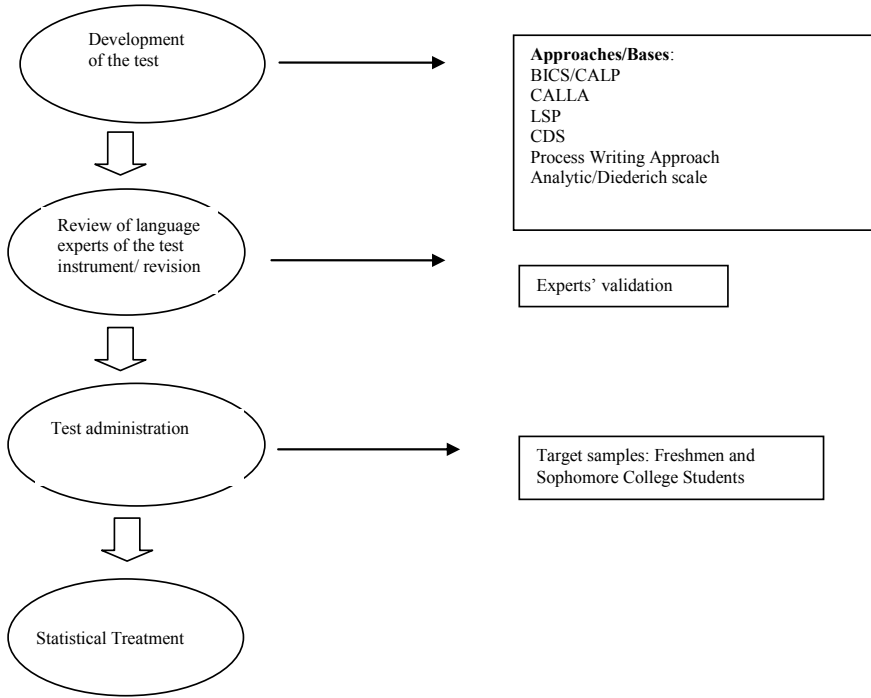


Figure 3 – Flow of the Study

Presentation, Interpretation and Analysis of Data

A. Approaches/Bases used for the development of the instrument

1. **Cognitive Academic Language Proficiency - CALP** – In semantics, a single word may mean differently and different words may mean the same based on contexts. For example, in the Vocabulary item of the test instrument, the test-taker has to analyze which word is the most correct answer among the given choices based on how the word is used in the sentence. Two example test items which the students have difficulty were presented below:

Example 1:**1. Iyon ang tawag sa kanya ng balana.**

- a. sinuman b. masa c. mayayaman d. anuman

1. That was his name by the populace/general public.

- a. anyone b. mass c. rich d. whatever

The right answer is a. sinuman (anyone). The populace/general public may refer to any person. It does not refer only to lower bracket of the population or the mass (masa), nor to the rich (mayayaman) only, and to whatever (anuman) since the word refers to a person/s and not to an object.

Example 2:**9. Ang estudyante ay may kilik na libro.**

- a. kipkip b. nakatago c. bitbit d. dala

9. The student carries a book under his/her armpit.

In Filipino, the word “kilik” may denote differently but based on context a reader can derive the closest meaning of the word. “Kilik”, though, is not a popular Tagalog term. There is no direct translation of this word (in one word) in English. The word “carry” though may mean differently, such as “kipkip” (tuck in), “bitbit” denotes a small thing is being carried with a hand; and carry (dala) which is generic. “Nakatago” means hidden from view. “Kilik” means something is carried under the armpit or something is being held in the hand closer to one’s chest. It may also mean other things.

Items such as these show knowledge in vocabulary, i.e. a certain word has a specific meaning based on context. In the item analysis, only 13% of the respondents got these items in the test right.

Items in Correct Usage, Phonology and Morphology require linguistic knowledge. Recall for grammar rules is a basic communicative skill.

2. ***Cognitive Academic Language Learning Approach*** – CALLA (Chamot at O’ Malley, 1994) - This enumerates a pedagogical model to meet the academic needs of the students in language learning. CALLA program believes that cognitive development of a learner hastens language learning. This is true to all other content area subjects since language is a medium of instruction. This emanates from CALP theory, thus, it also lists down functions of the language, as follows:

Academic Language Function	Student Uses Language to:	Examples	Quality Thinking Skill
Seek Information	observe and explore the environment, acquire information inquire, describe information	Use who, what, when, where, and how to gather informatio	Knowledge (Kaalaman)
Inform	identify, report	Recount information presented by teacher or text, retell a story or personal experience	Comprehension (Pag-unawa)
Infer	make inferences; predict implications; hypothesize,	Describe reasoning process (inductive or deductive) or generate hypothesis to suggest causes or outcomes	Comprehension

Relate	Use of facts, rules, principles	Give an example of in relation to the ideas presented, give the significance of an idea, situation	Application (Paglalapat)
Compare	describe similarities and differences in objects or ideas	Make/explain a graphic organizer to show similarities and contrasts	Analysis (Pag-susuri)
Order	sequence objects, ideas, or events	Describe/make a timeline, continuum, cycle, or narrative sequence	Analysis
Classify	group objects or ideas according to their characteristics	Describe organizing principle(s), explain why A is an example and B is not	Application, Analysis
Analyze	separate whole into parts; identify relationships and patterns	Describe parts, features, or main idea of information presented by teacher or text	Analysis
Solve Problems	define and represent a problem; determine solution	Describe problem-solving procedures; apply to real life problems and describe	Synthesis (Pag-bubuod)
Synthesize	combine or integrate ideas to form a new whole	Tell why A is important and give evidence in support of a positio	Evaluation (Pagtataya)
Evaluate	assess and verify the worth of an object, idea or decision	Identify criteria, explain priorities, indicate reasons for judgment, confirm truth	Evaluation

Figure 4 – CALP, CALLA and Quality Thinking Skills and Sample Techniques

Examples:

Knowledge – Ano ang tinutukoy na dalawang mukha ng syensya sa akda? (What refers to two faces of science in the selection?)

Comprehension – Higit bang nakatutulong ang syensya sa tao kaysa nagiging sanhi ng kapahamakan? Ipaliwanag ang sagot. (Does science pose more benefit to humankind than harm? Explain your answer.)

Application - Sa palagay mo, sampung taon mula ngayon, magdudulot kaya ang patuloy na pag-unlad ng syensya at teknolohiya ng higit na kapakinabangan sa tao o kapinsalaan? Magbigay ng maaaring sitwasyon o dahilan ng iyong mga palagay. (Would you say that, ten years from now, the continual advancement of science and technology bring about progress or doom? Enumerate possible occurrences to prove your hypothesis.)

Analysis - Gumawa ng dayagram o web ng sanhi at bunga ng dalawang mukha ng syensya at ipaliwanag ang ginawang dayagram. (Make a diagram or cause and effect web of the two faces of science. Explain the diagram or web.)

Synthesis - Anong suliranin ang inilahad? May solusyon ba? Pangatwiranan. Iugnay ang paglalahad sa kabutihan at di-kabutihan ng mga argumento sa akda. Ilapat ito sa karanasan, obserbasyon, o pagbabasa at hindi mula sa sariling opinyon lamang. (What is the problem being laid down in the selection? Is a solution presented? Justify your answer. Base your arguments from experiences, observations or readings and not only from your own opinion.)

Evaluation - May magagawa ba ang tao upang maiwasan ang masamang dulot ng syensya? Ipaliwanag ang ibig sabihin ng huling talata. Sang-ayon ka ba rito o hindi? Bakit? (Is there anything any human being can do to avert the ill effects of science? Explain the last paragraph in the selection. Do you subscribe to the idea presented here or not? Why?)

3. ***Curriculum Development System*** – CDS (Palma, 1992) - This specifies the need for Table of Specifications for a test. This includes objectives of each item, skills or concepts being developed, type of items, no. of items and percentage of each item based on objectives. The test instrument in this study aims to measure the cognitive domain in relation to language learning, thus, academic language functions based on CALLA are considered in the writing of each test item. Some aspects were also considered in writing the test items such as: 1) format or user-friendliness – “eye appeal”, spacing and margin; 2) gradation of test items - from easy to difficult 3) clarity, conciseness and comprehensibility of instructions; 4) textual aspect – level of difficulty of words/ vocabulary, type fonts and illustrations, clarity and accuracy of reproduction.
4. ***Language for Specific Purposes (Douglas at McNamara, 2000)*** - This is considered in the writing in each test component, prompt, rubrics and descriptive table of specifications.
5. ***Process Writing Approach (Roe et al., 1998)*** - The analytic scale was used in the evaluation of the essay. There are certain point values in each criterion of the test. The Diederich scale is an example of an analytic scale. The scores are: poor, weak, average, good, and excellent. This measures the following: 1) quality, relevance, flow and organization of ideas; 2) style, elements, individuality and word and phrase choice; and 3) grammar and syntax, punctuations, spelling and legibility.

B. Validation process

1. ***Experts review*** - The experts came from the two prime universities in the Philippines – University of the Philippines and De La Salle University. They are seasoned professors and have been in the academe for more than 40 years. They already held key positions in their respective institutions.

They have been members of Commission on Higher Education Technical Panel. They have been test reviewers for major institutional, academic and civil examinations.

2. Tests Validity and reliability indices

- **validity** – measures how much the students learned based on what was taught by the teacher (Palma, 1992). Test items should correspond to the teaching objectives (Popham, 1971; Douglas, 2000). This is more concerned with the accuracy of the inferences based on the scores obtained in the test. A test should be able to analyze the usefulness of the instrument as variable and predictors of what is being measured (Brown, 1980; Douglas, 2000). There are different types of validity: Criterion-related validity, content validity, construct validity, face validity, and consequential validity (McNamara, 2000).
- **criterion-related validity**, measures individual skills based on certain criterion, e.g., proficiency. Some of the implications of the scores obtained in the test can be attributed to the effectiveness of the teaching strategies and the curriculum (Brown).
- **content validity**, measures learning outcomes, specific objectives and if the test reflects the target learning objectives. This needs experts' validation of the test items (Brown, 1980).
- **construct validity** – based on the series of researches done. It is the gathering of evidences to prove what is being measured (Brown, 1980).
- **face validity** – pertains to the format of the test and overall physical aspects of the instrument if it is user-friendly (McNamara, 2000).
- consequential validity takes stock of the consequences on the test-takers. This can be done if there is a change in the teaching strategies, content and syllabus based on the strengths and weaknesses of the students evident on the scores of the test (McNamara, 2000).
- **reliability** – measures consistency of test objectives based on item analysis. The results of the test should be consistent

when administered to the same group of the same profile even at different times. Therefore, the test is dependable and can still be administered in the future (Palma, 1992; Popham, 1971; McNamara, 2000). There are types of reliability: reliability coefficient and coefficient of equivalence (Brown, 1980).

- The reliability coefficient looks into the scores of two sets of sample population on the same instrument. Physical condition, sampling of the test items and the test-takers that can affect margin of errors should be taken into consideration.
- coefficient of equivalence poses question on whether the test-taker can get the same scores in the test if given repeatedly.
- **objectivity** – takes into consideration all factors that may affect the results of the test such as internal or physical condition of the test taker whether it be health, emotional, psychological, or external which may be the environmental factors like ventilation, noise, lighting, and other things (Palma, 1982).

C. Results of the Study

There are 11 students out of 214 samples who got a Good mark, 5.1% of the population; 201 got Average, 93.92% and 2 (two of the samples got Weak mark, .93%. No one got Poor nor Excellent rating.

On the essay, generally, the students got Good mark on cognitive dimensions: Knowledge (76%), Comprehension (68%), Application (63%); Average on higher-order thinking skill - Analysis (56%), Synthesis (56%) and Evaluation (54%).

On the linguistic component using scantron sheet, students got scores on items, from highest to lowest: Phonology (84.11%), Morphology (72.73%), Vocabulary (56.41%), cloze test (51.32%) and Correct Usage (46.52%).

As to item analysis, degree of difficulty: Item 1 registered .28 – which means 28% of respondents got the item right. The rubrics: 50% means difficult, 51%-84% average, 85% easy. The results show that the students found most of the items difficult:

	Difficult	Average	Easy
A. Vocabulary (Items 1-25)	10	12	3
B. Correct Usage (26-45)	10	9	1
C. Phonology (46-50)			5
D. Morphology (51-65)	5	8	2
E. Cloze test (66-75)	5	4	1

Figure 5 – Percentage of degree of difficulty of test items

Referring to point biserial correlation, there are items that can be deleted or revised on the linguistic component to give more weight on the essay writing part which is the focus of the study. In general, KR-20 is .77 (base 50) which is considered standard and acceptable statistics, more so, in a pilot test.

E. Analysis

Based on the results of the degree of difficulty, the respondents found Correct Usage and Vocabulary the most difficult items in the test. This determines the linguistic knowledge of the students. Most of the interviewees found this assumption true as the students really found their Filipino subject difficult which requires academic language and not just conversational language. Thus, it can be deduced that the level of Filipino language proficiency of the students is conversational (BICS). Their cognitive academic language is on the average.

Therefore, there is a need to require grammar lessons and critical thinking tasks in the syllabus, especially to freshmen, and not just incidental teaching. There should always be scaffolding, corrective measures on the process of teaching grammars and critical thinking, more so, in speaking and writing tasks.

Summary and Conclusions

A. The following are theories, approaches and indices to develop a test instrument measuring the cognitive language proficiency of the students:

- Cognitive Academic Language Proficiency - CALP
- Cognitive Academic Language Learning Approach (CALLA) – Process Writing Approach
- Language for Specific Purposes
- Table of Specifications
- Curriculum Development System
- Validity and reliability indices
 - validity
 - criterion-related validity
 - content validity
 - construct validity
 - face validity
 - consequential validity
 - reliability
 - reliability coefficient
 - coefficient of equivalence
 - objectivity

B. Most of the students found items on grammar and correct usage difficult. Teaching strategies should gear toward the enhancement of the students' knowledge on this aspect. Curriculum should require grammar lessons and corrective measures should be employed on the process.

REFERENCES

- American Council on the Teaching of Foreign Languages (ACTFL)*. (n.d.) New York: White Plains.
- Anderson, L.W. & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. USA: Addison Wesley Longman, Inc.
- Assessment Analytics, Inc. Online Solutions*. Makati City: Ayala Avenue, MM Phils.

- Asia-Pacific Computer Technology Center Inc. (The APC Center).
 Bloom, B. S. Ed. (1956). *Taxonomy of educational objectives: The classification of educational goals*. USA: David McKay Co.
- Brown, F. G. (1980). *Guidelines for test use: A commentary on the standards for educational and psychological tests*. Iowa State University National Council on Measurement in Education.
- Cabuhat, A. S. (2002) *"Ang Filipino bilang subjek sa batayang edukasyon tuon sa interaktiv kurikulum: Tungo sa pagpapaunlad at pagbabago."* Pasig City: Department of Education.
- Center for Educational Measurement. *Admission Test for Colleges and Universities (ACTU)*. (1996). Makati City MM Phils.
- Chamot, A. U. & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Massachusetts: Addison-Wesley Publishing Co.
- Cummins, Jim & Swain, Merrill. (1986). *Bilingualism in education*. London: Longman Group Limited.
- Douglas, D. (2000). *Assessing languages for specific purposes*. United Kingdom: Cambridge University Press.
- Douglas, D. & Chapelle, C. Eds. (1993) *A new decade of language testing research: Selected papers for the 1990 language testing research colloquium. Teachers of English to Speakers of other Languages, (TESOL) Inc.* USA: Alexandria, Virginia, 98-122.
- Maminta, R. E. (2001). *In Focus; Selected writings in applied linguistics. The Philippine Association for Language Teaching, Inc.*, 11-35, 69-119.
- McNamara, T. (2000). *Language Testing*. Oxford University Press.
- Montañano, R. L. (1993). *"Higher-order cognitive skills in Filipino towards measurable criteria for describing Cummin's CALP"* PhD diss. Manila: De La Salle University.
- Palma, J. C. (1992). *Curriculum development system: A handbook for school practitioners in Basic Education*. Mandaluyong City: National Book Store.
- Popham, W. J. Ed. (1971). *Criterion-referenced measurement*. New Jersey: Englewood Cliff Educational Technical Publications.
- Roe, B. D. et. al. (1998). *The content areas*. New York: Houghton, Mifflin Co. <http://www.iteachlearn.com/cummins/bicscalp.html>. (wikipedia, ^ TOEFL® iBT Locations and Dates).