" IS A PICTURE WORTH A THOUSAND WORDS? " VISUAL STIMULI FOR THE ASSESSMENT OF COMMUNICATIVE-LANGUAGE COMPETENCE IN A SL

Mirella Pederzoli

HAL Id: hal-01252119

https://hal.science/hal-01252119

Submitted on 8 Jan 2016

"Is a picture worth a thousand words?" Visual stimuli for the assessment of communicative-language competence in a SL

Mirella PEDERZOLI
mi.pederzoli@gmail.com
Dipartimento di Ateneo per la Didattica e la Ricerca
Università per Stranieri di Siena
Italie

Abstract : The research described in this article investigates the construct validity of the Certificate of Italian as Second Language (CILS) speaking module, B2 Level of the CEFR, by the *Università per Stranieri di Siena*. The reason of this selection deals with the peculiar structure of the test that in the monological part provides for the possibility to choose among verbal inputs and visual ones, thus considered parallel forms for the completion of a descriptive-argumentative task.

1.   Introduction

This article describes a PhD project concerning the validity of an Italian speaking exam, part of one of the official certificate of Italian as foreign language, named CILS and administered by the *Università per Stranieri di Siena*. Specifically, we refer to validity in terms of construct validity (Messick, 1996), that is the degree to which the test adequately captures the construct of interest. Thus, the area of applicability is the speaking module of the CILS DUE-B2 test that corresponds to the B2 Level of the *Common European Framework of Reference for Languages* (Council of Europe, 2001).

Demonstrating parallel validity in a speaking test consisted of two different input types – verbal and visual- implies demonstrating that they measure the same ability providing evidence from the test itself, from the subjects involved and from statistical analysis. Therefore, the research project reported here drew on established methods for investigating these areas, including the validation process described in the *Manual* (Council of Europe, 2009), the qualitative data gathering through interviews, the socio - semiotic approach (Kress & van Leeuwen, 2006) for the analysis of the visual input, the discourse analysis and the statistical correlations.

2.   Literature review: visual stimuli for the assessment of communicative-language competence

The renowned adage "picture is worth a thousand words" highlights how the usage of visual material is, from a long time, part of the reflections in the field of teaching and learning first languages, but also and foremost, second languages. The broad literature confirms the attention given to this topic, especially concerning the relationship between verbal material and visual material (Mandl & Levin, 1989; Mollica, 2009, 2010, 2011; Paivio, 1971, 1986; Salomon, 1989) and the role of the visual one in the processes of learning, retaining information, problem solving, motivating students, etc. In addition, in the wide range of the applied linguistics, visual stimulus has become a common practice in language assessment or language testing, considering that it is extensively used as an input, to activate a linguistic performance by the test taker (Bae & Lee, 2010; Bax, 2013; Fulcher, 1996, 2003; Ginther, 2001; Ockey, 2007; Reynolds-Keefer, 2011; Suvorov, 2011; Xi, 2005).

« Is a picture worth a thousand words ? » Visual stimuli for the assessment of communicative-language competence in a SL

2

From the analysis of the widespread official language certifications[1] it has emerged a considerable usage of visual support in the tasks provided in all proficiency levels. Indeed, as we can see from Figure 1, it is a quite common practice applying visual material at low levels (A1-A2), confirming the theory that visual supports should have the role to facilitate learning and the process of retaining information, as supporting the cognitive process in the elicitation of the language competence and stimulating a linguistic performance (Mandl & Levin, 1989). It has emerged, in addition, that in the upper levels (C1-C2) the visual materials are used in some cases (Cambridge, Pearson, DELE and CELI tests) as a support for the production of an argumentative test. Indeed,

> «An obvious way of grading difficulty in a suite of speaking is to gradually reduce this support as the level of the exam increases. However, that seems no to have been the policy adopted in the speaking tests of the Cambridge ESOL suite: quite detailed written or visual support for conceptualization is provided in all five speaking tests from KET to CPE. The benefit of this is that it ensures comparability between the performances of candidates at a given level since the concepts and the areas of lexis upon which they draw are similar. Importantly, it also avoids the danger of weighting assessment at the higher levels too heavily in favour of the test taker's imagination rather than their language» (Taylor, 2011: 89)

It is often said that in the process of assessing language proficiency, the visual material validates the test itself, in particular concerning the speaking production because it would provide a rich source of ideas and an efficient means of conveying information without language that might be reformulated by the candidate in their response. However, it is «also time-consuming to develop, because they need to be inspiring enough to generate talk in the first place, structured enough so they really generate the talk that the developers intend, and unpredictable enough so the examinees cannot rehearse their performance in these particular topics and tasks» (Luoma, 2004: 53). The debate about the use of visual material as a way to assess just the speaking ability is still open as it deals with how "speaking production" is understood, that is as a simple unidirectional activity or a way of establishing some kind of interaction among the participants. In the context of this project, the speaking ability is seen as an activity assessed both in interactive and in monological way, consisting of a process of reception, comprehension, production and mediation.

3. The research project

3.1. The speaking module of the CILS DUE-B2 exam

The speaking module of the CILS DUE-B2 Level exam consists of two tasks: the first one is an interaction (2-3 minutes) between the examiner and the test taker about one of the two general topics proposed (e.g. if you could choose a country to live in, where would you go?; what are the advantages and disadvantages of living alone?); the second one is a monologue and the test taker can choose to speak about one of the two general topics proposed (e.g. your plans for the future; a special moment in your life, etc.) or about one of the two pictures arranged (1 minute and a half). Thus, the test taker has the possibility to choose among four

---

[1] Italian :CILS, CELI, PLIDA, IT, Test di conoscenza della lingua italiana. English Cambridge Main Suite Examinations (KET, PET, FIRST, CAE, CPE), IELTS, TOEFL, TOEIC, PTE GENERAL, PTE ACADEMIC, TRINITY, ISE, GESE. Spanish: DELE (Instituto Cervantes). French : DELF, DALF , DILF, TCP. German : FIT IN DEUTCH, GOETHE ZERTIFIKAT. Chinese: HSK, YCT.

different prompts. The same test design is also characteristic of the speaking module of the levels A1, A2 and B1.

3.2.   Context

The research described in this article has its origins in the consideration of three main problems dealing with this kind of test.

1. The low percentage of test takers choosing the visual input in the monological part of the speaking test in all CILS Level exams (6% in December 2013, that is 383 out of 6042 candidates)
2. Inter-rater reliability: some raters tend to assign a low score to the performance that manifests to use the visual input as a springboard for an argumentative rather that descriptive discourse.
3. Practicality: the test needs to be practical in terms of available and required resources (Bachman & Palmer, 1999). In the case of the speaking module of the CILS DUE-B2 exam, the low percentage of test takers choosing the visual input makes the test not practical in terms of test development.

Therefore, considering these three problematic sides of the speaking module of the CILS DUE-B2 exam, this research aims to give a contribution concerning some aspects of the visual inputs applied and move a step forward in the topic of parallel validity.

3.3.   Research questions

This project addresses the following research questions:

1. Does the use of two input types – visual and verbal – in the speaking module of the CILS DUE-B2 Level validate the implied construct in terms of parallel validity?
2. Does the presence of verbal input and visual input in the speaking module of the CILS DUE-B2 Level have any influence on the performance of the test takers in terms of language functions expected in that level of linguistic competence?
3. What are the main facets that influence the test takers' choice of the input type in the speaking module of the CILS exams (A2, B1 and B2).

3.4.   Research approach, procedure and method

Considering the interplay among different disciplines and approaches, the research applies a design called "mix methods research" (Tashakkori & Teddlie, 2003), that is:

> "a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases on the research process. As a method, it focuses on collecting, analyzing and mixing both quantitative and qualitative data in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone" (Creswell, Plano Clark, 2011: 5).

To this end, the study employs a triangulation method (Campbell & Fiske, 1959), that is a combination of two or more approaches and the use of more data sources in order to obtain a deep understanding of the phenomenon investigated. The research procedure follows the subsequent three steps:

- Statement of the problem
- Data collection
- Data analysis

First of all, starting from the winter session of the CILS Examination, it was collected the percentage of candidates that chose the visual input in all the CILS exams and the annotations given by the raters in the rating process. These are internal documents that help CILS Centre to identify possible problems during the exams. Indeed, some problems arose in the speaking module of the CILS DUE-B2 exam in relation to the input types. Thus, in this first part of the study, the researcher also analyzed the different item types applied in the CILS DUE-B2 speaking module from 2004 to 2014, according to the CILS Guidelines, syllabus and Kress and van Leeuwen's visual grammar (2006). Thus, the pictures are analyzed not only in terms of language functions, but also applying the socio-semiotic approach by Kress & van Leeuwen. In their study, the Halliday's metafunctions are applied to the visual products of the western culture. The definition of a "grammar of visual design" begins from the consideration that not much attention has been paid to the meanings of regularities in the way image elements are used in the western cultures.

Considering the *multy-strategy approach* (Bryman, 2001) adopted here, during the second phase of the research data was collected through a qualitative instrument, that is through the administration of two different questionnaires: one for the test takers and one for the examiner. Thus, starting from December 2013 (to December 2014) a questionnaire is given to all the candidates in the CILS exams for the A2, B1 and B2 Levels. In this way, a verbal protocol analysis is applied in order to make inferences on the cognitive and behavioral processes enabled by the test takers. The data collection is currently underway.

Concerning the questionnaire addressed to the examiners, the focus is on the collection of information about the speaking exam context, as the choice of the input type (verbal or visual) by the candidates, the reasons of this choice, the usage of visual material in courses of Italian as L2/LS and in the preparation courses for the CILS exams and the type of activities administered with this material. To this day, the number of questionnaire collected from the examiners are 235.

In order to give evidence for the second research question, during the summer session (June 2014) of CILS DUE-B2 exams, n=90 randomly selected audio-recorded performances were collected and transcribed. The performances belong to 45 candidates that in the speaking module of the exam chose the verbal input and 45 candidates that chose the visual one. The data analysis, in fact, is then conducted through the examination of the transcriptions in terms of language functions expected in this level of linguistic competence (B2). In this way it will be possible to give evidence about differences caused by item types.

Finally, statistical correlation of the scores obtained by the candidates is investigated as a tool adopted by the classical theory regarding the validity of parallel tests.

3.5.   *Expected results*

This project hopes for a valid contribution to the constant activity of investigation in the field of testing and assessment conducted by the CILS Centre, concerning the validity of the administered tests. In particular, it wishes for enhance the awareness regarding the reasons of the low percentage of candidates choosing the visual input, in order to force a deep consideration in terms of test design and its practicality. As a matter of fact the presence of four prompts in the speaking module of the CILS DUE-B2 exam – dealing with the willingness to provide the candidates with all the possibilities to demonstrate their speaking ability- is a time consuming choice in the process of test design. Thus, the analysis of the different items and task types of speaking test in the official certifications of foreign languages mentioned before and the results of this research will provide the CILS Centre with sufficient material to think about the sustainability of this approach and eventually the possibility of revising the speaking module of the CILS DUE-B2 exam.
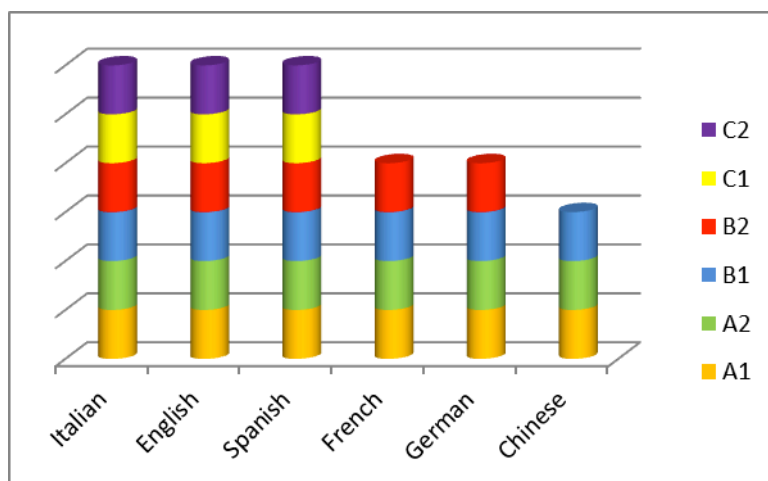
4.   Pictures

Figure 1 – Visual input for language certifications analyzed and proficiency level

References

BAE, Jungok. & LEE, Yae Sheik (2010). The validation of parallel test forms: "mountain" and "beach" pictures series for assessment of language skills, *Language Testing, 28*, 155-177

BAX, Stephen (2013). The cognitive processing of candidates during reading tests: evidence from eye tracking, *Language Testing*, *30*: 441-465

BRYMAN, Alan (2011). *Social Research methods*, Oxford: Oxford University Press

COUNCIL OF EUROPE (2001), *Common European Framework of Reference for Language. Learning, Teaching, Assessment – CEFR*, Cambridge: CUP

COUNCIL OF EUROPE (2009). *Relating language examinations to the Common European Framework of Reference for languages: learning, teaching, assessment. A manual,* Cambridge: CUP

CRESWELL, John W. & PLANO CLARK, Vicki L. (2007), *Designing and conducting mixed methods research*, Sage Publications, Thousand Oaks.

FULCHER, Glenn (1996). Testing tasks: issues in task design and the group oral, *Language Testing, 13*: 23-51

FULCHER, Glenn & MARQUEZ, Reiter (2003). Task difficulty in speaking tests, *Language Testing, 20*: 321-344

GINTHER, April (2001). Effects of the presence and absence of visuals on performance on TOEFL CBT listening comprehensive stimuli, *Educational testing service*, *research reports*, *6*: 1-53

GOLDSTEIN, Ben (2008). *Working with images*, Cambridge: CUP

HALLIDAY, Michael Alexander (1978). *Language as social semiotic: the social interpretation of language and meaning*, London: Arnold

KRESS, Gunther & VAN LEEUWEN, Theo (2006). *Reading images: the grammar of visual design*, London: Routledge

LEVIN, Joel R. & MANDL, Heinz (1989). *Knowledge acquisition from text and pictures*, New York: Elsevier Service

LUOMA, Sari (2004). *Assessing speaking*, Cambridge: CUP

MESSICK, Samuel (1996). Validity in performance assessment, in *Technical issues in large-scale performance assessment*, Washington DC: National Center for Educational Statistics, 1-19

MOLLICA, Anthony (2009). *Teaching and learning languages*, Perugia: Guerra-Soleil

MOLLICA, Anthony (2010). *Ludolinguistica e glottodidattica*, Perugia: Guerra-Soleil

MOLLICA, Anthony (2011). L'uso didattico del testo umoristico: la vignetta, in *Insegnare italiano a stranieri*, Milano: Le Monnier, 406-423

OCKEY, Gary J. (2007). Construct implications of including still image or video in computer-based listening tests, *Language Testing, 24*: 517-537

PAIVIO, Allan (1971). *Imagery and verbal processes*, New York: Rinehart & Co

PAIVIO, Allan (1986). *Mental representations: a dual coding approach*, New York: Oxford University Press

SALOMON, Gavriel (1989). Learning from texts and pictures: reflections on a meta-level, in *Knowledge acquisition from text and pictures*, Amsterdam: Elsevier, 73-82

STURKEN, Marita & CARTWRIGHT, Lisa (2001). *Practices of looking. An introduction to visual culture*, Oxford: Oxford University Press

SUVOROV, Ruslan (2009). Context visuals in L2 listening tests: the effects of photographs and video vs. audio-only format, in *Developing and evaluating language learning materials*, Ames, IA: Iowa State University, 53-68

TAYLOR, Lynda (2011). Examining speaking: research and practice in assessing second language speaking, in *Studies in language testing 30*, Cambridge: CUP

TASHAKKORI, Abbas & TEDDLIE, Charles (2003). *Handbook of Mixed Methods in Social and Behavioral Research*, Calif: Thousand Oaks

WEIDENMANN, Bernd (1989). When a good pictures fail: an information-processing approach to the effect of illustrations, in *Knowledge acquisition from text and pictures*, New York: Elsevier Service, 157-170

WEIR, Cyril (2005). *Language testing and validation: an evidence-based approach*, Basingstoke: Palgrave Macmillian

XI, Xiaoming (2005). Do visual chunks and planning impact performance on the graph description task in the SPEAK exam?, *Language Testing, 22*: 463-508