Using Comparisons within Language Examination Results to Validate Language Teaching Assumptions

Béatrice Boufoy-Bastick

Abstract. Assessing language learning outcomes is an essential grading requirement that can be used to monitor students' progress as well as to evaluate teaching effectiveness and curriculum quality (Gallavara et al., 2008, pp.40-41; OCDE, 2007, 2010). To satisfy the requirements of course assessment, language instructors use an array of language testing tools to methodically assign summative grades to assess their students' language proficiency. This research reports on a Year 1 Semester 1 French language course programme for 2009-2010 enrolling 33 students. The course used a set of six sub-tests which were administered as Test 1 in week 6. It also used a second parallel set of six sub-tests which were given at the end of the course as Test 2. Students' grades were determined by summing their performances on all 12 sub-tests. This represented considerable teacher effort and resulted in only one grade for each student. To add value to the assessment effort expended by language instructors this paper shows how to also use the students' results for course evaluations. The paper describes four language course assumptions, the pedagogic theory and advantages of each assumption, how these assumptions were implemented and how the students' results were re-used to evaluate how successful or unsuccessful were the assumptions. The pedagogic theory is significance in that it helps to resolve a major conflict in the language learning literature and introduces the Neuro-linguistic Learning Model (NLM).

Key words: *language proficiency; language testing; quality assurance; course evaluation; Neuro-linguistic Learning Model.*

Introduction

Student cohorts in the Humanities, often in excess of 200 students at a time, are routinely and efficiently examined in one sitting using low resource-intensive machine marked objective multiple choice questions. In contrast, language teachers invest considerable effort in assessing the languages competencies of their students, one student at a time. The rigorous assessment of written and oral components of language competence and of students' cultural proficiencies in language communication require significantly more human resources, including an extended amount of time for at least two examiners to attend to the consecutive language productions of individual students. Yet these considerable investments of human resources in the assessment of each individual language student result in a summative grade of the same final utility as the low resource computer generated objective assessments from simultaneous examinations of large cohorts of non-language students in the Humanities.

The problem being researched here is how language teachers can utilise their assessments to add value that is more commensurate with the relatively high effort they have invested in their assessment processes. The solution proposed in this paper is to use valuable information contained in the assessment results of separate course assignments, prior to their aggregation as a final grade, in order to evaluate teaching effectiveness and curriculum quality for the purpose of quality assessment and course improvement. The corresponding aim of this paper is to demonstrate how language teachers can re-value the efforts they have expended on assessing their students by re-using student marks on component assignments to also evaluate teaching effectiveness and curriculum quality for the purpose of quality assessment and course improvement.

Statistical analysis of summative grades is considered to produce objective evidence to support systemic change and this is one reason it is so commonly used by government consultants in large-scale educational evaluations. The scientific position of the author is that these objective methods of evaluation can also be applied to the benefit of language courses at the level of the individual language instructor. In particular the author's position is that curriculum intentions of language courses and pedagogic intentions of language teachers should, where possible, be made explicit and objectively and empirically evaluated for effectiveness and then modified accordingly as indicated by objective evidence from their empirical evaluation. The literature on large-scale education evaluation is replete with such statistical use of summative assessments promoting systemic change (Abedi & Gándara, 2006; Baldauf Jr & Kaplan, 2005; Bayley & King, 2003; Coleman, Galaczi & Astruc, 2007; Hamp-Lyons, 2009; Liddicoat et al., 2003; Lorenzo, Casal & Moore, 2010; Romero-Little et al., 2007).

In contrast to large-scale assessments of final grades by external bodies, it is more novel to use the same evaluative processes on the results of component assignments that are only available to the classroom teacher. This paper is particularly novel in promoting teachers of language, as opposed to teachers of more quantitative subjects, to use the results of their component assignments in this way. The four-component structure used here to explain and

demonstrate how language teachers can apply these methods is unique to this paper. These four components are: (a) a description of one language course assumption that has been chosen for validation, (b) the pedagogic theory and advantages and disadvantages of that assumption, (c) how that assumption is implemented, and (d) how students' results on component assignments are re-used to statistically and objectively evaluate and report the validity of the course assumption and/or the success, or otherwise, of its implementation. The combined use of these four components explicitly shows how the theory results in statistically testable hypotheses and objective results.

Four demonstrations of this structure are given, each being applicable to the evaluation of a different chosen language course assumption. Some of these course assumptions might be novel to other language teachers who could then also consider the benefits of their adoption. The four language course assumptions whose evaluations are demonstrated are:

1. *Standard constancy:* the assumption that the course promotes a consistent standard of language proficiency against which students may judge themselves;
2. *Test validity:* the assumption that course assessments are measuring the intended language competences;
3. *Course structure:* the assumption that the course structure builds all four language skills in unison;
4. *Pedagogic intentions:* the assumption that the teaching emphasis promotes oral communicative skills.

This paper makes two further novel contributions. Firstly, in the exposition of pedagogic theory, and the advantages and disadvantages of course structure – cross-category 3b above – the paper suggests a resolution to the debate on co-dependence v independence of language skill development. Secondly, a significant novel contribution of this paper is reported in the implementation of course pedagogic intentions to promote oral communicative skills – cross-category 4c above - in that this is the first published account of the Neuro-linguistic Learning Model (NLM) and its use of component assignment marks in evaluating the success of teaching to emphasise oral communication skills of language students.

This paper demonstrates ways in which language instructors can efficiently re-use their students' assignment results for their own course evaluations and for contributing to Quality Assurance requirements. As an illustrative example, the paper presents results of an outcomes evaluation conducted by the language instructor in the first semester of a first year undergraduate French language course during 2009-2010. The instructor re-used the students' results from two parallel sets of language proficiency tests to also evaluate four fundamental assumptions of the French Language programme; namely, constancy of standards, test validity, course structure and pedagogic assumptions. This paper gives detailed examples, with simple explanations of the statistics, showing how other language educators can also add value by doing this.

The summative assessment for the course consisted of Test 1 given in week four to 33 students and Test 2 given in week 12, at the end of the course, to 32 students – one student had left the course. Test 1 was composed of six language sub-tests, detailed in Table 1 below, and Test 2 was composed of a parallel set of six sub-tests that tested the same language abilities as in Test 1. Comparisons between the results of these sub-tests were used to evaluate four fundamental course assumptions. This was done by predicting a pattern in the results that would occur if each assumption had been successfully applied and then using statistical methods to see if the pattern was there in the students' results.

The four assumptions are first described, with the theory behind them, and how the teaching was modified and resources selectively used to put these assumptions into practice. The pattern in the sub-tests that would validate each assumption was identified and then appropriate statistical tests were used to determine if the patterns were present in the students' results so that the success, or otherwise, of the course assumptions could be inferred and reported.

Four language teaching assumptions for validation

Standard constancy. Each skill standard required by the end of the course is set at the beginning of the semester. It is presented as an unmoving target at which students should aim. Each skill tested, with the exception of grammar, is graded to this constant criterion standard. The teaching does not start with lower expectations and easier work. Assessment does not 'grade to the curve' or award marks for effort, although there are motivational reasons why these moving standards could be used (Guskey & Bailey, 2001; Sadler, 2005).

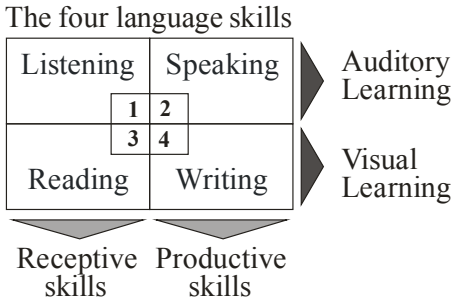
In contrast to the motivational uses of moving standards, students are thus assumed to understand the effort that is expected of them and be able to judge their expected improvement from their test results. Hence, tests of similar difficulty are used for Test 1 and Test 2, which comes after a further 8 weeks of learning. Consequently it is expected that marks will improve over the semester to indicate improved attainments. For Grammar, however, the standard is cumulative, as new grammatical rules are introduced throughout the semester (Purpura, 2004, p.100, 146). To evaluate teaching to these standard assumptions, the pattern in the students' results to be tested was therefore an improvement in marks from Test 1 to Test 2 on all sub-tests, except for sub-tests of grammar.

Test validity. It was assumed that the language tests used by the course assessments were valid. That is, they tested what they were intended to test (Brown, 2004; Fulcher & Davidson, 2007; McNamara, 2000; McNamara & Roever, 2006). In particular, reading comprehension skills were assessed through two reading comprehensions sub-sets in Test 1 and again in Test 2. Hence, as both sub-sets were assumed to test the same ability of 'reading comprehension', the pattern in the results to be expected is that the mark each student received on one sub-test of 'reading comprehension' should match the mark received on the other sub-test of 'reading comprehension'.

Course structure. The course is structured to teach and develop the four language skills in concert. This is achieved through a teaching emphasis on 'language immersion' and French is the language of instruction for all components, including grammar. There is some pedagogic disagreement on this in the language learning literature. For example, it is recognised that each language skill requires a specific ability which can develop independently of the others hence each skill is taught and assessed independently of the others (Alderson, 2000; Brown, 2004; Buck, 2001; Fulcher & Davidson, 2007; Luoma, 2004; Purpura, 2004; Woods, 2006). Yet, being interdependent, language skills intimately build on one another using common underlying proficiencies (Cummins, 1991a, 1991b; 2000; Genesee & Jared, 2008) and are intricately intermingled.

The approach used by the course of developing the four language skills in concert is made effective by using small learning blocks of separate skills to develop the four language skills together, and in doing so it operationalises the language pedagogy for oral communication skills emphasised by the teaching. The author's theory, which goes some way to resolving the above conflict is that the sequential sensory dependence of language skills - e.g. that one must first hear vocabulary before one can voice it, or first read vocabulary before one can write it - interacts with language learning through the concept of 'block learning' and the size of the blocks learnt. Obviously, it would hinder the integrative learning of all four skills if large sequential block learning were mandated, e.g. if a one year course were to be subdivided into four immense blocks matching the sensory sequence of language skills dependency such that vocabulary was only heard in the first 3 months, followed by only speaking that vocabulary during the next 3 months, followed by only reading that vocabulary in the next 3 months and only writing it in the last 3 months. The premise of interdependent language skills teaching is, in reality, based on using consecutive sequences of the smallest sensory dependent language skill blocks possible for the purpose of learning, e.g. phonemics or morphemes (McShane, 2005), speech acts (Fujimori & Houck, 2004) or idioms (Erman, 2007). Using this pedagogy it was assumed the teaching was developing all four abilities 'simultaneously'. This assumption was evaluated by checking for the expected pattern in the results that all sub-tests of language skills contributed positively to students' summative assessment of language proficiency.

Pedagogic intentions. The teaching is consistent with current global emphases on communicative language objectives. It recognises that the first year students need essential communication language tools for engaging in the substantially more academic content of the French undergraduate programme in the subsequent two years of the course. In practice, this is done by placing a stronger emphasis on the development of combined listening and speaking skills in the first year of the undergraduate language programme. Emphases on these communicative skills is operationalised by teaching to the Neuro-linguistic Learning Model (NLM) shown in Picture 1.



Picture 1. Neuro-linguistic Learning Model (NLM)

The author's Neuro-linguistic Learning Model (NLM) is neurological extension of the common differential receptive and productive skills model traditionally used in language learning (Davies, 1980; Strevens, 1988, 2007). In practice, the NLM is applied by devoting more class-time, computer resources and practice facilities to the combined Auditory communicative language skills of Listening and Speaking than to the combined Visual language skills of Reading and Writing, in the ratio of four-to-one. That is the teaching devoted four times the amount of course resources to combined Auditory language skills than to combined Visual language skills. This was assumed to be a superior pedagogy for teaching communicative language skills because it avoided the cognitively more difficult and ontologically later developed cross-sensory language representations necessary to the traditional teaching emphasis on developing receptive and productive language skills. It is a principle of assessment that assessment output parallels learning input (Brown, 2004; Council of Europe, 2001; Genesee & Upshur, 1996; Woods, 2006). Hence it would be expected that the output pattern of students learning results would parallel this input emphasis in the teaching. That is, the results should show more learning in the combined Auditory skills of Speaking and Listening than in the combined Visual learning skills of Reading and Writing, in the same ratio as the teaching of 4 to 1.

Methodology

Students' summative examination results on the sub-test component assignments were re-used to evaluate the four course assumptions.

Test materials and administration. The course examination consisted of two parallel tests, each of six language ability sub-tests. These were university course requirements and served as the official summative assessments for the course. The first test T1 was given in week 6 and the parallel test T2 was given in week 12. This is parallel test structure illustrated in Table 1.

Subjects. Thirty-three students enrolled in the first semester of a first year undergraduate French language course sat Test 1 and thirty-two of these students sat Test 2. The student results for the 6 sub-tests of Test 1 and the student results for the 6 sub-tests of Test 2 were re-used for this course evaluation.

Research design and analyses. The four assumptions were tested by comparing results from the 12 sub-tests as follows.

Table 1. The course examination comprised six language sub-tests given in week 4 and in week 12, at the end of the course

#	Assignment	Description	Language Ability Assessed	Two Parallel tests Test 1	Test 2
1	ORAL	Oral	Speaking	T1 #1	T2 #1
2	Anal	Analyse littéraire	Reading comprehension of literary texts	T1 #2	T2 #2
3	CE	Compréhension écrite	Reading comprehension of articles	T1 #3	T2 #3
4	CA	Compréhension auditive	Listening comprehension	T1 #4	T2 #4
5	EE	Expression écrite	Written expression	T1 #5	T2 #5
6	GRAMM	Grammaire	Grammar	T1 #6	T2 #6
# sub-test number				*Week 6*	*Week 12*

Standard constancy. The final standard is maintained throughout the course for abilities #1 to #5, *Oral, Analyse littéraire, Compréhension écrite, Compréhension auditive* and *Expression écrite*. If Test 1 and Test 2 were marked to the same standard, then because students were taught for 12 weeks for Test 2 and only for 6 weeks for Test 1, it is expected that students would score higher in Test 2 than in Test 1. That is, the marks for sub-tests T2#1 to T2#5 should be higher than the marks for sub-tests T1#1 to T1#5. The standard for *Grammaire*, sub-test #6, was assumed to be cumulative so the mean for T2#6 would not be expected to be higher than the mean for T1#6. This was tested for all language six abilities by comparing the mean scores for each T1 sub-test with the mean scores for each sub-test on the parallel Test 2. Null hypotheses were that mean scores on each of the first 6 sub-tests were the same on T1 and T2. Independent samples t-test was used for *Expression écrite* because it was marked as group assignments. Paired t-tests were used to compare the other sub-tests because they were individually marked assignments. The percentage improvements were graphed.

Test validity. Although the two sub-tests of reading comprehension used different genre, which might have produced different mean results, they were assumed to both measure reading comprehension. This concurrent validity was tested by comparing what each student scored on T1#2 with his/her score on T1#3 using correlation. The null hypothesis is that results from the two tests of comprehension are uncorrelated. This was tested for the comprehension sub-tests on Test 1 and replicated for the comprehension sub-tests on T2.

Course structure. Time and materials were allotted to all language abilities in concert (rather than developing them sequentially) so that all language abilities were expected to contribute positively to students' overall language proficiency. This was tested by checking that the sub-test scores for each student did contribute positively to his/her total score of language proficiency. The Cronbach-alpha statistic was used for this test and it was expected that C-alpha would be greater than 0.741 corresponding to an effect size of 50 %.

Pedagogic intentions. Teaching emphasised oral communication giving priority to learning via the Auditory sensory systems over learning via the Visual systems in the ratio of 4:1. The testing outcomes should match the teaching inputs, so this expectation was used to test the Pedagogic assumptions by factor analysis of all 12 sub-tests. It was expected that varimax rotation of factors whose Eigen values were greater than one would identify a main auditory sensory factor and a secondary Visual sensory factor accounting for total language proficiency in the ratio of 4:1 rather than showing any Receptive and/or Productive skills factors.

Analysis and results

Standard constancy. Table 2 shows the expected positive increase in abilities against the constant standard set throughout the course, with a mean percent improvement of 18 %. Figure 1 compares the mean sub-test results for Test 1 with Test 2. As expected, the only decrease in Test 2 results was for Grammar. The standard expected from the students was higher on Test 2 than on Test 1 because Grammar was taught as cumulative skill.

Table 2. Significance of increases in test results against the constant standard and significant decrease in Grammar results against the higher standard

#	Test	T1 Means	T2 Means	Sig	Improvement
1	ORAL	8.64	10.08	0.031	17%
2	Anal	3.18	5.84	0.000	84%
3	CE	4.87	5.64	0.067	16%
4	CA	6.76	7.23	0.234	7%
5	EE	10.50	11.27	0.390	7%
Totals		**33.94**	**40.06**		**18%**
6	*GRAM*	*10.77*	*9.64*	*0.005*	*-11%*

The unexpectedly high improvement of 84 % in *Analyse littéraire* was investigated and found to be due to an unplanned punitive marking scheme for T1#2.

Test validity. The correlations between the Comprehension sub-tests on Test 1 and Test 2 were positive as expected at r=0.450 (n=32, p=0.010) and r=0.329 (n=32, p=0.066) respectively.

Course structure. It was assumed that all abilities would contribute in concert to students' French Language Proficiency. This was tested using C-alpha to ensure that the result for each language ability contributed to the reliability of students' overall assessment of Language Proficiency.

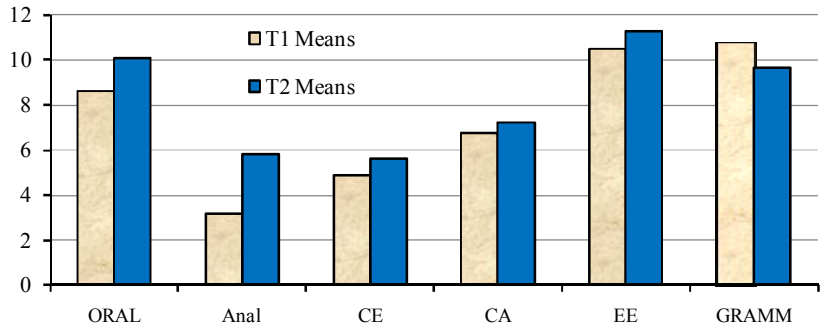


Figure1: Expected increases in test results and expected decrease in Grammar results

Table 3. Very high reliability of students' language proficiency assessment

Reliability Statistics

Cronbach's Alpha Based on Standardized Items	N of Items
0.896	12

Table 4. Contribution of sub-test assessments to the reliability of total assessment

Item-Total Statistics

Sub-test	Corrected Item-Total Correlation	Cronbach's Alpha if Sub-test Deleted
T1ORAL	0.653	0.871
T1Anal	0.469	0.881
T1CE	0.642	0.877
T1CA	0.688	0.868
T1EE	0.395	0.884
T1GRAMM	0.718	0.866
T2ORAL	0.582	0.882
T2Anal	0.676	0.875
T2CE	0.432	0.882
T2CA	0.787	0.860
T2EE	0.581	0.875
T2GRAMM	0.727	0.865

Table 4 shows that the sub-test assessments contribute positively and optimally to the reliability of the overall assessment of language proficiency in that if any assignment had NOT been included then the overall reliability of the aggregated assessments would have been reduced from a C-alpha of 0.896 as shown in Table 3 to the C-alpha value listed in Table 4 opposite whichever sub-test was deleted.

Pedagogic intentions. The scree plot in Figure 2 shows that the main two factors account for most of the variance in the students' results. They account for 58 % of the variance. Table 5 confirms this expected two-factor structure of students' total examination results for language proficiency.

The sub-tests whose sorted factor loadings most define the first factor are listening and speaking tests accounting for 48 % of the variance in the results. This is the expected Auditory factor accounting for 48 % of the variance in learning. The sub-tests whose sorted factor loadings most define the second factor are reading and writing tests. This is the expected Visual factor accounting for 10 % of the variance in learning. The relative emphasis in achieved learning of 48 % for Auditory language skills and 10 % for Visual language skills is in the ratio of 48 %:10 % and so confirmed the assumption by matching the 4:1 emphasis in teaching. The last two sub-tests in Table 5, T1Anal and T1EE, were identified as contributing very little to these sensory learning systems. On further enquiry the reasons for this were found to lie in the marking of these two assignments.

Table 5. Auditory and Visual factor structure of students' language proficiency

Varimax Rotated Factor Matrix

	Factor		
	1	**2**	**3**
T1CA	.789	.358	-.054
T2ORAL	.789	.053	.176
T2Anal	.725	.214	.242
T1ORAL	.663	.090	.477
T2CA	.654	.387	.383
T1CE	.538	.457	.126
T2CE	-.008	.777	.178
T2GRAMM	.400	.700	.329
T2EE	.490	.651	-.152
T1GRAMM	.381	.629	.391
T1Anal	.157	.118	.861
T1EE	.127	.171	.693
Eigen value	**5.75**	**1.26**	**1.03**
% variance	**48%**	**10%**	**9%**
Cum var	48%	**58%**	67%

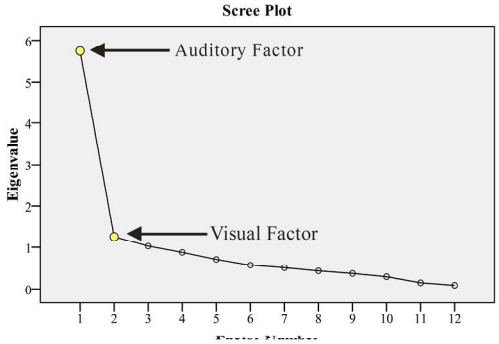


Figure 2. Scree plot for Factor structure of language Proficiency showing the expected two main NLM factors

Conclusions and Discussion

This paper has demonstrated and explained how language teachers can add value to the relatively large effort they expend in assessing students' language proficiencies by 'recycling' the same sub-test assignment results that they are obliged to produce for the summative assessment of their courses to also evaluate teaching effectiveness and curriculum quality for the purpose of quality assessment and course improvement. The paper explained and demonstrated these outcome evaluations using the language proficiency results from a first-year course of an undergraduate French language programme.

In explaining the pedagogic assumptions of the programme the paper has offered some resolution to the dependent/independent language skills controversy and introduced the Neuro-linguistic Learning Model (NLM) and its application to evaluating teaching of oral communication skills which can also be of benefit to the language teaching community.

The evaluations demonstrated tests of (i) standard constancy, (ii) test validity assumptions, (iii) course structure to develop language skills in concert and (iv) pedagogic assumptions that course teaching intentionally developed primarily students' auditory language skills privileging verbal communication, and developed secondly students' visual language skills of written communication. Where the analyses did not validate course assumptions, the results were used to identify parts of the course that had not progressed as planned and the problems were investigated so that future teaching efforts will result in the optimum learning outcomes intended for the students. This language course outcomes evaluation was included in a university quality assurance report (DLA, 2010) on the four-year French language programme of which the international external examiners wrote "…we applaud the French Section's clear formulation of the aims, objectives and outcomes of its Programme" (Quality Assurance Unit/OBUS, 2010, pp.19-20).

References

1. Abedi, J., Gándara, P., 2006. Performance of English Language Learners as a Subgroup in Large-Scale Assessment: Interaction of Research and Policy. *Educational Measurement: Issues & Practice*, 25(4), pp.36-46.
2. Alderson, J. C., 2000. *Assessing Reading*. Cambridge: Cambridge University Press.
3. Baldauf Jr, R. B., Kaplan, R. B., 2005. Language-in-Education Policy and Planning. In: E. Hinkel, ed. *Handbook of Research in Second Language Teaching and Learning*. Mahwah, N.J.: L. Erlbaum Associates, pp.1013-1034.
4. Bayley, R., King, R., 2003. Languages Other than English in Canada and the United States. *American Speech*, 88, pp.163-229.
5. Brown, D. H., 2004. *Language Assessment. Principles and Classroom Practice.* White Plains, NY: Pearson Education.
6. Buck, G., 2001. *Assessing Listening.* Cambridge: Cambridge University Press.
7. Coleman, J. A., Galaczi, Á., Astruc, L., 2007. Motivation of UK School Pupils Towards Foreign Languages: a Large-scale Survey at Key Stage 3. *Language Learning Journal*, 35(2), pp.245-281.
8. Council of Europe, 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.
9. Cummins, J., 1991a. Language Development and Academic Learning. In: L. M. Malave & G. Duquette, eds. *Language, Culture and Cognition.* Philadelphia: Multilingual Matters.
10. Cummins, J., 1991b. Interdependence of First- and Second Language Proficiency in Bilingual Children. In: E. Bialystok, ed. *Language Processing in Bilingual Children.* Cambridge: Cambridge University Press.
11. Davies, N. F., 1980. Putting Receptive Skills First. *Canadian Modern Language Review,* 36, pp.461–67.
12. DLA, 2010. *2010 Outcomes Evaluation Report for FREN1401*. Department of Liberal Arts, French section, University of the West Indies, Trinidad & Tobago.
13. Erman, B., 2007. Cognitive Processes as Evidence of the Idiom Principle. *International Journal of Corpus Linguistics,* 12, pp.25-53.
14. Fujimori, J., Houck, N., 2004. Practical Criteria for Teaching Speech Acts. *The Language Teacher,* 28(5), pp.3-8.
15. Fulcher, G., Davidson, F., 2007. *Language Testing and Assessment.* Abingdon: Routledge.
16. Gallavara, G., Hreinsson, E., Kajaste, M., Lindesj Ö, E., Sølvhjelm, C., Sørskån, A.K., & Zadeh, S. M., 2008. Learning Outcomes: Common Framework – Different Approaches to Evaluating Learning Outcomes in the Nordic Countries. *Occasional Paper 15.* Joint Nordic project 2007-2008 Nordic Quality Assurance Network for Higher Education (NOQA).
17. Genesee, F., Jared, D., 2008. Literacy Development in Early French Immersion Programs. *Canadian Psychologist*, 49(2), pp.140-147.
18. Genesee, F., Upshur, J., 1996. *Classroom-based Evaluation in Second Language Education.* Cambridge: Cambridge University Press.
19. Guskey, T. R. & Bailey, J. M., 2001. *Developing Grading and Reporting Systems for Student Learning.* Thousand Oaks, CA: Corwin Press Inc.
20. Hamp-Lyons, L, 2009. Principles for Large-Scale Classroom-Based Teacher Assessment of English Learners' Language: An Initial Framework From School-Based Assessment in Hong Kong. *TESOL Quarterly*, 43(3), pp.524-530.
21. Liddicoat, A. J., Papademetre, L., Scarino, A., & Kohler, M., 2003. *Government Report on Intercultural Language Learning.* Research Centre for Languages and Cultures Education at the University of South Australia and the School of Languages and Linguistics at Griffith University, Australia.
22. Lorenzo, F., Casal, S., Moore, P., 2010. The Effects of Content and Language Integrated Learning in European Education: Key Findings from the Andalusian Bilingual Sections Evaluation Project. *Applied Linguistics*, 31(3), pp.418-442.
23. Luoma, S., 2004. *Assessing Speaking*. Cambridge: Cambridge University Press.
24. McNamara, T., Roever, C., 2006. *Language Testing: The Social Dimension.* Malden, MA: Blackwell Publishing.
25. McNamara, T., 2000. *Language Testing*. Oxford: Oxford University Press.
26. McShane, S., 2005. *Applying Research in Reading for Adults: First Steps for Teachers.* Washington, DC: National Institute for Literacy.
27. OCDE, 2007. Enseignement supérieur: qualité, équité, efficience. In: *Enseignement et compétences*, 2006 (14), pp.12-81.
28. OCDE, 2010 – Hénard, F., 2010. Monitoring and Measuring Quality Teaching. In: *Learning Our Lesson: Review of Quality Teaching in Higher Education.* Ch. 4.
29. Purpura, J. E., 2004. *Assessing Grammar*. Cambridge: Cambridge University Press.
30. Quality Assurance Unit/OBUS, 2010. *Report of the Review of French.* Faculty of Humanities and Education. University of the West Indies, Trinidad and Tobago.
31. Romero-Little, M. E., McCarty, T. L., Warhol, L., Zepeda, O., 2007. Language Policies in Practice: Preliminary Findings from a Large-Scale National Study of Native American Language Shift. *TESOL Quarterly*, 41(3), pp.607-618.
32. Sadler, R.D., 2005. Interpretations of Criteria-based Assessment and Grading in Higher Education. *Assessment and Evaluation in Higher Education*, 30(2), pp.175-194.
33. Strevens, P., 1988. Learning English Better Through More Effective Teaching: Six Postulates for a Model of Language Learning/Teaching. *World Englishes,* 7(1), pp.51-61.
34. Woods, C., 2006. *Teaching and Assessing Skills in Foreign Languages*. Cambridge: Cambridge University Press.

Béatrice Boufoy-Bastick

Studentų kalbos gebėjimų įvertinimo naudojimas kalbos dėstymo ir mokymo teorinėms prielaidoms pagrįsti

Santrauka

Kalbos gebėjimo rezultatų vertinimas yra esminis vertinimo reikalavimas, kurį galima naudoti stebint ir vertinant studentų pažangą, mokymo ar dėstymo efektyvumą, o taip pat ir programos kokybę (Gallavara et al., 2008, pp.40-41; OCDE, 2007, 2010). Remdamiesi dalyko vertinimo reikalavimais, kalbos mokytojai ir dėstytojai naudoja daugelį gebėjimo tikrinimo įrankių ir sumuoja įvertinimus, taip tikrindami studentų kalbos įgūdžius. Šiame straipsnyje nagrinėjama 2009–2010 mokslo metų prancūzų kalbos studentų pirmojo kurso pirmojo semestro programa, kurioje yra 33 studentai. Kurso dėstymo metu buvo naudojama šešių tarpinių testų komplektas, kurio rezultatai buvo apjungti į pirmąjį testą 6-tos kurso savaitės metu. Taip pat buvo naudojamas dar vienas komplektas tarpinių testų, kurių rezultatai buvo apibendrinti kurso pabaigoje ir pavadinti antruoju testu. Studentų pažymiai parašyti susumuojant visų 12 tarpinių testų rezultatus. Tam prireikė nemaža dėstytojo pastangų, o rezultatas – tik vienas pažymys vienam studentui. Siekdami padidinti kalbos mokytojų ir dėstytojų įdėtų pastangų vertę, šiame straipsnyje norime parodyti, kaip panaudoti studentų pažangumo rezultatus pačiam mokymo kursui įvertinti. Straipsnyje pateikiame keturias kalbos kurso prielaidas, pedagoginę teoriją, kiekvienos prielaidos pranašumus ir kaip tie pranašumai buvo panaudoti. Studentų rezultatai buvo dar kartą pasitelkti siekiant įvertinti, kiek pasiteisino šios prielaidos. Pedagogikos teorija yra svarbi ta prasme, kad ji padeda spręsti prieštaravimą, atsirandantį kalbos ugdymo literatūroje ir supažindina su neurolingvistiniu mokymosi modeliu NLM (*Neuro-linguistic Learning Model*).

Straipsnis įteiktas 2010 05
Parengtas spaudai 2010 11

About the author

Béatrice Boufoy-Bastick is a Culturometrician and Linguist (MA (FLE) Rouen. MA (by research) Sorbonne-Paris IV, PGCE London Institute of Education, Ph.D at the University of the West Indies in Trinidad).
Research interests: culture, language, culturometrics.
Address: Department of Liberal Arts, University of the West Indies, St. Augustine Campus, Trinidad, Trinidad and Tobago, West Indies.
E-mails: bboufoybastick@gmail.com, beatrice.boufoy-bastick@sta.uwi.edu

DOI: 10.5755/j01.sal.1.17.43390