

Paola Masillo

Linguistic integration and residence policies in Italy: issues and perspectives

Abstract: This paper reports on the development of policies and practices designed to support the linguistic integration of adult migrants in Italy. The focus is on the use of language requirements: in order to secure a permanent residence permit, non-EU citizens must pass a compulsory Italian language test. The purpose is to highlight the lack of coherence and effectiveness of Italian policy with regard to the issues of *fairness* and *ethics*.

Résumé : L'article informe sur le développement des politiques et pratiques italiennes pour favoriser l'intégration linguistique des migrants adultes. L'accent est mis sur le recours à des conditions de connaissances linguistiques, via la réussite du test d'italien obligatoire exigé des ressortissants non-UE qui demandent un titre de séjour permanent. Le but est de mettre en avant l'incohérence et l'inefficacité des politiques italiennes, dues à leur manque d'équité et d'éthique.

1 Introduction

Language assessment has a long history of use as an instrument of social policy and practice, in particular in the role of “linguistic gate-keeping” (Hogan-Brun, Mar-Molinero, and Stevenson 2009). Language tests and their results always have consequences for test takers in educational, professional or social contexts, but sometimes tests may be used to take high-stakes decisions in respect of migration, such as permission to migrate to a new country and receiving an official document to stay there (Shohamy 2001).

Paola Masillo, University for Foreigners of Siena, Italy, E-mail: masillo@unistrasi.it

DOI 10.1515/9783110477498-035,  © 2017 Paola Masillo, published by De Gruyter.

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.

Unauthenticated

Download Date | 6/21/18 4:39 PM

2 Research context

The starting point of this study was reflection on the concept of integration as it is defined in some fundamental European documents.¹ In particular, the focus was on one of the most frequently used parameters for measuring the level of integration in a host country, proficiency in its language. The research investigated testing regimes in which language tests are both assessment tools with a diagnostic purpose and “powerful tools” because they are required to access residence rights (McNamara 2000; Shohamy 2001). The research was also concerned with the impact of the *Common European Framework of Reference for Languages* (CEFR; Council of Europe 2001) on language policies, mostly in Europe. Since its publication, the CEFR has become the main reference document in the fields of language learning, teaching and assessment. Nowadays it has even assumed an operational function, being used to justify choices in language policies of integration and immigration (Extra, Spotti, and Van Avermaet 2009).

3 Overview of the study

In the last decade, successive surveys have demonstrated that language and cultural knowledge play a central role as prerequisites for obtaining residence permits and/or citizenship in many countries, particularly in Europe. As demonstrated by the most recent survey (Pulinx, Van Avermaet, and Extramania 2014), many European countries are increasingly requiring adult migrants to attain certain minimum levels of proficiency in the language of the country before they are granted the right to long-term residence. Consequently, an increase in the use of language tests based on the CEFR descriptors has been found across the whole of Europe. In Italy, non-EU citizens who apply for a permanent residence permit are asked to take an official language test in order to demonstrate that they have reached CEFR level A2 (Law no. 94/2009; MD 4 June 2010).

3.1 Aims and objectives

Good practice in test design requires that test developers meet the following conditions: “results are standardised and reliable, which means that it is easy to

¹ See among others: Resolution 68 (18); Resolution 1437 (2005); Resolution 1511 (2006); COMM (2003) 336.

compare candidates across the same or different administrations; candidates are assessed with a high degree of independence and objectivity; large numbers may be tested in a short space of time; test validity helps to ensure fairness” (Council of Europe/Language Policy Unit, Project LIAM). Starting from these theoretical preconditions, the scope of this study is to highlight the lack of coherence and effectiveness of Italian policy with regard to the issues of fairness and ethics. The study also aims to investigate *test usefulness* (Bachman and Palmer 1996), taking into account three factors: *validity*, *reliability* and *impact*.

3.2 Research questions

The crux of the matter underpinning the research questions concerns the issue of ethics in adopting a test not only as a tool to measure linguistic competences, but also as an instrument to discriminate and even exclude.

Based on this preliminary observation, three research questions were developed. The first one arises from the Italian government’s decision not to use a national standardised test, but to leave test development and administration in the hands of each adult education centre, which raises questions of fairness, validity and reliability.

The second research question takes into account test reliability, in terms of comparability of test structure, content, and results, considering the uneven score distribution obtained in the first year of administration.

The third research question focuses on test design: the construct is mainly based on receptive and written skills² and the assessment criteria focus on accuracy, consistency and appropriateness. The test design makes no reference to speaking or effective communication, which raises problems of test validity.

3.3 Study procedures and sample characteristics

This PhD study was carried out over three years. In the first year, a database was established of 83 potentially equivalent test forms developed for migrants throughout Italy. Following a survey of outcomes, we selected two test forms developed by the two most representative areas, according to the pass rates: Piemonte (90.2%) and Veneto (71.6%).

² The test content consists of three sections: listening, reading and written interaction (MIUR 2010).

In the second year, we administered the two selected tests to a representative sample of test takers, following the official guidelines given by the Italian Ministry of Education (MIUR 2010).

In the third year, we focused on data analysis. In the first stage, a content analysis was carried out via correlation with the CEFR descriptors, according to the *Standard Setting* method. In the second stage, a selection of statistical analyses was run in order to measure test reliability following *Classical Test Theory* and focusing on the two receptive skills, listening and reading. In the last stage, the focus was on the validation of the assessment scale by assessing a sample of written performances and running a judgment analysis.

4 Findings and discussion

The aim of this paper is to show the main results reached during the three-year project, followed by a discussion and final considerations.³

4.1 Unfair testing

The first research question can be confirmed, first of all, in the light of the results obtained through the *standard setting* process, in particular the *specification* stage⁴ (Table 1). Looking at the mean values obtained, there is a difference in the level of item difficulty recognised, in particular between the Reading 1 components of the two tests.

Table 1: Standard setting – specification

| Mean | Test 1 | Test 2 |
|-------------|--------|--------|
| Listening 1 | 2.38 | 2.7 |
| Listening 2 | 2.2 | 2.78 |
| Reading 1 | 1.54 | 3.09 |
| Reading 2 | 2.3 | 2.3 |

Secondly, if we look at the *frequency distribution* (Table 2), the outcome confirms the unfairness of the two tests. 63.7% of the test takers achieved the maxi-

³ For research purposes, we named the two tests Test 1 (Piemonte) and Test 2 (Veneto).

⁴ For research purposes, we named the CEFR levels 1 (A1), 2 (A2), and 3 (B1).

num score (10) for listening in Test 1, while only 28.7% of the same sample achieved the maximum score (10) in Test 2.

Table 2: Frequency distribution

| Test 1 | | Test 2 | |
|-----------|--------|-----------|--------|
| Listening | 63.7 % | Listening | 28.7 % |
| Reading | 28.7 % | Reading | 17.8 % |

Descriptive statistics (Table 3) gave us the third confirmation of the *unfairness* of the two tests. The difference between the mean values of the two tests confirms that Test 1 is easier than Test 2.

Table 3: Descriptive statistics – mean

| Test 1 | | Test 2 | |
|-----------|------|-----------|------|
| Listening | 9.31 | Listening | 8.46 |
| Reading | 8.20 | Reading | 7.13 |

4.2 Non-equivalent testing

The first confirmation of low test reliability is the *Cronbach's alpha* value (Table 4), since it appears generally below the minimum threshold of acceptability (.7).

Table 4: Cronbach's alpha

| | Listening 1 | Reading 2 |
|--------|-------------|-----------|
| Test 1 | .388 | .566 |
| Test 2 | .512 | .627 |

Correlations are weak (.390) between the listening tests and moderate (.483) between the reading tests. In addition, a very low value was detected for *shared variance*: there is a degree of overlap of 15% for listening and 23% for reading.

The final confirmation of the second hypothesis is given by the *paired-samples t-tests* and the *Wilcoxon signed rank tests*, since they confirm that the corre-

lations and differences found between the mean values of the two tests were statistically significant.

4.3 The low validity and low reliability of the assessment scale

As regards the third research question, we selected a sample of 20 written performances by test takers who took both tests. The *standard deviation* value for Test 1 indicates more similar and more clustered scores around the mean; by contrast, for Test 2, it confirms a greater spread both above and below the mean. For both tasks, the *range* is very low, considering the 35-point assessment scale, as if the judgment process takes into consideration only one third of the scale (Table 5).

Table 5: Rating process (statistics)

| Written interaction | | Test 1 | Test 2 |
|---------------------|---------|--------|--------|
| N | Valid | 20 | 20 |
| | Missing | 0 | 0 |
| Mean | | 29.80 | 29.25 |
| Mode | | 27 | 29 |
| Std. Deviation | | 2.707 | 3.810 |
| Range | | 10 | 16 |

The outcome obtained by the judgment analysis was the following: 11 raters out of 11 considered the scale slightly acceptable and they stated the need to focus on other assessment criteria to achieve the assessment task.

5 Final considerations

In line with the recent Recommendation 2034 (Strik 2013), the research attempted to provide scientific evidence about the (mis-)use of language tests for purposes of integration, and to share potential good practice to support the linguistic integration of adult migrants.

The crux of the matter is the lack of a language testing and assessment culture in Italy, which is reflected in the choices that underpin the activated procedures (Barni 2012). Test developers should ensure that they follow appropriate

testing procedures, considering the high-stakes decisions that will be based on their test (Van Avermaet and Rocca 2011).

A second critical point concerns the lack of fairness and transparency caused by decentralised procedures and the involvement of test developers who lack appropriate training. Test fairness is relevant to all types of language test and all test takers, but it is especially important when used for migration purposes (ALTE 2016).

Our case study confirms that Italy's *testing regime* uses tests as instruments of power and only apparently to support integration. As regards new perspectives to focus on, the first is the need to monitor the assessment tools and carry out validation studies, in order to ensure that the tools measure the abilities they claim to measure in a valid and reliable way. A collaborative effort is needed to guarantee the quality of assessment procedures so that they meet current professional standards (validity, reliability, and transparency). Finally, it is necessary to promote studies that focus on the impact of language policies on the social integration and social inclusion of adult migrants in the host country (Pulinx, Van Avermaet and Extramiana 2014).

References

- Bachman, Lyle F. & Adrian S. Palmer. 1996. *Language testing in practice*. Oxford: Oxford University Press.
- Barni, Monica. 2012. *Diritti linguistici, diritti di cittadinanza: l'educazione linguistica come strumento contro le barriere linguistiche*, 213–23. Roma: Bulzoni.
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Council of Europe/Language Policy Unit (Strasbourg) – Project LIAM. Tests (language test). <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680494490>.
- Extra, Guus, Max Spotti & Piet Van Avermaet (eds.). 2009. *Language testing, migration and citizenship: Cross-national perspectives on integration regimes*. London: Continuum.
- Hogan-Brun, Gabrielle, Clare Mar-Molinero & Patrick Stevenson (eds.). 2009. *Discourse on language and integration*. Amsterdam: John Benjamins.
- McNamara, Tim. 2000. *Language testing*. Oxford: Oxford University Press.
- MIUR – DG IFTS. 2010. *Vademecum. Indicazioni tecnico-operative per la definizione dei contenuti delle prove che compongono il test, criteri di assegnazione del punteggio e durata del test*. <http://hubmiur.pubblica.istruzione.it/alfresco/d/d/workspace/SpaceStore/d6686cab-4f36-4c32-acb3-97f2090ead92/vademecum.pdf> (accessed 20 June 2016).
- Pulinx, Reinhilde, Piet Van Avermaet & Claire Extramiana. 2014. *Linguistic integration of adult migrants: Policy and practice. Final report on the 3rd Council of Europe survey*.

- Strasbourg: Council of Europe. <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016802fc1ce> (accessed 28 July 2016).
- Shohamy, Elana. 2001. *The power of tests: A critical perspective on the use of language tests*. Harlow: Pearson.
- Strik, Tineke. 2013. Integration tests: Helping or hindering integration? Report of the Committee on Migration, Refugees, and Displaced Persons. Strasbourg: Council of Europe, Parliamentary Assembly, Doc. 13361.
- Van Avermaet, Piet & Lorenzo Rocca. 2011. Language testing and access. LAMI forum at the 4th ALTE International Conference, Kraków, 8 July, 2011). <https://biblio.ugent.be/publication/4284783/file/6809549> (accessed 20 June 2016).