Materiali e documenti

The "MATEL" Project: Research Results

edited by
Maria Antonietta Pinto

SAPIENZA
UNIVERSITÀ EDITRICE

Collana Materiali e documenti 9

The "MATEL" Project: Research Result

edited by

Maria Antonietta Pinto

SAPIENZA
UNIVERSITÀ EDITRICE
2015

This project has been funded with support from the European Commission. This publication reflects the views of the authors only, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Sapienza Università Editrice
Piazzale Aldo Moro 5 – 00185 Roma

www.editricesapienza.it
editrice.sapienza@uniroma1.it

Iscrizione Registro Operatori Comunicazione n. 11420

ISBN 978-88-98533-72-5

Pubblicato a dicembre 2015

In copertina: la *word cloud* è stata generata attraverso il sito www.wordle.net

Contents

INTRODUCTION

Maria Antonietta Pinto

This publication is a shortened version of a Special Issue of the RIVISTA DI PSICOLINGUISTICA APPLICATA / JOURNAL OF APPLIED PSYCHOLINGUISTICS (Ed. by M.A. Pinto, XV, 2, 2015), entirely devoted to the presentation of the results of the "MATEL" [1](Metalinguistic Awareness Tests in European Languages) Longlife Learning Project, financed by the European Union in the years 2013-2015. The project had two major aims, a research and an educational aim. The Special Issue focuses on the outcomes of the research part, whose objectives were to validate two metalinguistic ability tests (MATs, in English acronym, Pinto, Titone & Trusso, 1999), translated and adapted from the original Italian version (Pinto, 1999) into Spanish (Pinto, Titone & Gonzáles Gil, 2000) and French (Pinto & El Euch, 2015), and to translate them into a further language, namely German.

The first article presents the theoretical background and the structural characteristics of the two tests in their original form, as TAM (Italian acronym for *Test di Abilità Metalinguistiche*) and overviews the extensive research that has been conducted with these instruments during the past twenty years, approximately, in all the linguistic versions available and in various countries.

These tests address different developmental stages and educational levels. The first, named TAM-2, covers the age range from 9 to 14, and is therefore applicable from the end of primary to the beginning of secondary level, whereas the second, named the TAM-3, addresses late adolescence up to adulthood, from 16 onwards. The numbers 2 and 3 are due to the fact that there also exists a TAM-1 for children aged 4 to 6, that was developed as part of a comprehensive project on metalinguistic awareness promoted and supervised by Titone and Pinto, since the late 80s (Titone, 1994). The article highlights the structural features that are common to the two TAMs, beyond differences in linguistic complexity and general culture requirements. These features express the particular conception of the theoretical construct measured by the tests, namely metalinguistic awareness (MLA, henceforth). This construct posits a definite distinction between an intuitive form of MLA, as explored by a first, global question in a given metalinguistic task, and an explicit, argumented form of MLA, which comes to light when subjects are requested to justify their answers. This is where we can ascertain whether subjects are really able to "go beyond" (according to the meaning of the

1

[1] Metalinguistic Awareness Tests in European Languages, Project number 543277-LLP-1-2013-1-IT-KA2-KA2MP. www.matelproject.com).

For copyright reasons, the Special issue of the RiPLA / JAPL (XV-2-2015) is not accessible on the website of the MATEL Project. Based on an agreement between the Coordinator of the Project, Maria Antonietta Pinto and he publisher of the journal, Fabrizio Serra, the Coordinator has been authorized to make this shortened version available on the Project's website.

greek prefix “meta”) the surface of the item, and reach a consistent “meta” level. To this end, they necessarily have to analyse forms and meanings and find a suitable solution to the metalinguistic problem posed by the task they are confronted with. This distinction, which calls upon cognitive processes of a different level, is reflected in two different scores, the L and the ML, respectively. L is the acronym for Linguistic, as the first response requires the retrieval of linguistic rules and conventions, whereas ML is the acronym for Metalinguistic as ML questions require an in-depth analysis of forms and meanings. The coding system for this more abstract level is based on specific theoretical grounds, different from those that underlie the coding system of the L level. And here we reach the core of the construct.

What is claimed is that a *metalinguistic task is inherently based on a conflict between linguistic data.* It can be a conflict between two or more different forms of the same meaning or between two or more meanings expressed by the same form, or between two words simply differing by one letter or by an accent or by their position in a sentence, etc. The way the conflict is solved is not reductible to a yes-or-no, or false/right answer. It does not simply reveals knowledge about language but *reasoning about language.* To assess this type of reasoning on linguistic conflicts, a precise cognitive model has been adopted, derived from the latest equilibration model of Piaget (1975). According to Piaget (1975), solving conflicts between data requires to recompose contradictions into a new equilibrium, a new synthesis. To face these contradictions, subjects may use three types of cognitive processes, called “mental regulations”. At the lowest level, with an “alpha” regulation, subjects simply ignore the perturbation created by a given conflict. At an immediately higher level, that of “beta” regulations, subjects take into account the perturbation and face it but they limit their action to local and successive adjustements, without an overall view of the problem at hand. Only at the “gamma” regulation level is the conflict consistently recomposed, and the representation of the problem restructured in such a was as to justify the place of each particular elment in a comprehensive, abstract system.

The coding system created for the TAM is a three-step scale where alpha regulations, under the form of blanks, or “I don’t know” answers, or of pseudo-justifications, are scored 0, beta regulations, under the form of relevant but partial linguistic elements, are scored 1 and gamma regulations, under the form of relevant and exhaustive answers, are scored 2. The idea that metalinguistic tasks are intrinsically based on conflicts between this or that aspect of language, and that this conflict may be faced at qualitatively different levels is the common conceptual thread that underlies the coding system of the ML responses in the TAM-2 and the TAM-3.

The complex architecture of this MLA construct makes it challenging, especially if we consider that it addresses different developmental stages with different linguistic materials. The challenge is further amplified by the linguistic variety of the versions of the tests currently existing, This variety was one of the factors that motivated to perform the validation of these tests in other languages, and translate them in one

more European language by means of the MATEL project.

The next three articles deal precisely with the outcomes of the validation studies performed respectively with the Spanish THAM-2 (*Test de Habilidades Metalingüísticas n.2,* Núñez Delgado & Pinto, 2015) in Granada (Spain) and surroundings, the Spanish THAM-3 (*Test de Habilidades Metalingüísticas n.3,* Lasagabaster, Merino & Pinto, 2015), in Soria (Spain), and the French THAM-3 (*Test d'Habiletés métalinguistiques n.3, Pinto,* El Euch, 2015) in Québec (Canada). The psychometric studies followed the same pattern adopted for the validation of the corresponding Italian tests, except for the study of construct validity. In the original tests (TAM-2, Pinto, Candilera & Iliceto, 2003; TAM-3, Pinto & Iliceto, 2007) construct validity was studied with Exploratory Factor Analyses (EFA), which revealed a two-factor model matching the above distinction between the L and the ML levels of awareness. The studies of construct validity conducted within the MATEL project considered this two-factor model as a starting point to be examined by means of Confirmatory Factorial Analyses (CFA, Brown, 2006).

Whenever possible, the values found in each of these new validations were compared with those found in the original Italian tests. Important commonalities emerged between the newly validated tests and the original tests. In all cases, performances at the L level were much higher than those at the ML level. As the former measure MLA at the intuitive level and the latter measure MLA at the explicit level, the more discriminant character of this explicit level is also a confirmation of its abstract, metacognitive nature. Another recurrent pattern highlighted by the results of these validations was that the Acceptability subtest in the Spanish THAM-3, French THAM-3, and Italian TAM-3 was by far the most difficult of the three subtests that compose the test, especially in the ML part, i.e. when subjects must justify their answers. Differently from the other subtests, namely, Comprehension and Figurative language, that measure metasemantic abilities, Acceptability is a metagrammatical subtest, that requires solid grammatical knowledge and the capability of detecting where and how grammatical rules are violated. This is what makes it a particularly demanding task, both on linguistic and metalinguistic grounds. The evident difficulties this subtest poses should probably induce researchers to modify something in the text in which errors are embedded, in the typology of errors, and in the strictness of the coding system. Nonetheless, the gaps found in performances in metasemantic and metagrammatical tasks could also be attributed to intrinsic aspects of the metalinguistic abilities required in each case. While categorizing and explaining grammatical errors requires a strong normative attitude towards language, this attitude is much less relevant when the issue is to elaborate some coherent interpretation of meanings in a metasemantic task. This raises a delicate theoretical point regarding the diversification of MLA in relation to the language areas it refers to (semantics, grammar, syntax, phonology, pragmatics, etc.). A third important commonality between the validated versions of the tests described in this issue and the original Italian tests is the robustness of the ML factor, that emerged from the EFA as well as from the CFA. Given the relevance of construct validity in a test, this is to be considered a

solid result. More importantly, given the architecture of the MLA construct we delineated, and the characteristics attributed to the ML dimension, consistency found in this dimension across all the linguistic versions available is especially significant.

Overall, the Spanish THAM-2, the Spanish THAM-3 and the French THAM-3 met the requirements for reliable and valid instruments. All three can be used in their respective countries with reference norms for populations having Spanish or French as a first language, in and out of Europe, and also for learners of Spanish or French as second or foreign languages. This will offer important opportunities for both research and education.

For research, the study of convergent validity of each metalinguistic test can be completed by examining the relation with basic language competence tests, mainly in the semantic, grammatical and syntactic areas, as these are involved in the two metalinguistic tests we addressed in the Project. Research on bilingualism involving Spanish or French or both languages can also greatly benefit from the validated tests because it will be possible to make comparisons between bilingual and monolingual samples on the basis of precise reference norms. Research on education will be able to use the metalinguistic tests as possible predictors of other language or cognitive competencies relevant to academic achievement, as it has been the case for the Spanish THAM-2 in relation to text production in the first language (De Haro, Núñez Delgado & López 2012), or to second language learning (Perales &Cenoz, 2002).

But the validated tests open new perspectives also for teachers, educational and clinical psychologists, as they permit to draw *profiles* of strenghts and weaknesses of single students or of entire class groups. On these grounds, it is possible to devise curricula for empowering metalinguistic abilities in weak areas, or exploiting them in stronger areas. Empowering students' ability of analysing language appropriately, as required in the ML parts of the metalinguistic tests, can help students to create a method they can autonomously apply to whatever type of text, be it a mathematics, history, geography, social sciences, or literary text. In this vein, the MATEL Project has devised educational resources[2] that are the natural application of the spirit of the metalinguistic tests, where reflection on the relationship between form and meaning is the central focus. It is beyond the scope of this introduction to describe these resources. We will just say that they consist in metalinguistic exercices where teachers draw students' attention on critical aspects of texts, in mathematics as well as in history, geography, social sciences, etc., by means of group activities conducted in the classroom.

The translation of the Italian TAM-2 and TAM-3 into German (Jessner, Hofer, & Pinto, 2015; Jessner, Pellegrini, Moroder, Hofer, & Pinto, 2015) raised issues of a different kind, but always related to the distinctive characteristics of the tests and their constructs. The focus was not only theoretical but also

2 [!] available at www.matelproject.com

empirical because the translated versions were used in pilot studies on samples of German-speaking students to see the impact of the tests on participants of the same age as those of the original test. The two last articles describe the challenges encountered during the translation process and the results of these pilot studies. Compared to the other translations of the same tests, from Italian into Spanish, French and English, German posed further linguistic specificities that required a larger number of adaptations and reformulations. The outcomes of the pilot studies seem to indirectly confirm the appropriatedness of the linguistic renderings. The German equivalent of the TAM-2, i.e. the MKT-2, was administered in a class of 7th graders (Mittel schule) in a small village in South Tyrol (North of Italy), and the equivalent of the TAM-3, i.e. the MKT-3, was administered in two High school classes of the same region. In both cases, participants were native speakers of German. In general terms, the *type* of answers provided by the participants were the same as those elicited by the Italian TAM-2 and TAM-3, which indicates that the *processes* elicited by the items and the questions were of the same nature as those elicited by the original version. In addition, in quantitative terms, the L performances of these native German speakers were totally comparable to those of the Italian counterparts, and the ML performances were even superior. This happened at each age level, with the MKT-2 as well as with the MKT-3. If the participants were able to process even the most demanding questions at a high level the translations must have captured the essence of the metalinguistic tasks successfully. In addition to basic descriptive statistics, with means, standard deviations and percentages of score levels, some further statistical analyses were performed for each of the two tests. These showed normally distributed curves, good internal consistence of the items, significant correlations between subtests and, as with the validated Spanish, French and Italian tests, a consistent ML factor. Although preliminary, these outcomes stimulate further validation studies of both the MKT-2 and MKT-3 on larger samples, in other German-speaking countries in Europe (Germany, Austria, Switzerland).

The MATEL Project offered a special chance to broaden the range of the metalinguistic tests currently available by adding three validated versions in Spanish and French, which will consolidate and extend their usability, and a new linguistic version of the same tests in German that yielded promising outcomes. We believe that each linguistic version of these tests and the study of their empirical impact will illuminate our understanding of the concept of MLA, and provide further ideas for exploiting its educational potentialities.

ACKNOWLEDGMENTS

The authors of the studies described in this publication wish to thank the E.U. for the funding received through the Lifelong Learning Project "MATEL" (Metalinguistic Awareness Tests in European Languages, www.matelproject.com). Project number 543277-LLP-1-2013-1-IT-KA2-KA2MP.

REFERENCES

BROWN, T.A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York, NY: Guilford.

DE HARO, E. F., DELGADO, M. P. N., & LÓPEZ, A. R. (2012). Consciencia metalingüística y enseñanza-aprendizaje de la composición escrita en educación primaria: Un estudio empírico. *Rivista di Psicolinguistica Applicata, 12*(1-2), 25-46.

JESSNER, U., HOFER, B., PINTO, M.A., *(2015). MKT, Metalinguistischer Kompetenztest Teil 2*. Innsbruck, Studia UniversitätsVerlag. Innsbruck.

JESSNER, U., PELLEGRINI, C., MORODER, V., PINTO, M.A (2015*). MKT. Metalinguistischer Kompetenztest Teil 3*. Innsbruck. Studia UniversitätsVerlag Innsbruck.

LASAGABASTER, D., MERINO, J.A., & PINTO, M.A. (2015). *Test de habilidades metalingüisticas para adolescentes y adultos: THAM-3.* Bilbao: Universidad del País Vasco/Euskal Herriko Unibertsitatea. www.matelproject.com.

NÚÑEZ DELGADO, P., & PINTO, M.A. (2015). *THAM-2. TEST DE HABILIDADES METALINGÜÍSTICAS N.2 (9-14 AÑOS).* www.matelproject.com.

PERALES, J., & CENOZ, J. (2002). The effect of individual and contextual factors in adult second-language acquisition in the Basque country. *Language, Culture and Curriculum*, 15 (1), 1-15.

PIAGET, J. (1975). *L'équilibration des structures cognitives.* Paris, P.U.F.

PINTO, M.A. (1999). *La consapevolezza metalinguistica. Teoria, sviluppo e strumenti di misurazione.* Pisa-Roma. Istituti Editoriali e Poligrafici Internazionali.

PINTO, M.A., CANDILERA, G., & ILICETO, P. (2003). *TAM-2. Test di abilita`metalinguistiche n.2 (9-14 anni). La valutazione dello sviluppo metalinguistico tra scuola elementare e scuola media*. Roma: Scione Editore.

PINTO, M.A., & EL EUCH, S. (2015). *La conscience métalinguistique. Théorie, développement et instruments de mesure*. Québec, P.U.L.

PINTO, M. A., ILICETO P. (2007). *TAM-3 Test di abilità metalinguistiche n° 3*, *Fascia adolescente adulta.* Roma: Carrocci Faber.

PINTO, M. A., TITONE, R. & GONZÁLES GÍL, M.D. (1999). *La consciencia metalingüística. Teoría, desarrollo e instrumentos de medición.* Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali.

PINTO, M. A., TITONE, R. & TRUSSO, F. (1999). *Metalinguistic Awareness. Theory, Development and Measurement Instruments*. Pisa-Roma: Istituti Editoriali e Poligrafici

TITONE, R. (1994). Bilingual education and the development of metalinguistic abilities: a research project. *International Journal of Psycholinguistics.* 10 (5), 5-14.

THE ITALIAN METALINGUISTIC ABILITY TESTS TAM-2 AND TAM-3 (*Pinto & Titone 1989; Pinto, 1995, 1999*) AND THEIR USE IN RESEARCH: AN OVERVIEW

MARIA ANTONIETTA PINTO*
Università di Roma "SAPIENZA"

ABSTRACT:

This article describes two metalinguistic ability tests, the TAM-2 and the TAM-3 (Pinto & Titone, 1989; Pinto, 1995, 1999), which address two different age ranges, 9-14 and late adolescence-adulthood, respectively. In both tests, there is a clear distinction between L questions, which elicit an intuitive and global form of metalinguistic awareness (henceforth, MLA), and ML questions, which request justification of the previous L questions, and thus elicit MLA at the explicit level. To assess these more complex ML processes, a three-step scale has been created, based on Piaget's final equilibration model (Piaget, 1975), that posits three types of mental regulations to face cognitive conflicts: alpha, beta, and gamma. As metalinguistic tasks are intrinsically generated by cognitive conflicts between linguistic elements of different nature (in form and in meaning), the alpha, beta, and gamma pattern has been transposed to the metalinguistic domain. This coding system ensures continuity between the TAM-2 and the TAM-3, beyond differences in metalinguistic complexity (See also the Introduction, in this publication). The article also offers an overview of the research that has been conducted with the TAM-2 and the TAM-3, not only in their original linguistic version, i.e. Italian, but in all the other linguistic versions available (English, French, Spanish, and German).

THE USE OF THE TAM-2 AND THE TAM-3 IN RESEARCH: AN OVERVIEW

Extensive empirical research has been conducted with the TAM-2 and the TAM-3, in the original language as well as in other linguistic versions, in European countries, but also in Canada and Argentina. In Italy, from 1992 to 1998, the TAM-2 was used in a broad research design where bilingual 5th graders (mean age: approximately 10.6) enrolled in international schools in Rome (The French Lycée "Chateaubriand"; the "Deutsches Schule" and the "Schweitzer Schule" –where German is the language of the curriculum– and an Italian-English school) were compared with Italian monolingual controls on MLA. Participants were matched by school grade, age mean, and nonverbal intelligence, as measured by the Raven's Progressive Matrices (Raven, Raven & Court, 1998), gender

* Corresponding author: Maria Antonietta Pinto. Università di Roma "SAPIENZA". Via dei Marsi, 78. 00185. Roma. E-mail; mariantonietta.pinto@uniroma1.it.

balance, sociocultural background, residential area and number of hours of Italian classes. In all comparisons, results (Pinto, 2002: Pinto & Melogno 2014) showed metalinguistic superiority of the Italian-French, Italian-German and Italian-English children versus their Italian monolingual counterparts in a test that, at the time, only existed in Italian, i.e. the native language of the monolinguals.

The TAM-2 was also used to test the impact of experimental language curricula on MLA. These included: enriched foreign language teaching (national languages), heritage language teaching (various local vernaculars, Sardinian), anticipated teaching of a classical language (Latin), and planned language teaching (Esperanto). In these studies, experimental classes were compared with control groups enrolled in the same schools, and were matched by sociocultural and intellectual indicators. In all cases, the impact of the experimental language curricula proved significantly effective on MLA, as measured by the comparison between the experimental and control classes' performances on the TAM-2.

Selected items of the TAM-2 were used by Hofer (2013, 2015) to study the relationship between MLA and bilingual education in South Tyrol. This region of Italy is officially bilingual, with Italian and German as the two main languages, both compulsory in the school system, but with a different balance according to the type of curriculum. Some schools have implemented a bilingual education model where the two languages are assigned an equal number of hours. The results of Hofer's study showed that metalinguistic performances were significantly higher in pupils enrolled in this balanced model of bilingual education than in control groups enrolled in a curriculum where Italian is the L1 and German the L2.

The whole battery of the TAM-2 was used as a pre-test/post-test measure to assess MLA in socioculturally deprived children enrolled in a seven months programme aimed at enhancing their metalinguistic abilities (Pinto & Fulgenzi, 2014). The program significantly improved the MLA of this experimental group, as measured by the pre/post-test differences in performances on the TAM-2 in comparison with a control group.

As forms of MLA also exist in atypically developing subjects, the TAM-2 was used as a diagnostic tool in a case study on a gifted child with autistic spectrum disorders (Melogno, Pinto & Levi, 2014). The test pointed out strengths and weaknesses that made it possible to highlight core aspects of the intellectual functioning in this type of deficit.

In an Argentinean school (Rosario), where Spanish is the L1 and Italian the L2, Pafumi (2005) administered some subtests of the TAM-2 and of its Spanish version, the THAM-2, to children from 11 to 14 years, to examine their MLA in their two main languages. Performances on the two versions of the test reflected the imbalance between the L1 and the L2, as expected, but showed acceptable levels also in the less known language, Italian, in this case. In Spain, the first researcher who used the Spanish THAM-2 was Lasagabaster, in the Basque country (1998, 2001), to examine the role of MLA, on the one hand as an outcome of Spanish-Basque bilingualism and, on the other hand, as a predictor of a series of abilities in

English. In a first study (1998), conducted with 5th and 8th graders, results showed that MLA performances in Spanish were significantly better in the trilingual education model (Spanish, Basque and English) than in less demanding educational models. In a further study (2001), the outcomes obtained from the THAM-2 in the previous study were processed as predictors of English linguistic abilities. The results confirmed the predictive power of metalinguistic abilities as expressed by the THAM-2. Perales and Cenoz (2002) also used parts of the Spanish THAM-2 and THAM-3 as predictors of the ability to learn Basque in adult monolinguals whose first language was Spanish. Among all the independent variables (age, gender, exposition to Basque, motivation, etc.) MLA was by far the most effective predictor. Again, in Spain, De Haro, Delgado and López (2012) showed that MLA, as measured by all the subtests of the THAM-2, is a predictor of text production skills in 7th graders.

Research with the TAM-3 started later, during the 2000s, associated with metacognition, bilingualism, and educational aspects. The relationship between the TAM-3 and metacognition was explored in high school and university students (Di Santo, Pinto, Iliceto & Melogno, 2007; Pinto, Iliceto & Melogno, 2012), where participants were administered the Progressive Matrices SPM38 (Raven *et al.* 1998) and the TAM-3. After the usual, nonverbal administration of the Matrices, participants were asked to justify their choices of the figures of the test, which is a highly demanding metacognitive task. Very strong correlations were found between the ability to verbally justify a nonverbal answer and the metalinguistic abilities measured by the TAM-3, which highlights the metacognitive nature of the TAM-3.

The relationship between the TAM-3 and bilingualism was explored by means of a series of studies (Pinto, 2011; Pinto, El Euch, Lombardo, Caucci & Iliceto, submitted; Pinto, Trusso & Bevilacqua, 2004; Pinto, Trusso & Kristiansen, 2002) based on designs where multiple comparisons were possible: a) bilinguals'performances on the TAM-3 versus those of each monolingual group: b) within the bilingual group, performances as a function of the linguistic version chosen; and c) within the bilingual group, as a function of the distinction between consecutive and simultaneaous bilinguals. These more complex designs were made possible due to the existence of various linguistic versions of the test (in the above studies, the English and the French ones), which also made possible more in-depth analyses of the results. As in the studies conducted on bilinguals with the TAM-2, metalinguistic superiority in bilinguals systematically appeared, but this general gap could be better interpreted when considering each of the monolingual groups. In some studies (Pinto, El Euch, Lombardo, Caucci, & Iliceto, submitted; Pinto, Trusso & Bevilacqua, 2004), the gap with the bilinguals was more reduced in one of the monolingual groups, while in another study (Pinto, Trusso & Kristiansen, 2002), the two monolingual groups performed at very similar levels. While there were considerable differences between bilinguals and monolinguals on the metalinguistic level, (ML scores) there were no differences between them at the linguistic level (L scores). Differences, however, were found among simultaneous and consecutive bilinguals, but only moderate, in favour of the former.

The administration of the TAM-3 was also considered with an interactive modality in a focus group with three participants and a tutor expert in this test (Pinto & Micale, 2014). The three participants had been previously assessed with the whole test and then selected on the basis of the dominant ML level shown in their performances: 75% of ML0 answers, or 75% of ML1 answers or 75% of ML2 answers. During the interaction, participants reconsidered the answers they had given individually and commented those given by the others. Differences in their points of view triggered lively discussions, which resulted in more advanced MLA in the participants who were initially weaker (ML0 and ML1) and in more refined and sharper formulations in the participant with the highest level (ML2). The pedagogical value of these discussions in enhancing MLA has stimulated the creation of didactic videos on the interactive use of the metalinguistic ability test n.2 and 3. These videos are part of the educational resources offered by the MATEL project and, as such, they are accessible on the MATEL website (www.matelproject.com). These are, respectively, the videos based on the interactive use of the Italian TAM-2 (Micale, Bracone, & Pinto), on the Spanish THAM-2 (Núñez Delgado & Santamarino Sancho), on the Italian TAM-3 (Micale & Pinto), on the French THAM-3 (Monette & El Euch) and on the Spanish THAM-3 (Merino & Lasagabaster).

In Canada, as well as in Italy, research on MLA has been carried out with several linguistic versions of the metalinguistic ability test n.3 by El Euch (2010). A group of young adult French-English bilinguals and of French-English-Spanish trilinguals were tested with the English version of the test, the MAT-3 (Pinto, Titone & Trusso, 1999), and the Spanish version, the THAM-3 (Pinto, Titone & González Gil, 1999). Beyond differences in the linguistic versions of the tests and in the number of languages known by the participants, no differences in MLA appeared, a result that the author interpreted in light of Cummins' hypothesis of a Common Underlying Proficiency (Cummins, 2000). El Euch (2010, 2015) also found that the learners' attitudes and motivation towards their additional languages (English and Spanish) had no effect on their MLA in these languages. She concluded that this cognitive ability is not sensitive to such affective factors as attitudes and motivation. El Euch (2012) also examined the question of whether MLA, through the use of the French THAM-3 (Pinto & El Euch, 2015), affects reading proficiency in all the languages of a bilingual or a plurilingual individual. She found significant correlations between MLA and reading in the first language and the second language of the participants in the study. However, there was no contribution of MLA in reading achievement in their third language. El Euch concluded, in line with Cummins' (2000) threshold hypothesis, that reaching a threshold level of proficiency in a language might be necessary for a correlation MLA-reading to come up. El Euch, Pinto and Ostiguy (2015) found the same results as far as writing is concerned. They came to the same conclusions when they examined the question of whether MLA, through the use of the French THAM-3, affects writing proficiency in all the languages of a bilingual or a plurilingual individual.

We hope that the validations of the French THAM-3, of the Spanish THAM-2 and THAM-3, the

presence of the German MKT-2 (Jessner, Hofer & Pinto, 2015) and MKT-3 (Jessner, Pellegrini Moroder, Hofer & Pinto, 2015), the encouraging results obtained with these German versions (Candilera, Iliceto, Hofer, Pellegrini & Pinto, 2015), and the insights we have gained through the interactive use of these metalinguistic tests will significantly enrich our understanding of the construct of MLA and of its pedagogical applications.

REFERENCES

CANDILERA, G., ILICETO, P., HOFER, B., PELLEGRINI, C., PINTO, M.A. (2015).The pilot studies on the MKT-2 (Metalinguistischer Kompetenztest Teil 2, *Jessner, Hofer & Pinto,* 2015) and the MKT-3 (Metalinguistischer Kompetenztest Teil 3, *Jessner, Pellegrini, Moroder, Hofer & Pinto*, 2015). *RIVISTA DI PSICOLINGUISTICA APPLICATA/JOURNAL OF APPLIED PSYCHOLINGUISTICS*, Special Issue, ed. by M.A.Pinto: *The "MATEL"(Metalinguistic Awareness Tests in European Languages) Project: Validations and translations"*. XV, 2.

CUMMINS, J. (2000). *Language, power and pedagogy*. Clevedon: Multilingual Matters.

DE HARO, E. F., DELGADO, M. P. N., & LÓPEZ, A. R. (2012). Consciencia metalingüística y enseñanza-aprendizaje de la composición escrita en educación primaria: Un estudio empírico. *Rivista di Psicolinguistica Applicata, 12*(1-2), 25-46.

DI SANTO, A., ILICETO, P., PINTO, M.A., MELOGNO, S. (2007). Metacognitive processes in justifying the solutions of the Raven's Spm38 Test and metalinguistic abilities. Studies on high school students, *Lifespan and disability*, Vol 10, 1, 105-132.

EL EUCH, S. (2010). Attitudes, motivations et conscience métalinguistique chez des bilingues et des trilingues adultes: effets, similarités et différences. *Language Awareness,* 19 (1), 17-33.

EL EUCH, S. (2012). Effect of metalinguistic awareness on reading in the first language, second language and third language of university students. *INTED2012 Proceedings, ISBN: 978-84-615-5563 5.*

EL EUCH, S. (2015). Attitudes, motivations et conscience métalinguistique: le rôle de l'affectif dans le cognitif *MELT Proceedings: The successful language learner* (pp. 242-261). Montréal: Université du Québec à Montréal, École de langues.

EL EUCH, S., PINTO, M. A., & OSTIGUY, L. (2015). *Metalinguistic awareness and multi-literacy in undergraduates: Does time matter?* Paper presented at the Joint AAAL and CAAL conference, Toronto.

HOFER, B. (2013). *On the Dynamics of Early Multilingualism. A Psycholinguistic Study of Bilingual Education in South Tyrol.* (Unpublished doctoral dissertation). University of Innsbruck, Innsbruck Austria.

HOFER, B. (2015). *On the Dynamics of Early Multilingualism. A Psycholinguistic Study of Trilingual Education in South Tyrol*. Berlin/Boston: Mouton de Gruyter.

JESSNER, U. (2006). *Linguistic awareness in multilinguals. English as L3.* Edinburgh, Edinburgh Press.

JESSNER, U., HOFER, B., PINTO, M.A., *(2015). MKT, Metalinguistischer Kompetenztest Teil 2.* Innsbruck, Studia Universitätsverlag. Innsbruck.

JESSNER, U., PELLEGRINI, C., MORODER, V., HOFER, B., & PINTO, M.A (2015*). MKT. Metalinguistischer Kompetenztest Teil 3*. Innsbruck. Studia Universitätsverlag Innsbruck.

LASAGABASTER, D. (1998). The threshold hypothesis applied to three languages in contact at school. *International Journal of Bilingual Education and Bilingualism*, 1 (2), 119-133.

LASAGABASTER, D. (2001). The effect of knowledge about the L1 on foreign language skills and grammar. *International Journal of Bilingual Education and Bilingualism,* 4 (5), 310-331.

MELOGNO, S., PINTO, M.A., LEVI, G. (2014). Profile of the linguistic and metalinguistic abilities of a gifted child with autism spectrum disorder: A case study. *Child Language Teaching and Therapy,* 1-14.

MERINO, J.A., & LASAGABASTER, D. (2015). *Utilización interactiva del Test de Habilidades metalingüísticas-3 (THAM-3) (Pinto, Titone, Gonzáles Gil, 2000; Lasagabaster, Merino, Pinto, 2015) Vídeo didáctico sobre un grupo de discusión.* www.matelproject.com.

MICALE, F., BRACONE. I, PINTO, M.A. (2015). *Video didattico sull'uso interattivo del TAM-2 (Pinto, Candilera, Iliceto, 2003)* www.matelproject.com.

MICALE, F., & PINTO, M.A.(2015). *Video didattico sull'uso interattivo del TAM-3 (Pinto & Iliceto, 2007)* www.matelproject.com.

MONETTE, I., & EL EUCH, S. (2015). *Livret d'accompagnement de la vidéo didactique visant le développement de la conscience métalinguistique (CML) à partir d'items du THAM-3 (Pinto & El Euch, 2015).* www.matelproject.com.

NÚÑEZ DELGADO, P., & SANTAMARINO SANCHO, M. (2015). *Utilización interactiva del Test de Habilidades metalingüísticas-2 (THAM-2) (Pinto, Titone, Gonzáles Gil, 2000; Núñez Delgado y Pinto, 2015). Vídeo didáctico sobre un grupo de discusión.* www.matelproject.com.

PAFUMI, M.L. (2005). La comprensión de metáforas en alumnos bilingües español-italiano años : un estudio translingüístico en Argentina. *RIVISTA DI PSICOLINGUISTICA APPLICATA*, V, 1-2, 59-80.

PERALES, J., & CENOZ, J. (2002). The effect of individual and contextual factors in adult second-language acquisition in the Basque country. *Language, Culture and Curriculum*, 15 (1), 1-15.

PIAGET, J. (1975). *L'équilibration des structures cognitives.* Paris, P.U.F.

PINTO, M.A. (a cura di) (1995). *La consapevolezza metalinguistica. Teoria, sviluppo e strumenti di misurazione.* Numero speciale di *Rassegna Italiana di Linguistica Applicata*.

PINTO, M.A. (1999). *La consapevolezza metalinguistica. Teoria, sviluppo e strumenti di misurazione.* Pisa-Roma. Istituti Editoriali e Poligrafici Internazionali.

PINTO, M. A. (2002). Bilinguismo e sviluppo metalinguistico. in G. Di Stefano & R. Vianello (a cura di), *Psicologia dello sviluppo e problemi educativi. Studi e ricerche in onore di Guido Petter*. Firenze: Giunti.
PINTO, M.A. (2011). Long-term effects of early bilingualism on metalinguistic awareness: a study on young adults. Paolo Valore Ed., *Multilingualism. Language, Power, and Knowledge*. Pisa, Edistudio.

PINTO, M.A., CANDILERA, G., & ILICETO, P. (2003). *TAM-2. Test di abilita`metalinguistiche n.2 (9-14 anni). La valutazione dello sviluppo metalinguistico tra scuola elementare e scuola media*. Roma: Scione Editore.

PINTO, M.A., & EL EUCH, S. (2015). *La conscience métalinguistique. Théorie, développement et instruments de mesure*. Québec, P.U.L.

PINTO, M.A., EL EUCH, S., LOMBARDO, M.C., CAUCCI, M., ILICETO, P, (submitted). Metalinguistic abilities in Italian-English and Italian-French young adult bilinguals: A comparison with English, French and Italian monolingual controls.

PINTO, M.A., FULGENZI, D. (2014). *QUASI UN BIG BANG ! Potenziare le abilità di riflessione sulla lingua.* Valore italiano. Lilamé.

PINTO, M. A., ILICETO P. (2007). *TAM-3 Test di abilità metalinguistiche n° 3*, *Fascia adolescente adulta.* Roma: Carocci Faber.

PINTO, M.A., ILICETO, P., MELOGNO, S. (2012). Argumentative abilities in metacognition and in metalinguistics: a study on university students, *European Journal of Psychology of Education*, 27: 35 58.

PINTO, M.A., & MELOGNO, S. (2014). *Lo sviluppo metalinguistico. Modelli teorici, strumenti e applicazioni cliniche*. Firenze: SEID Editori.

PINTO, M.A., MICALE, F. (2014). Promoting Metalinguistic Awareness in group discussions. A study with university students. *Paper presented at the 2014 Annual ACLA/CAAL Conference. Brock University.* May, 26-28, 2014.

PINTO, M.A., TITONE, R. (1989). Uno strumento di misurazione delle abilità metalinguistiche : il T.A.M. *Rassegna Italiana di Linguistica Applicata*, Vol. XXI. N.1-2. pp 59-128.

PINTO, M. A., TITONE, R. & GONZÁLES GÍL, M.D. (1999). *La consciencia metalingüística. Teoría, desarrollo e instrumentos de medición.* Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali.

PINTO, M. A., TITONE, R. & TRUSSO, F. (1999). *Metalinguistic Awareness. Theory, Development and Measurement Instruments.* Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali.

PINTO, M.A., TRUSSO, F., & BEVILACQUA, A. (2004). Bilingualism in university students. Further evidences of metalinguistic benefits. *Rivista di Psicolinguistica Applicata*, 11 (2), 77-90.

PINTO, M.A., TRUSSO, F., & KRISTIANSEN, K. (2002). Metalinguistic abilities in Italian-English adult bilinguals. A comparison with Italian-speaking and English-speaking monolinguals. *Rivista di Psicolinguistica Applicata,* II (2-3), 77-90.

RAVEN, J., RAVEN, J.C., & COURT, J.H. (1998a). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview*. San Antonio, TX: Harcourt Assessment.

THE VALIDATION OF THE THAM-2
(TEST DE HABILIDADES METALINGÜÍSTICAS N. 2, *Núñez Delgado & Pinto, 2015*)

GABRIELLA CANDILERA*, PAOLO ILICETO**,
PILAR NÚÑEZ-DELGADO***, MARIA ANTONIETTA PINTO****

S&P. Statistics & Psychometrics Ltd
Universidad de Granada
Università di Roma "SAPIENZA"

ABSTRACT:

The present study reports the results of the validation of a Spanish Metalinguistic Ability Test, the THAM-2 (*Test de Habilidades Metalingüísticas n. 2)* for children aged from 9 to 14, translated and partially adapted from the Italian original version of this test, the TAM-2 (Pinto & Titone, 1989; Pinto, Candilera & Iliceto 2003;). The test measures metalinguistic awareness (MLA) at two levels, intuitive and explicit. These two levels are embodied in two types of scores named L (acronym for Linguistic ability factor) and ML (acronym for Metalinguistic ability factor), and are computed for each of the 6 subtests included in the THAM-2 (Comprehension, Synonymy, Acceptability, Ambiguity, Grammatical function, Phonemic segmentation). Three hundred and one Spanish-speaking children between 9 and 16 (mean age = 12.5; SD = 1.3) participated in the study, which was conducted in the city of Granada Spain) and surroundings. Based on the Spanish school system, the sample was subdivided into four grades from 4-5 primaria (4th and 5th grade) to 2° ESO (8th grade) (Table 1).

We presented descriptive statistics in terms of means and standard deviations for the scores of each subtest and for the totals of the L and ML of the test, and in terms of percentages of L and ML score levels (Table 2, Figures 1 and 2). Skewness and Kurtosis, as measures of the shape of the distributions, were calculated. Only a few values in some of the L scales exceeded the conventional criterion for normality, contrary to the rest of the data which were normally distributed (Table 3). Therefore, the sample was considered normative, and the standardization of the test could be developed. Group differences in the sample were also studied as a function of two main factors: education level (school grades) and gender (males/females). Factorial analyses of variance (ANOVA) with a 4 x 2 design were computed: education (4-

* Corresponding author: Gabriella Candilera, S&P Statistics & Psychometrics Ltd.Via Tuscolana, 458. 00181. Roma. E-mail: gabriellacandilera@libero.it.
** Paolo Iliceto, S&P, Statistics & Psychometrics Ltd. E-mail: paolo.iliceto@statpsyc.com
***. Pilar Núñez-Delgado. Universidad de Granada. E-mail: ndelgado@ugr.es
**** Maria Antonietta Pinto. Università di Roma "SAPIENZA". E-mail; mariantonietta.pinto@uniroma1.it

5 primaria/6 primaria/1° ESO/2° ESO). Both factors significantly affected the performances since we observed a marked gender superiority of the females in all the totals of the test, and a definite incremental pattern associated to the school grades. Inter-rater reliability measured by Cohen's Kappa (Cohen, 1960) was highly satisfactory; the Cronbach's alpha coefficients presented acceptable internal consistency estimates for the ML scales. Convergent validity, assessed by means of Pearson's correlations between all the THAM-2 scales and SPM38 scores (Raven, Raven & Court, 1998), showed several significant correlations, although moderate (Table 4). The intercorrelations between the THAM-2 ML scores were statistically significant, although moderate, whereas the intercorrelations between the L scores were weaker (Table 5).

Confirmatory Factor Analyses (CFA, Brown, 2006) was performed to study construct validity. The results supported the theoretical two-factor structure represented by Linguistic (L) and Metalinguistic ability ML) factors - an expression of MLA at the intuitive and the explicit level, respectively - providing a good fit to the empirical data. However, this result remains theoretically problematic. While the ML factor emerged as a unitary construct, well measured by the all of the ML scores, the L factor appeared weaker, well measured by only two scores (Synonymy-L and Grammatical function-L).

Despite the limitations that have been pointed out above, the THAM-2 can be considered a valid metalinguistic ability test that can be reliably used in all Spanish-speaking contexts, in and out of Spain, by 9 to 14 year-old subjects for whom Spanish is their first language. Given the wide age range addressed by this test and the richness of its contents, many avenues of research can be envisaged both in the psychometric and the educational areas.

From the psychometric point of view, the sample should be enlarged so as to include bigger groups with a similar number of subjects at each age level. A more balanced stratification of the groups would permit to calculate more precise group norms for monolingual Spanish-speaking children. These norms, in turn, would enable researchers and educators to make more reliable comparisons with children from officially bilingual regions in Spain. In addition, as the THAM-2 comprises a variety of metasemantic, metagrammatical and metaphonological subtests, the study of convergent validity could be pursued by correlating the corresponding metalinguistic performances to a variety of more basic semantic, grammatical and phonological competences, as measured by other tests. To this end, contrasted groups characterized by low and high performances in these language tests could be created. The pattern of associations that might result from these studies are likely to shed some light on the still problematic status of the L dimension in the THAM-2.

From the educational point of view, the specific metalinguistic character of the THAM-2, based on the analysis of the MLA at the most complex level, i.e. explicit and argumented, makes it a relevant tool for the assessment of metacognitive competences in language domains at school, such as text comprehension

and production, as De Haro, Núñez Delgado, & López (2012) have pointed out.

REFERENCES

BROWN, T.A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York, NY: Guilford.

COHEN, J. (1960). A coefficient of agreement for nominal scales. *Education and Psychological Measurement, 20*, 37-46.

DE HARO, E. F., DELGADO, M. P. N., & LÓPEZ, A. R. (2012). Consciencia metalingüística y enseñanza-aprendizaje de la composición escrita en educación primaria: Un estudio empírico. *Rivista di Psicolinguistica Applicata, 12*(1-2), 25-46.

NÚÑEZ DELGADO, P., PINTO, M.A. (2015). *THAM-2. TEST DE HABILIDADES METALINGÜÍSTICAS N.2 (9-14 AÑOS)*. www.matelproject.com.

PINTO, M.A., CANDILERA, G., & ILICETO, P. (2003). *TAM-2. Test di abilità metalinguistiche n. 2*. Roma: Scione.

PINTO, M.A., & TITONE, R. (1989). Uno strumento di misurazione delle abilità metalinguistiche: il T.A.M. *Rassegna Italiana di Linguistica Applicata*, Vol. XXI. N.1-2. pp 59-128.

PINTO, M.A., TITONE, R., GONZÁLES GIL, M.D. (2000). *La consciencia metalingüística. Teoría, desarrollo e instrumentos de medición*. Pisa-Roma. Istituti Editoriali e Poligrafici Internazionali.

RAVEN, J., RAVEN, J.C., & COURT, J.H. (1998a). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 1: General Overview*. San Antonio, TX: Harcourt Assessment.

Tables and Figures

Table 1. Distribution of the participants by school grade and gender				
School grade		**Gender**		
		Males	**Females**	**Total**
4-5° primaria	N	22	31	53
	%	15,3%	19,7%	17,6%
6° primaria	N	51	43	94
	%	35,4%	27,4%	31,2%
1° ESO	N	23	34	57
	%	16,0%	21,7%	18,9%
2° ESO	N	48	49	97
	%	33,3%	31,2%	32,2%
Total	N	144	157	301
	%	100,0%	100,0%	100,0%

Table 2. Descriptive statistics for all the scores of the THAM-2 (Núñez Delgado & Pinto, 2015) vs the TAM-2 (Pinto, Candilera, & Iliceto, 2003)

Subtests	Mean (Spanish sample)	SD (Spanish sample)	Mean (Italian sample)	SD (Italian sample)
Comprehension L (15)	13.67	1.24	14.31	1.11
Comprehension ML (32)	8.92	5.61	12.86	5.07
Synonymy L (5)	3.83	1.14	4.47	.66
Synonymy ML (10)	4.17	3.04	2.47	1.99
Acceptability L (28)	25.03	2.50	25.71	2.69
Acceptability ML (26)	7.72	4.67	9.42	4.52
Ambiguity L (7)	5.84	1.11	5.68	1.10
Ambiguity ML (14)	4.91	2.31	5.32	2.25
Gr Function L (6)	5.46	1.41	5.76	.49
Gr Function ML (24)	4.30	3.83	8.34	3.47
Ph. Segmentation L (33)	25.65	5.12	27.25	5.82
Ph.Segmentation ML (24)	13.18	4.00	10.22	2.69
THAM-2 total L (94)	79.49	7.57	82.78	8.83
THAM-2 total ML (130)	43.20	15.50	48.63	15.14
THAM-2 total score (224)	122.68	20.95	131.41	21.64

Table 3. Symmetry and shape of the distribution of the THAM-2 (Núñez Delgado & Pinto, 2015) vs the TAM-2 (Pinto, Candilera & Iliceto, 2003)

Subtests	Skewness (Spanish sample)	Kurtosis (Spanish sample)	Skewness (Italian sample)	Kurtosis (Italian sample)
Comprehension L	-1.016	.890	-2.120	6.025
Comprehension ML	.817	.986	-.180	-.473
Synonymy L	-1.733	3.697	-1.647	6.043
Synonymy ML	.085	-1.200	.576	.031
Acceptability L	-2.059	8.089	-1.505	2.362
Acceptability ML	.363	-.562	.049	-.712
Ambiguity L	-1.112	1.638	-.931	1.145
Ambiguity ML	.249	.779	.015	.005
Gr.Function L	-3.119	9.137	-2.117	4.464
Gr. Function ML	.757	.120	-.383	-.397
Ph.Segmentation L	-1.194	1.144	-1.578	3.752
Ph.Segmentation ML	-.875	.297	-.293	.885
THAM-2 total L	-1.082	1.235	-1.509	3.107
THAM-2 total ML	.158	.296	-.091	-.402
THAM-2 total score	-.270	.171	-.634	.576

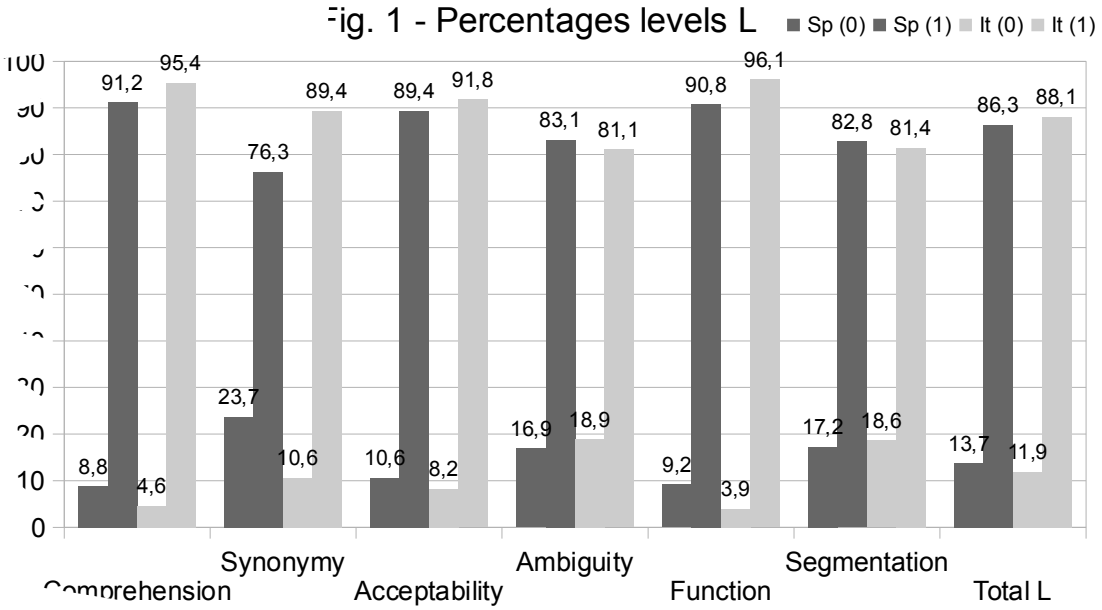

Fig. 1: Percentages of L score levels (0,1), in each subtest, and in the total THAM-2 (Núñez Delgado, *&* Pinto, 2015)

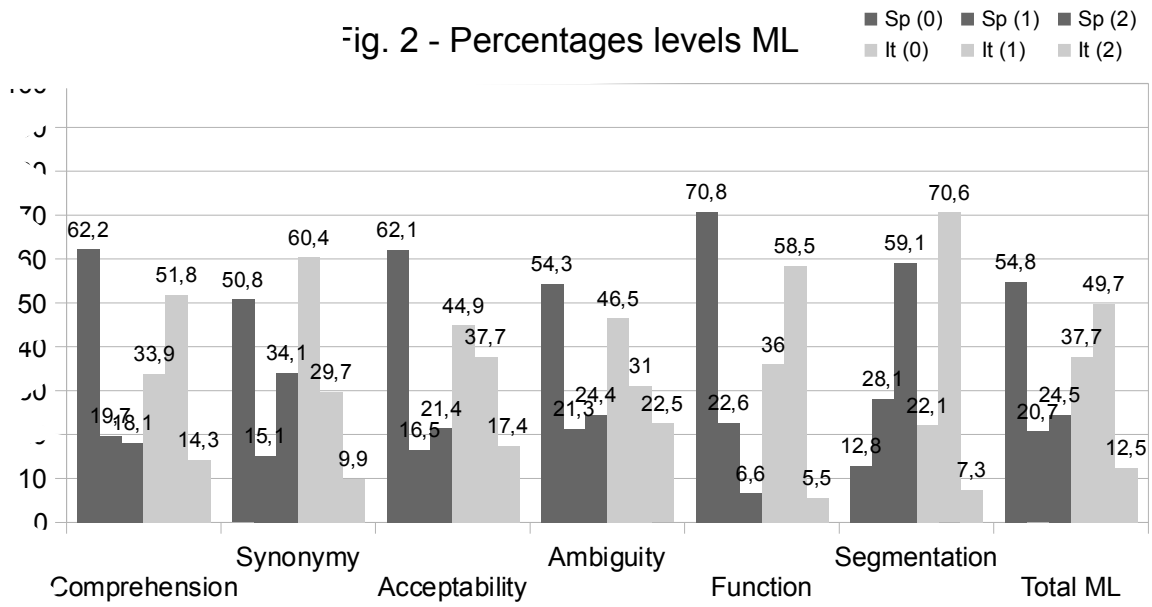

Fig. 2: Percentages of ML response levels (0,1,2) in each subtest and in the total THAM-2 (Núñez Delgado & Pinto, 2015)

Table. 4. Correlations matrix between the THAM-2 (Núñez Delgado & Pinto, 2015) and the SPM38 (Raven, Raven & Court, 1998)

Subtests	SPM38
Comprehension L	.164*
Comprehension ML	.209**
Synonymy L	.212**
Synonymy ML	.151*
Acceptability L	.085
Acceptability ML	.138
Ambiguity L	.069
Ambiguity ML	.127
Gr.Function L	-.031
Gr.Function ML	.027
Ph.Segmentation L	.308***
Ph.Segmentation ML	.302***
THAM-2 total L	.310***
THAM-2 total ML	.250***
THAM-2 total score	.295***

* p <.05; ** p <.01; *** p <.001

Table 5 THAM-2. (Núñez Delgado & Pinto, 2015) Inter-correlations matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1) Comprehension L	1													
2) Comprehension ML	.257**	1												
3) Synonymy L	.033	-.155*	1											
4) Synonymy ML	.155*	.269**	.432**	1										
5) Acceptability L	.034	.195*	.092	.145*	1									
6) Acceptability ML	.073	.386**	-.022	.182*	.310**	1								
7) Ambiguity L	.041	.248**	-.006	-.012	.094	.259**	1							
8) Ambiguity ML	-.006	.177*	.196*	.182*	.208**	.315**	.384**	1						
9) Gr.Function L	.034	-.145*	.587**	.276**	.117*	-.045	.009	.140*	1					
10) Gr.Function ML	.025	.354**	.266**	.385**	.212**	.331**	.149*	.326**	.246**	1				
11) Ph.Segmentation L	.165*	.308**	.236**	.266**	.214**	.347**	.214**	.361**	.115*	.271**	1			
12) Ph.Segmentation ML	.176*	.297**	.334**	.366**	.280**	.271**	.159*	.355**	.367**	.385**	.485**	1		
13) THAM-2 total L	.305**	.301**	.456**	.368**	.531**	.375**	.330**	.424**	.398**	.366**	.862**	.592**	1	
14) THAM-2 total ML	.196*	.723**	.204**	.566**	.349**	.677**	.301**	.516**	.165*	.699**	.515**	.668**	.601**	1
15) THAM-2 total score	.256**	.643**	.315**	.552**	.450**	.636**	.342**	.535**	.266**	.650**	.693**	.708**	.807**	.957**

THE VALIDATION OF THE THAM-3
(TEST DE HABILIDADES METALINGÜÍSTICAS N. 3, *Lasagabaster, Merino & Pinto,* 2015)

GABRIELLA CANDILERA*, PAOLO ILICETO**, DAVID LASAGABASTER***,
JON ANDER MERINO****, MARIA ANTONIETTA PINTO*****

S&P Statistics & Psychometrics Ltd
University of the Basque Country UPV/EHU
Università di Roma "SAPIENZA"

ABSTRACT:

This study describes the statistical analyses performed to validate the THAM-3 (*Test de Habilidades Metalingüísticas n.3*), a metalinguistic test aimed at late adolescence and adulthood, translated from the original Italian TAM-3 (Pinto & Iliceto, 2007) into Spanish (Pinto, Titone, & Gonzáles Gil, 2000; Lasagabaster, Merino, & Pinto, 2015) with partial adaptations. The test is composed of three subtests, Comprehension (*metasemantic* task), Acceptability *(metagrammatical* task) and Figurative language (*metasemantic* task), and the scoring system is based on the distinction between Linguistic scores (L, measuring metalinguistic awareness at the implicit level) and Metalinguistic scores (ML, measuring metalinguistic awareness at the explicit level).

One hundred and fifty university students aged 18 to 37 years (mean age: 21; SD = 2.7) were recruited at the University of Valladolid (Soria-Spain), all enrolled in the Teaching Training undergraduate program and sharing the same education background. They were administered the THAM-3, as a measure of metalinguistic abilities, and the Raven's SPM38 (Raven, Raven & Court, 1998), as a measure of nonverbal intelligence which shares relevant metacognitive components with the metalinguistic test. Skewness and Kurtosis values as measures of the shape of the distributions, showed that the data were normally distributed (Table 2).

The distribution of percentages of score levels in the total L scores (0, 1) highlighted some difficulties in metalinguistic awareness at the intuitive level (Fig.1). Although correct answers prevailed over the incorrect ones, their percentages were very close, contrary to what happened in the Italian sample where the gap was largely in favour of the correct answers. The participants' performance on the Acceptability subtest contributed heavily to this overall result, as half of the answers in this subtest were incorrect.

---

* Corresponding author: Gabriella Candilera, S&P Statistics & Psychometrics Ltd, e-mail: Via Tuscolana, 458. 00181. Roma. E-mail: gabriellacandilera@libero.it.

** Paolo Iliceto, S&P Statistics & Psychometrics Ltd, e-mail: paolo.iliceto@statpsyc.com
*** David Lasagabaster, University of the Basque Country UPV/EHU, e-mail : david.lasagabaster@ehu.eus
**** Jon Ander Merino, University of the Basque Country UPV/EHU, e-mail : jon.merinovillar@gmail.com
***** Maria Antonietta Pinto, Università di Roma "SAPIENZA", e-mail: mariantonietta.pinto@uniroma1.it

This pattern was confirmed at the more abstract level of metalinguistic awareness, as shown by the distribution of percentages of the ML scores (Fig. 3). Again, the Acceptability subtest revealed the most evident difficulty, particularly at the highest level of answer, a pattern also observed in the Figurative Language subtest. Overall, only Comprehension results were comparable to those of the Italian sample in both L and ML scores. No differences were found regarding gender, except for the Acceptability ML score in favour of female students, which indicates a fundamental homogeneity of the sample (Table 3).

Inter-raters' agreement, as measured by Cohen's Kappa (Cohen, 1960), was satisfactory in the majority of cases, ranging from moderate to perfect, with the exception of some cases were it was found to be low. The internal consistency of each L and ML scale was assessed by means of Cronbach's Alpha coefficients. One out of the three L scales (Acceptability) and two out of the three ML scales (Acceptability and Figurative language) yielded satisfactory results.

The correlations between all the THAM-3 scales and SPM38 scores (Table 4) as a measure of convergent validity showed moderately significant associations between the performances at the nonverbal intelligence test and the performances at the explicit level of metalinguistic awareness, i.e. in the ML scores. These results highlight the metacognitive character of the metalinguistic measures. The inter-correlations between the THAM-3 scales (Table 5) ranged from low to strong.

CFA (Brown, 2006) results supported the theoretical two-factor structure of the original test, assigning the ML subtests' scores to the corresponding Metalinguistic latent factor, the L subtests' scores to the corresponding Linguistic latent factor, and explaining the inter-correlations between the observed variables. The final model provided a good fit to the empirical data. However, we must point out that some differences were brought to light in the two main factors: while the Metalinguistic ability factor (which expresses metalinguistic awareness at the explicit level) emerged as an unitary construct, well measured mainly by the Comprehension scale, the Linguistic ability factor (which expresses metalinguistic ability at the intuitive level**)** resulted in a weaker construct with only two weak observed variables retained in the analysis. These results call for further research into this issue.

Overall, on the basis of the results reported above, the THAM-3 can be considered a valid and reliable test for all speakers – late adolescents or adults – of Spanish as their first language, both in Spain and in other Spanish-speaking countries all over the world, as well as for students of Spanish as a second, foreign or additional language.

Many possibilities can be explored by future research. For instance, as the test is foreseen for a large age range beyond late adolescence, the normative sample could be considerably extended in such a way as to include older participants, as it has been done with the original Italian test. Some structural points of the test could be reconsidered, as highlighted by the evident difficulties shown in the Acceptability subtest, but also in the ML part of the Figurative language subtest. The coding system of the former could be less strict and

consider the very correction of the error as the first step of the ML score, as it expresses more elaborated processes than the mere detection of the error. In the latter, as most of the difficulties appeared in the interpretation of the poetic verses, other verses could be included, more adapted to the cultural characteristics of the targeted participants. The whole issue of the relative weakness of the L factor could be better reconsidered in light of these changes. Convergent validity, which at the moment, is only measured by means of correlations with a nonverbal intelligence test, the SPM38, could be further studied through correlations with other language tests measuring basic semantic and grammatical proficiency. This could clarify to what extent the "meta" dimension needs to be grounded on basic language proficiency.

Furthermore, as there are currently three versions of the test in three Romance languages (namely Italian, French and Spanish) which have already been validated, and two in another linguistic family (i.e. English and German) still to be validated, comparative studies could be conducted. The results of these comparative studies are likely to illuminate core aspects of metalinguistic awareness at the wide age range aimed at by the THAM-3, and also help to shed light on the possible differences to be detected in different languages and diverse educational systems around the globe.

REFERENCES

BROWN, T.A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York, NY: Guilford.

COHEN, J. (1960). A coefficient of agreement for nominal scales. *Education and Psychological Measurement, 20*, 37-46.

LASAGABASTER, D., MERINO, J.A., & PINTO, M.A. (2015). *Test de habilidades metalingüisticas para adolescentes y adultos: THAM-3*. Bilbao: Universidad del País Vasco/Euskal Herriko Unibertsitatea. www.matelproject.com.

PINTO, M. A., & ILICETO P. (2007). *TAM-3 Test di abilità metalinguistiche n° 3*, *Fascia adolescente adulta.* Roma: Carrocci Faber.

PINTO, M. A., TITONE, R. & GONZÁLEZ GIL, M.D. (2000). *La consciencia metalingüística. Teoría, desarrollo e instrumentos de medición.* Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali.

RAVEN, J., RAVEN, J.C., & COURT, J.H. (1998a). *Manual for Raven's Progressive Matrices and Vocabulary Scales . Section 1: General Overview*. San Antonio, TX: Harcourt Assessment.

Tables and Figures

Table 1. Descriptive statistics for all the scores of the THAM-3 (Lasagabaster, Merino & Pinto, 2015) vs the TAM-3 (Pinto & Iliceto, 2007)

Scales	Means (Spanish sample)	SD (Spanish sample)	Means (Italian sample)	SD (Italian sample)
Comprehension L (4)	2.72	.80	3.07	.96
Comprehension ML (16)	5.52	2.41	5.48	3.14
Acceptability L (13)	6.42	2.21	9.93	2.85
Acceptability ML (26)	4.59	3.49	5.28	4.98
Figurative Lang L (2)	1.53	.65	1.37	.80
Figurative Lang ML (12)	3.72	1.89	5.57	3.16
THAM-3 total L (19)	10.67	2.56	14.37	3.40
THAM-3 total ML (54)	13.83	5.47	16.32	9.14
THAM-3 total score (73)	24.49	7.13	30.69	11.46

Table 2. Skewness and Kurtosis for the sample of the THAM-3 (Lasagabaster, Merino, & Pinto 2015) vs the sample of the TAM-3 (Pinto & Iliceto, 2007)

Scales	Skewness (Spanish sample)	Kurtosis (Spanish sample)	Skewness (Italian sample)	Kurtosis (Italian sample)
Comprehension L	-.155	-.440	.205	-.886
Comprehension ML	.239	-.346	-.177	.516
Acceptability L	.659	-.086	.480	-.648
Acceptability ML	1.010	.897	.054	.824
Figurative Language L	-1.056	-.024	-1.033	-.775
Figurative Language ML	.272	.215	-.812	.214
THAM-3 total L	.580	-.103	.134	-.555
THAM-3 total ML	.735	.596	-.012	.664
THAM-3 total score	.813	.862	-.126	.479

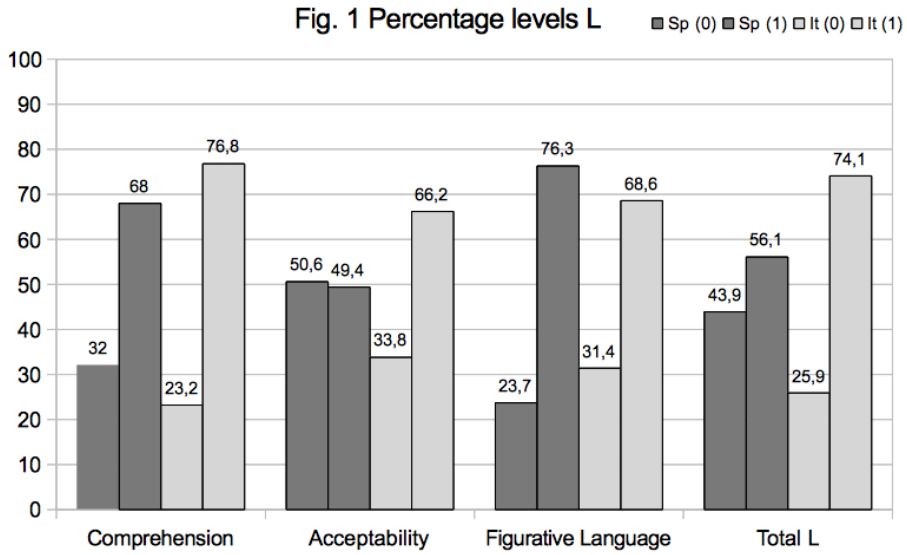

Figure 1. Percentage of L score levels (0, 1) in each subtest of the THAM-3 (Lasagabaster, Merino & Pinto, 2015) vs the TAM-3 (Pinto & Iliceto, 2007)

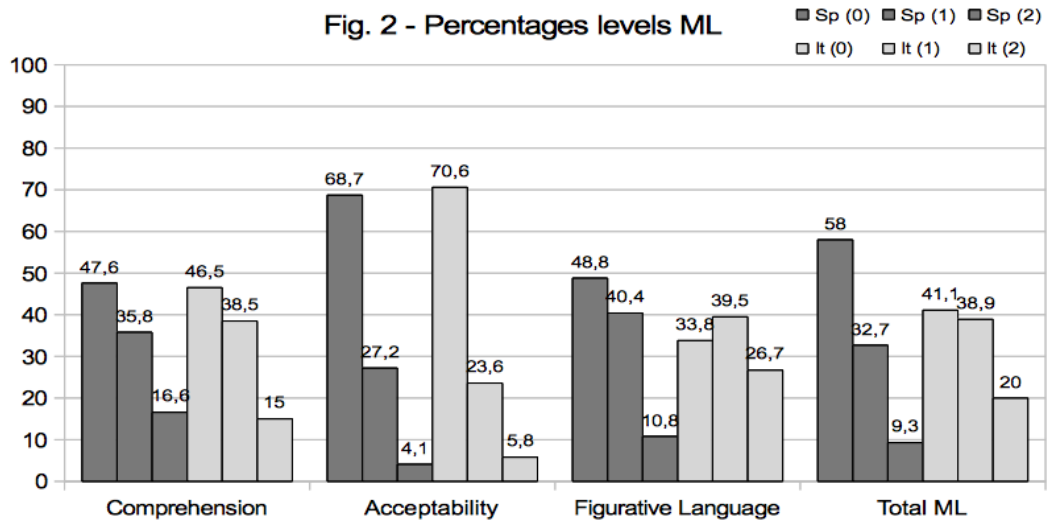

Figure 2. Percentage of ML levels (0, 1, 2) in each subtest of the THAM-3 (Lasagabaster, Merino & Pinto, 2015) vs the TAM-3 (Pinto & Iliceto, 2007).

Table 3. Student's t applied to the THAM-3 measures (Lasagabaster, Merino & Pinto, 2015). Comparisons between male and female students

	M (55)		F (95)			
	Means	SD	Means	SD	*t*(148)	p
Comprehension L	2.75	.84	2.71	.78	.29	.769
Comprehension ML	5.82	2.38	5.35	2.42	1.15	.251
Acceptability L	6.11	2.05	6.60	2.29	1.31	.192
Acceptability ML	3.62	3.17	5.15	3.56	2.63	.009
Figurative Language L	1.55	.63	1.52	.66	.26	.789
Figurative Language ML	3.98	1.61	3.57	2.02	1.29	.198
THAM-3 total L	10.40	2.57	10.82	2.56	.96	.335
THAM-3 total ML	13.42	5.16	14.06	5.66	.69	.489
THAM-3 total score	23.82	6.71	24.88	7.36	.88	.379
Bonferroni correction (α/9); p = .005						

Table 4. Correlations matrix between the THAM-3 (Lasagabaster, Merino & Pinto, 2015) and the Raven's SPM38

Scales	**SPM38**
Comprehension L	.056
Comprehension ML	.207**
Acceptability L	.108
Acceptability ML	.165*
Figurative Language L	.072
Figurative Language ML	.208**
THAM-3 total L	.129
THAM-3 total ML	.269***
THAM-3 total score	.253***
* p <.05; ** p <.01; *** p <.001	

Table 5. Inter-correlations matrix of the THAM-3 (Lasagabaster, Merino & Pinto, 2015).

	1	2	3	4	5	6	7	8
1) Comprehension L	1							
2) Comprehension ML	.369***	1						
3) Acceptability L	.134	.165*	1					
4) Acceptability ML	.104	.268***	.537***	1				
5) Figurative Language L	-.037	.170*	.060	.064	1			
6) Figurative Language ML	.014	.300***	.023	.088	.442***	1		
7) THAM-3 total L	.420***	.301***	.920***	.513***	.294***	.137	1	
8) THAM-3 total ML	.234***	.715***	.423***	.786***	.268***	.533***	.507***	1
9) THAM-3 total score	.331***	.658***	.656***	.788***	.312***	.459***	.749***	.951***
* p <.05; ** p <.01; *** p <.001								

THE VALIDATION OF THE THAM-3
(TEST D'HABILETÉS MÉTALINGUISTIQUES N° 3, Pinto & El Euch, 2015)

GABRIELLA CANDILERA*, PAOLO ILICETO**, SONIA EL EUCH***,
LUC OSTIGUY****, MARIA ANTONIETTA PINTO*****

S&P Statistics & Psychometrics Ltd
Université du Québec-à-Trois-Rivières
Università di Roma "SAPIENZA"

ABSTRACT

The present study describes the statistical processing used to validate the THAM-3 (*Test d'Habiletés Métalinguistiques n.3*, Pinto & El Euch, 2015), a metalinguistic ability test translated from the original Italian version, the TAM-3 (Pinto & Iliceto, 2007) into French, with partial adaptations. The test is composed of three subtests, Comprehension (*metasemantic* task), Acceptability *(metagrammatical* task) and Figurative language (*metasemantic* task), and the scoring system is based on the distinction between Linguistic scores (L, measuring metalinguistic awareness at the implicit level) and Metalinguistic scores (ML, measuring metalinguistic awareness at the explicit level). One hundred and fifty students aged from 19 to 52 years (mean age: 24.75; SD = 6.5) were recruited at the Université du Québec à Trois-Rivières and took the THAM-3 and the Raven's SPM38 (Raven, Raven & Court, 1998). The SPM38 is a measure of nonverbal intelligence that was used to examine convergent validity. The aim was to reach a satisfactory level of validation for the French version of the test.

Descriptive statistics measures, in terms of means and standard deviations (Table 1), and Skewness and Kurtosis values (Table 2) as measures of the shape of the distributions, showed that the data were normally distributed. The sample could therefore be considered as normative, which made it possible to develop the standardization of the test. The distribution of answers in the two L levels (0, 1) (Fig. 1) and in the three ML levels (0, 1, 2) (Fig. 2) showed satisfactory performances in the two metasemantic subtests, namely Comprehension and Figurative language, and less satisfactory performances in the metagrammatical subtest, Acceptability, a pattern already found in the Italian sample. No differences related to gender nor to education level were found – except for the Figurative

---

*Corresponding author: Gabriella Candilera, S&P Statistics & Psychometrics, Ltd; e-mail: gabriellacandilera@libero.it.
** Paolo Iliceto, S&P Statistics & Psychometrics Ltd, e-mail: paolo.iliceto@statpsyc.com
*** Sonia El Euch, Université du Québec à Trois-Rivières, e-mail: sonia.eleuch@uqtr.ca
**** Luc Ostiguy, Université du Québec à Trois-Rivières, e-mail: luc.ostiguy@uqtr.ca
***** Maria Antonietta Pinto, Università di Roma "SAPIENZA", e-mail: mariantonietta.pinto@uniroma1.it

language L score in favour of females – which indicates a fundamental homogeneity of the sample (Table 3).

Inter-raters' agreement, measured by Cohen's Kappa (Cohen, 1960), was fully satisfactory in the majority of cases, ranging from high to perfect, with the exception of a few cases where it was moderate. Internal consistency, as measured by Cronbach's alpha, was unsatisfactory as far as L scores are concerned, and acceptable as far as ML scores are concerned. Convergent validity, measured by correlations between all the THAM-3 measures and the SPM38 scores (Table 4), showed statistically significant correlations, although moderate. The inter-correlations between all the THAM-3 scores (Table 5) varied from moderate to strong.

CFA (Brown, 2006) results supported the theoretical two-factor structure of the original test, assigning the ML subtests' scores to the corresponding Metalinguistic latent factor, the L subtests' scores to the corresponding Linguistic latent factor, and explaining the inter-correlations between the observed variables. The final model provided a good fit to the empirical data. However, we must point to some differences between the two main factors: while the Metalinguistic ability factor (which expresses metalinguistic awareness at the *explicit level)* proved to be a *unitary construct*, well measured by the ML scale, the Linguistic ability factor (which expresses metalinguistic ability at the *intuitive level*) appeared as a *weaker construct*. This finding needs further in-depth studies.

Overall, on the basis of the results reported above, the THAM-3 can be considered a valid and reliable measure of metalinguistic abilities for adolescents and adults having French as their first language or learning French as a second or foreign language.

Many avenues of research can be considered for further studies on this test. A major line of investigation could be the administration of the THAM-3 to native speakers of French in French-speaking countries in Europe (France, Switzerland and Belgium) and in other continents (North America, Africa). Then, as there is another validated linguistic version of the same test, namely the Spanish THAM-3 (*Test de Habilidades Metalingüísticas n.3*) (Lasagabaster, Merino & Pinto, 2015), comparisons between the results of the validations of the French and the Spanish tests will surely bring further light on the construct underlying these instruments. The perspective of a possible validation of the English (MAT-3) (Pinto, Titone & Trusso, 1999) and the German (MKT-3) (Jessner, Pellegrini, Moroder, Hofer & Pinto 2015) versions could add further elements to our understanding of metalinguistic processes at the age level targeted by this test. Future studies could also investigate the most difficult and the least difficult items, or the effect of a possible modification of the scoring criteria in the subtest that proved to be the most demanding, i.e. Acceptability, especially in the justification of errors.

REFERENCES

BROWN, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.

COHEN, J. (1960). A coefficient of agreement for nominal scales. *Education and Psychological Measurement, 20*, 37-46.

JESSNER, U., PELLEGRINI, C., MORODER, V., HOFER, B. & PINTO, M. A. (2015). MKT. *Metalinguistischer Kompetenztest Teil 3*. Innsbruck. Studia Universitätverlag Innsbruck.

LASAGABASTER, D., MERINO, J.A., & PINTO, M.A. (2015). *Test de habilidades metalingüisticas para adolescentes y adultos: THAM-3.* Bilbao: Universidad del País Vasco/Euskal Herriko Unibertsitatea. Available at www.matelproject.com.

PINTO, M. A. & EL EUCH, S. (2015). *La conscience métalinguistique. Théorie, développement et instruments de mesure.* Québec: Presses de l'Université Laval.

PINTO, M. A., & ILICETO, P. (2007). *TAM-3. Test di abilità metalinguistiche n.3. Fascia adolescente adulta*. Roma : Carocci faber.

PINTO, M. A., TITONE, R., & TRUSSO, F. (1999). *Metalinguistic awareness: Theory, development and measurement instruments*. Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali.

RAVEN, J., RAVEN, J.C., & COURT, J.H. (1998a). *Manual for Raven's Progressive Matrices and Vocabulary Scales . Section 1: General Overview*. San Antonio, TX: Harcourt Assessment.

Tables and Figures

Table 1. Means and SD of all the scores of the THAM-3 (Pinto & El Euch, 2015) vs the TAM-3 (Pinto & Iliceto, 2007)

Subtests and totals L and ML (max. score)	Means (French sample)	SD (French sample)	Means (Italian sample)	SD (Italian sample)
Comprehension L (4)	2.99	1.09	3.07	.96
Comprehension ML (16)	7.80	2.77	5.48	3.14
Acceptability L (15)	6.99	2.45	9.93	2.85
Acceptability ML (30)	4.63	3.43	5.28	4.98
Figurative Language L (2)	1.62	.59	1.37	.80
Figurative Language ML (12)	5.45	1.98	5.57	3.16
THAM-3 total L (21)	11.60	2.97	14.37	3.40
THAM-3 total ML (58)	17.87	6.52	16.32	9.14
THAM-3 total score (79)	29.47	8.67	30.69	11.46

Table 2. Skewness and Kurtosis for the sample of the THAM-3 (Pinto & El Euch, 2015) vs the sample of the TAM-3 (Pinto & Iliceto, 2007)

Subtests and totals L and ML	Skewness (French sample)	Kurtosis (French sample)	Skewness (Italian sample)	Kurtosis (Italian sample)
Comprehension L	-1.019	.278	.205	-.886
Comprehension ML	.260	.212	-.177	.516
Acceptability L	-.129	-.630	.480	-.648
Acceptability ML	.798	.363	.054	.824
Figurative Language L	-1.333	.748	-1.033	-.775
Figurative Language ML	.180	-.648	-.812	.214
THAM-3 total L	-.212	-.320	.134	-.555
THAM-3 total ML	.513	.195	-.012	.664
THAM-3 total score	.307	-.072	-.126	.479

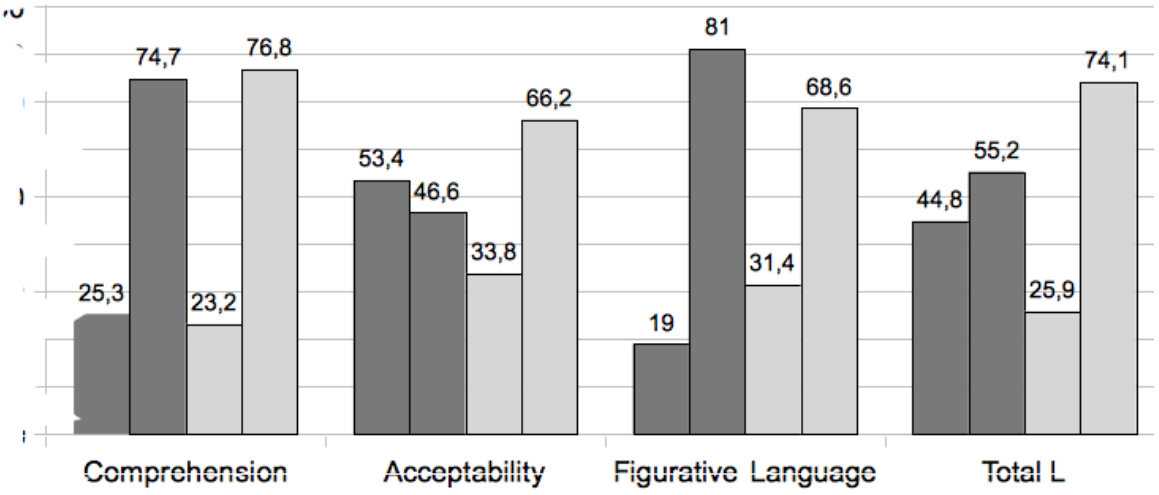

Figure 1. Percentage of L score level (0, 1) in each subtest of the THAM-3 (Pinto & El Euch, 2015) vs the TAM-3 (Pinto & Iliceto, 2007)

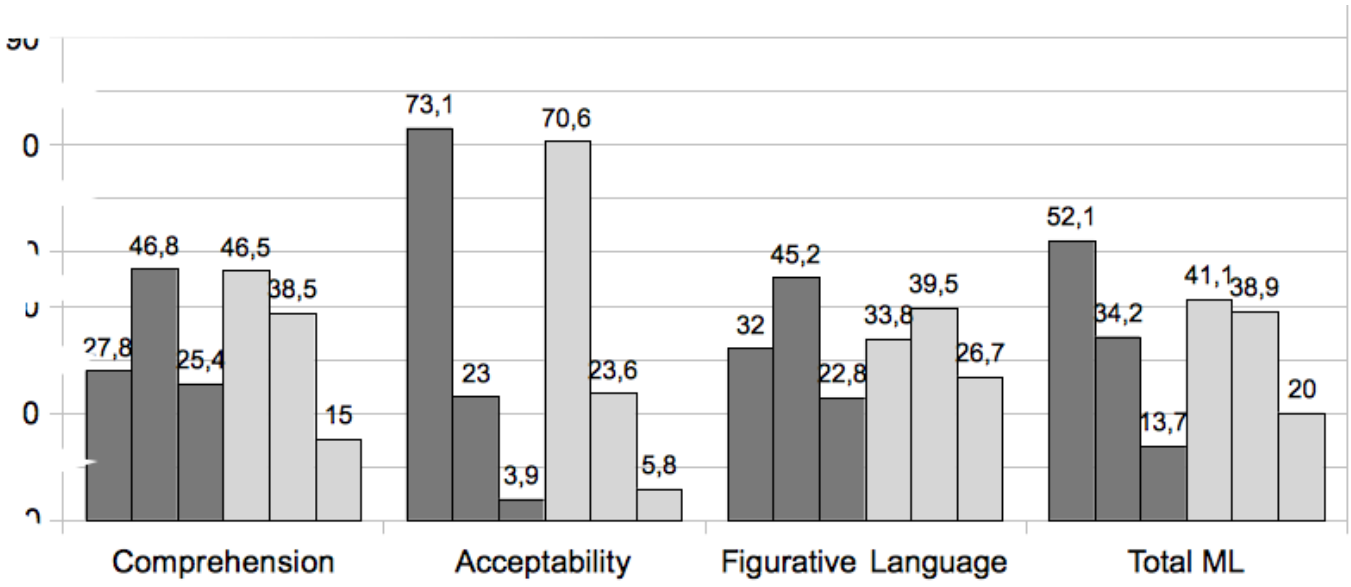

Figure 2. Percentage of ML score level (0, 1, 2) in each subtest of the THAM-3 (Pinto & El Euch, 2015) vs the TAM-3 (Pinto & Iliceto, 2007).

Table 3. ANOVA (Education level x Gender) in relation to the THAM-3 (Pinto & El Euch, 2015). Figurative Language L

Source		Means	SD	F	p	Partial Eta Squared	
Education	<=18	1.63	.59	.005	.941	.000	
	>18	1.50	.65				
Gender	Males	1.38	.74	6.810	.010	.045	
	Females	1.71	.51				
Education * Gender	<=18	Males	1.41	.75	.729	.394	.005
		Females	1.70	.51			
	>18	Males	1.25	.70			
		Females	1.83	.40			

Table 4. Correlations matrix between the THAM-3 (Pinto & El Euch, 2015) and the SPM38 (Raven, Raven & Court, 1998)

Subtests and totals L and ML	**SPM38**
Comprehension L	.159
Comprehension ML	.336***
Acceptability L	.291**
Acceptability ML	.107
Figurative Language L	.106
Figurative Language ML	.243**
THAM-3 total L	.328***
THAM-3 total ML	.271**
THAM-3 total score	.319***

** p <.01; *** p <.001

Table 5. Inter-correlations matrix of the THAM-3 (Pinto & El Euch, 2015)								
	1	2	3	4	5	6	7	8
1) Comprehension L	1							
2) Comprehension ML	.186*	1						
3) Acceptability L	.084	.431***	1					
4) Acceptability ML	.036	.472***	.632***	1				
5) Figurative Language L	.186*	.156	.176*	.189*	1			
6) Figurative Language ML	.067	.439***	.408***	.390***	.172*	1		
7) THAM-3 total L	.477***	.457***	.894***	.574***	.416***	.398***	1	
8) THAM-3 total ML	.119	.807***	.640***	.846***	.218**	.696***	.617***	1
9) THAM-3 total score	.253**	.763***	.787***	.832***	.306***	.659***	.806***	.963***
* p <.05; ** p <.01; *** p <.001								

THE TRANSLATION OF THE ITALIAN METALINGUISTIC ABILITY TESTS TAM-2 AND TAM-3 (Pinto, 1999) INTO THE GERMAN MKT-2 (*Jessner, Hofer & Pinto, 2015)* AND MKT-3 (*Jessner, Pellegrini, Moroder, Hofer & Pinto, 2015*)

ULRIKE JESSNER,* BARBARA HOFER,** CLAUDIA PELLEGRINI,*** MARIA ANTONIETTA PINTO****

Universität Innsbruck
Università di Roma "SAPIENZA"

ABSTRACT

In this article we describe the major challenges, strategies and solutions found in the process of translating two metalinguistic tests from the Italian original version (the TAM-2, *Test di Abilità Metalinguistiche n.2*; 9-14 years, Pinto, Candilera & Iliceto, 2003, and the TAM-3, *Test di Abilità Metalinguistiche n-3. Late adolescence-adulthood,* Pinto & Iliceto, 2007) into German (as, respectively, MKT-2, *Metalinguistischer Kompetenztest Teil 2,* Jessner, Hofer & Pinto, 2015*)* and MKT-3, *Metalinguistischer Kompetenztest Teil 3,* Jessner, Pellegrini, Moroder, Hofer, & Pinto, 2015).

Although these tests address two different developmental stages (childhood-early adolescence; late adolescence-adulthood), they share the same conception of metalinguistic awareness, as based on a definite distinction between an implicit, intuitive and global level, and an explicit, intentional and analytic level. This distinction is embodied in the coding system of both tests and most contributes to their complexity, in cognitive and linguistic terms, whatever the language. When this intrinsic complexity must be rendered in languages other than the original one, the problem is clearly amplified. Before the translation into German, three previous translations existed of the whole Italian work on metalinguistic awareness (Pinto, 1999), that included the presentation of the metalinguistic tests. There existed an English (Pinto, Titone, Trusso, 1999), a Spanish (Pinto, Titone, Gonzáles Gil, 2000), and a French version (Pinto, El Euch, 2015), that had each posed problems related to the specificities of the respective language. German posed further linguistic specificities that this article has reviewed.

One of the most salient is the fact that the German language resorts to Latin-based lexemes much less than Spanish, French and even English. This had heavy consequences on the choice of many lexemes in the items in both the MKT-2 and the MKT-3, and in some cases it entailed the reformulation of entire items. The new formulations had to capture the essence of the metalinguistic problem present in the original item

* Corresponding author: Ulrike Jessner, Innsbruck Universität. E-mail: Ulrike.Jessner@uibk.ac.at .
** Barbara Hofer. Innsbruck Universität. E-mail: hofbar@libero.it.
*** Claudia Pellegrini. Innsbruck Universität. E-mail: claudia.pellegrini65@tiscali.it.
**** Maria Antonietta Pinto. Università di Roma "SAPIENZA". E-mail: mariantonietta.pinto@uniroma1.it.

and find totally different forms for expressing it. This process also applied to the search for suitable poetic verses in the MKT-3. The issue was to find metaphors and other tropes in German poetry based on semantic characteristics comparable to those of the Italian poetic verses of the TAM-3.

How felicitous these renderings have been, we may partially infer from the impact that the MKT-2 and the MKT-3 had on native speakers of German in the pilot studies (Candilera, Iliceto, Hofer, Pellegrini & Pinto, 2015). These studies showed that the *type* of responses provided by the participants were the same as those elicited by the Italian TAM-2 and TAM-3. Therefore, the German-speaking participants' answers fell under the categories established by the test, which means that the *processes* elicited by the items and the questions were of the same nature as those elicited by the original version. In addition, in *quantitative* terms, the L performances of the German respondents were entirely comparable to those of the Italian counterparts, and the ML performances were even superior. This means that they were perfectly able to interpret the most demanding metalinguistic requests, which would have been impossible without a suitable rendering of the items.

Nevertheless, further studies with native German speakers are required to confirm this first, encouraging impact of the MKT-2 and the MKT-3.

REFERENCES

CANDILERA, G., ILICETO, P., HOFER , B., PELLEGRINI, C. & PINTO, M.A. (2015). The pilot studies on the MKT-2 (Metalinguistischer Kompetenztest Teil 2, *Jessner, Hofer & Pinto,* 2015) and the MKT-3 (Metalinguistischer Kompetenztest Teil 3, *Jessner, Pellegrini, Moroder, Hofer & Pinto*, 2015). *RIVISTA DI PSICOLINGUISTICA APPLICATA/JOURNAL OF APPLIED PSYCHOLINGUISTICS*, Special Issue, Ed. by M.A. Pinto: *The "MATEL" (Metalinguistic Awareness Tests in European Languages) Project: Validations and translations"*. XV, 2.

JESSNER, U., HOFER, B. & PINTO, M.A. (2015) *MKT. Metalinguistischer Kompetenztest Teil 2.* Innsbruck, Studia Universitätsverlag Innsbruck.

JESSNER, U., PELLEGRINI, C., MORODER, V., HOFER, B. & PINTO, M.A. (2015). *MKT-3. Metalinguistischer Kompetenztest Teil 3,* Innsbruck, Studia UniversitätsVerlag Innsbruck.

PINTO, M. A. (1999). *La consapevolezza metalinguistica. Teoria, sviluppo, strumenti di misurazione.* Pisa-Roma, Istituti Editoriali e Poligrafici Internazionali.

PINTO, M.A., CANDILERA, G. & ILICETO, P. (2003) *TAM-2. Test di abilita`metalinguistiche n.2 (9 14 anni). La valutazione dello sviluppo metalinguistico tra scuola elementare e scuola media*. Roma: Scione Editore.

PINTO, M. A., & EL EUCH. S., (2015) *La Conscience Metalinguistique. Théorie, développement et instruments de mesure*. Quebec: P.U.L.

PINTO, M. A., ILICETO P. (2007). *TAM-3 Test di abilità metalinguistiche n° 3*, *Fascia adolescente*

adulta. Roma: Carrocci Faber.

PINTO, M. A., TITONE, R. & GONZÁLES GÍL, M.D. (2000). *La consciencia metalingüística. Teoría, desarrollo e instrumentos de medición*. Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali.

PINTO, M. A., TITONE, R. & TRUSSO, F. (1999). *Metalinguistic Awareness. Theory, Development and Measurement Instruments.* Pisa-Roma: Istituti Editoriali e Poligrafici Internazionali.

THE PILOT STUDIES ON THE MKT-2 (METALINGUISTISCHER KOMPETENZTEST TEIL 2, *Jessner, Hofer & Pinto,* 2015) AND THE MKT-3 (METALINGUISTISCHER KOMPETENZTEST TEIL 3, *Jessner, Pellegrini, Moroder, Hofer & Pinto*, 2015)

GABRIELLA CANDILERA*, PAOLO ILICETO**, BARBARA HOFER ***,
CLAUDIA PELLEGRINI,**** MARIA ANTONIETTA PINTO*****

S &P. Statistics & Psychometrics Ltd
Universität Innsbruck
Università di Roma "SAPIENZA"

ABSTRACT:

This article presents the results of the pilot studies conducted on the MKT-2 (Jessner, Hofer & Pinto, 2015) and the MKT-3 (Jessner, Pellegrini, Moroder, Hofer & Pinto, 2015), two tests translated from the original Italian metalinguistic ability test TAM-2 (Pinto, Candilera & Iliceto, 2003) and TAM-3 (Pinto & Iliceto, 2007) into German, with partial adaptations. Given the difficulties in the translation processes (Jessner, Pellegrini, Hofer & Pinto, 2015) and the absolute novelty of the administration to native speakers of German, these pilot studies were quite challenging. For both the MKT-2 and the MKT-3, performances were measured with a range of descriptive and inferential statistics indicators and by comparison with subjects of the same age in the normative samples of the Italian TAM-2 (Pinto *et al.* 2003) and TAM-3 (Pinto & Iliceto, 2007). The expectation was that the performances were similar to those of the corresponding Italian normative samples.

The study on the MKT-2 was conducted on one class of 17 Junior school pupils (mean age: 12,7; SD: 6.7), all native speakers of German, living in a small village in South Tyrol (North of Italy), where German is the dominant language. Children showed good metalinguistic performances at the implicit level, as measured by L scores (Table 1 and Fig.1), which were very similar to those exhibited by Italian children of the same age of the normative sample. Their metalinguistic performances at the explicit level were less satisfactory but still acceptable (Fig.2), as the dominant percentage of ML answers was either at the intermediate or at the maximum level. In all ML measures (subtests and total ML), both the means

* Corresponding author: Gabriella Candilera. S&P Statistics & Psychometrics Ltd. Via Tuscolana, 458. 00181. Roma. E-mail: gabriellacandilera@libero.it.
**Paolo Iliceto, S&P Statistics & Psychometrics Ltd. Via Tuscolana, 458. 00181. Roma. E-mail: paolo.iliceto@statpsyc.com.
*** Barbara Hofer. Innsbruck Universität. E-mail: hofbar@libero.it
**** Claudia Pellegrini. Innsbruck Universität. E-mail: claudia.pellegrini65@tiscali.it
**** Maria Antonietta Pinto. Università di Roma "SAPIENZA". E-mail: mariantonietta.pinto@uniroma1.it.

(Table 1) and the distribution of the score levels showed superior performances in the German-speaking sample compared to the Italian normative sample. Skewness and Kurtosis values (Table 2), as measures of the normality of the curve, were within range. Reliability coefficients varied from medium to high and the intercorrelations matrix (Table 3) showed positive associations, varying from moderate to high. Principal Component Analysis (PCA) (Table 4) highlighted a major component that corresponds to ML scores, and embodies the core of the theoretical construct of the test.

A similar pattern appeared in the study of the MKT-3, conducted on 44 students (mean age: 17, 3; SD: 6.1) also native speakers of German, enrolled in two different High schools in South Tyrol where German is the language of instruction. Their performances at the implicit level of metalinguistic awareness (L scores) (Table 5) were quite satisfactory, with high percentages of good answers (Fig.3), and very close to those of the Italian normative sample. As with the MKT-2, the ML scores, that measure metalinguistic awareness at the explicit level, were less successful (Fig. 4) although the best score level (ML2) was dominant in two out of the three subtests and in the total, namely Comprehension and Figurative language, two metasemantic tasks. The metagrammatical subtest (Acceptability) proved more difficult, with a much lower percentage of good answers. Again, the ML scores of the German-speaking participants were clearly superior to those of the Italian normative group. Skewness and Kurtosis values (Table 6) were within range. Reliability coefficients varied from medium to high and the intercorrelations matrix (Table 7) showed positive associations, varying from moderate to strong. By means of PCA (Table 8) a major component was extracted, that efficiently represents the metalinguistic core of the MKT-3, as measured by the various ML subtests.

Overall, the MKT-2 and the MKT-3 had a good impact on the German-speaking subjects they addressed. In spite of difficult administration conditions, due to particular time constraints in each context, participants showed good metalinguistic abilities both at the intuitive level, as measured by L answers, and at a more demanding level, as measured by ML answers. Unexpectedly, while the L means and percentages were very close to those of the Italian normative sample, ML performances were clearly superior. This result was found at each age level, therefore *at different levels of metalinguistic complexity,* as the MKT-3 has more sophisticated items than the MKT-2 and the argumentative abilities required are much more elaborated. Although the comparison between the German-speaking sample and the Italian normative sample has limited value, due to the considerable gap in the number of participants in each case, the metalinguistic superiority of the pilot samples deserves attention. As these German-speaking groups both lived in small towns, the most plausible interpretation of their particularly good results in metalinguistic abilities could be attributed to the school system. This might have conveyed high sensitivity to the mother tongue, and to the other surrounding languages, Italian as L2 in the first place, but also foreign languages, as is surely the case for the one of the schools where the MKT-3 was administered, namely the “Sprachengymnasium”, a High school

with a curriculum focused on foreign languages.

Although preliminary, the good results concerning the normality of the curve, the internal consistence of the items, the correlations between subtests, and the metalinguistic component extracted with PCA, all encourage towards further validation studies of both the MKT-2 and MKT-3. The results of these future validations could provide other German-speaking countries in Europe (Germany, Austria, Switzerland) with metalinguistic tests of relevant educational impact.

REFERENCES

JESSNER, U., HOFER, B., PELLEGRINI, C. & PINTO, M.A. (2015). The translation of the Italian metalinguistic ability tests into German: the MKT-2 and the MKT-3. Special Issue, Ed. by M.A. Pinto, *RIVISTA DI PSICOLINGUISTICA APPLICATA/JOURNAL OF APPLIED PSYCHOLINGUISTICS, The "MATEL" Project (Metalinguistic Awareness Tests in European Languages): Validations and translations.* XV, 2.

JESSNER, U., HOFER, B. & PINTO, M.A. (2015). *MKT-2. Metalinguistischer Kompetenztest Teil 2.* Innsbruck, Studia Universitätsverlag Innsbruck.

JESSNER, U., PELLEGRINI, C., MORODER, V., HOFER, B. & PINTO, M.A. (2015). *MKT-3. Metalinguistischer Kompetenztest Teil 3.* Innsbruck, Studia Universitätsverlag Innsbruck.

PINTO, M.A. (1999). *La consapevolezza metalinguistica. Teoria, sviluppo e strumenti di misurazione,* Pisa-Roma, Istituti Editoriali e Poligrafici Internazionali.

PINTO, M.A., CANDILERA, G. ILICETO, P. (2003). *TAM-2. Test di abilità metalinguistiche n.2 (9 - 14anni). Manuale di istruzioni.* Roma, Scione.

PINTO, M.A., EL EUCH, S. (2015). *La conscience métalinguistique. Théorie, développement et Instruments de mesure.* Québec. P.U.L.

PINTO, M.A., ILICETO, P. (2007). *TAM-3. Test di abilità metalinguistiche n.3. Fascia adolescente adulta*. Roma, Carocci faber.

PINTO, M.A., TITONE, R. (1989). Uno strumento di misurazione delle abilita metalinguistiche: il T.A.M. *Rassegna Italiana di Linguistica Applicata*, Vol. XXI. N.1-2. pp 59-128.

Tables and Figures

Table 1. MKT-2. (Jessner, Hofer & Pinto, 2015). Means and standard deviations. German-speaking sample vs Italian normative sample.

Subtests	Mean (Ger. sample)	SD (Ger. sample)	Mean (Italian sample)	SD Italian sample)
Synonymy L **(5)**	4.06	.65	4.47	.66
Synonymy ML **(10)**	4.76	2.81	2.47	1.99
Acceptability L (**28)**	23.71	2.66	25.71	2.69
Acceptability ML **(26)**	14.41	5.10	9.42	4.52
Ambiguity L **(7)**	6.71	.68	5.68	1.10
Ambiguity ML **(14)**	8.76	2.88	5.32	2.25
Gr. Function L **(6)**	5.88	.33	5.76	.49
Gr. Function ML **(24)**	11.59	5.11	8.34	3.47
MKT-2 total L **(46)**	40.35	2.93	41.64	3.77
MKT-2 total ML (**74)**	39.53	13.19	25.55	9.73
MKT-2 total test **(120)**	79.88	13.62	67.19	12.44

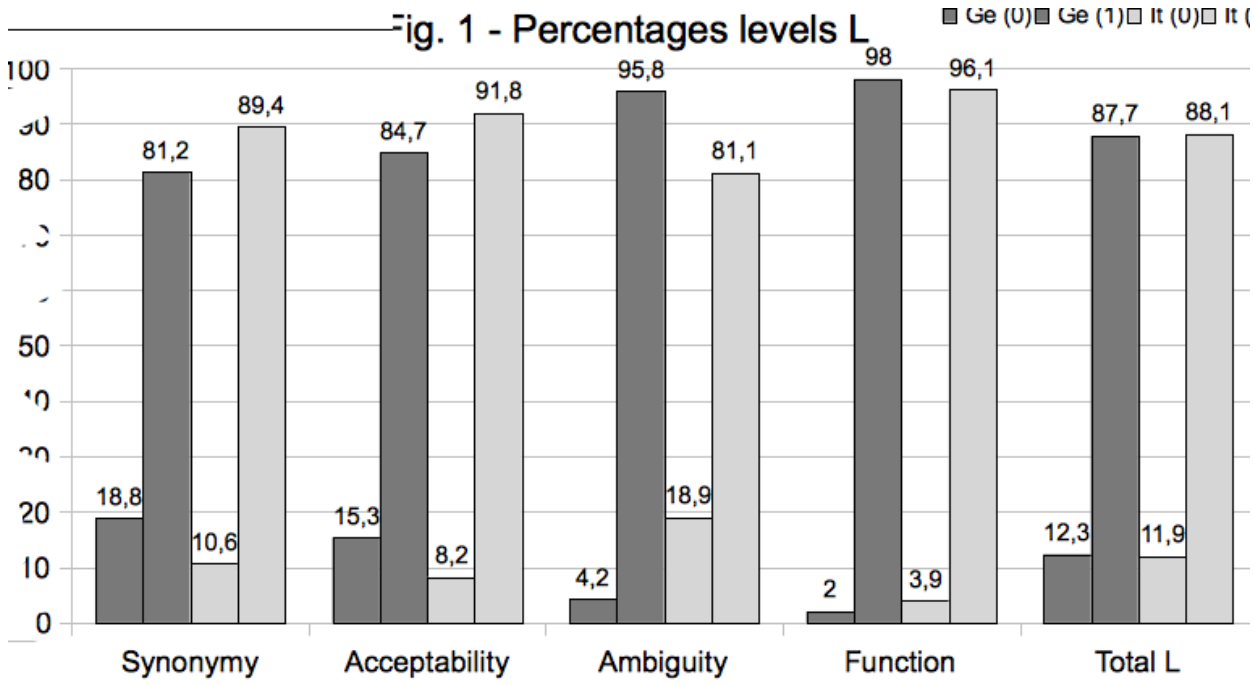

Fig. 1. MKT-2. (Jessner, Hofer & Pinto, 2015). Distribution of L score levels in the subtests and in the total L of the MKT-2. German-speaking sample vs Italian normative sample.

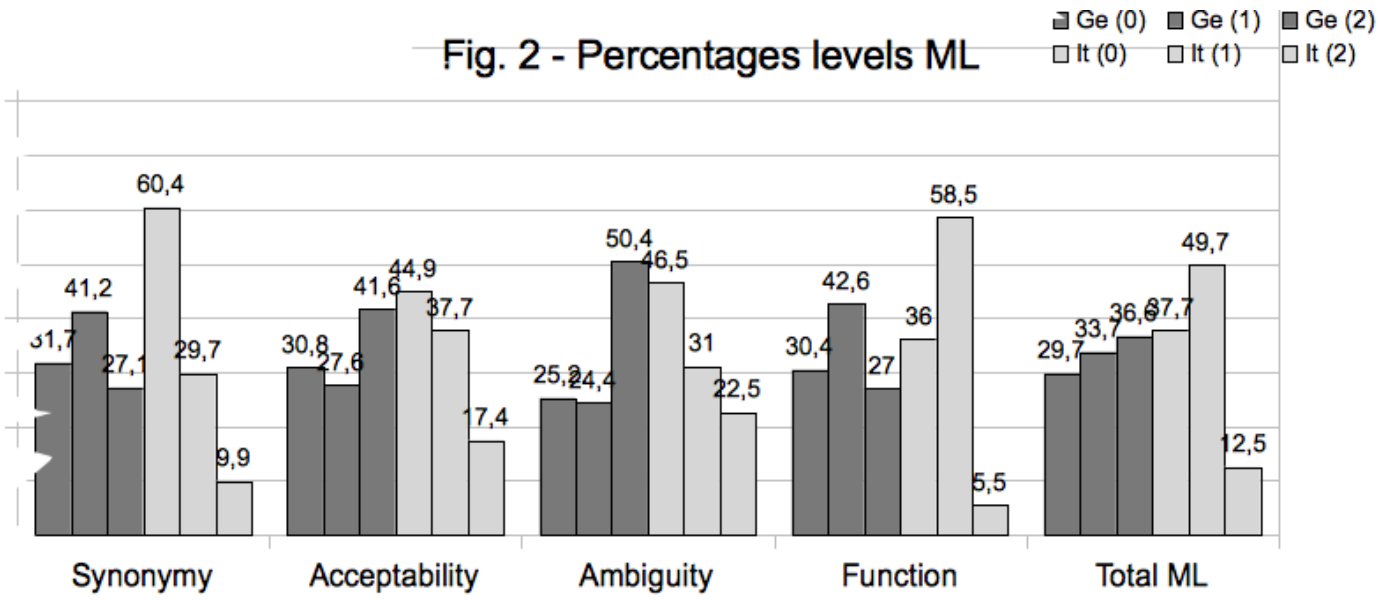

Fig.2.Distribution of ML score levels in the subtests and in the total ML - MKT-2 (Jessner, Hofer, Pinto). German-speaking sample vs Italian normative sample.

Table 2. MKT-2. (Jessner, Hofer & Pinto, 2015). Skewness and Kurtosis. German-speaking sample vs Italian normative sample.

Subtests	Skewness (Ger. sample)	Kurtosis (Ger. sample)	Skewness (Italian sample)	Kurtosis (Italian sample)
Synonymy L	-.057	-.314	-1.647	6.043
Synonymy ML	-.222	-.475	.576	.031
Acceptability L	.336	-.869	-1.505	2.362
Acceptability ML	-.534	-.434	.049	-.712
Ambiguity L	-2.177	3.453	-.931	1.145
Ambiguity ML	-.744	.431	.015	.005
Gr. function L	-2.610	5.440	-2.117	4.464
Gr.function ML	.233	-.812	-.383	-.397
MKT-2 total L	-.142	-1.214	-1.509	3.107
MKT-2 total ML	-.651	-.104	-.091	-.402
MKT-2 total score	-.594	-.136	-.634	.576

Table 3. MKT-2 (Jessner, Hofer & Pinto, 2015). Inter-correlations matrix

	1	2	3	4	5	6	7	8	9	10	11
1) Synonymy L	1										
2) Synonymy ML	.547*	1									
3) Acceptability L	-.203	-.260	1								
4) Acceptability ML	.178	.542*	.143	1							
5) Ambiguity L	.041	.415	.258	.198	1						
6) Ambiguity ML	.205	.646**	.080	.644**	.658**	1					
7) Gr. function L	-.252	-.031	.100	-.154	.387	.035	1				
8) Gr. function ML	.100	.644**	-.312	.479	.195	.578*	-.251	1			
9) MKT-2 total L	.021	-.020	.933**	.198	.520*	.276	.238	-.244	1		
10) MKT-2 total ML	.270	.814**	-.104	.829**	.384	.830**	-.156	.837**	.038	1	
11) MKT-2 total test	.265	.784**	.101	.845**	.484*	.863**	-.100	.758**	.253	.977**	1

* p <.05; ** p <.01; *** p <.001

Table 4. MKT-2 (Jessner, Hofer &s Pinto, 2015). PCA (Principal Component Analysis): factor loadings and commonalities

Subtests	ML Component	Commonalities
Synonymy	.855	.730
Acceptability	.796	.634
Ambiguity	.866	.751
Grammatical function	.809	.655

Table 5. MKT-3 (Jessner, Pellegrini, Moroder, Hofer & Pinto, 2015). Means and standard deviations. German-speaking sample vs Italian normative sample.

Subtests	Means (Ger. sample)	SD (Ger. sample)	Means (Italian sample)	SD (Italian sample)
Comprehension L (4)	3.14	.73	3.07	.96
Comprehension ML (16)	10.00	3.89	5.48	3.14
Acceptability L (15)	9.00	2.61	9.93	2.85
Acceptability ML (30)	11.45	7.07	5.28	4.98
Fig Language L (2)	1.89	.44	1.37	.80
Fig Language ML (12)	8.93	2.21	5.57	3.16
MKT-3 total L (21)	14.02	3.15	14.37	3.40
MKT-3 total ML (58)	30.39	11.66	16.32	9.14
MKT-3 total score (79)	44.41	14.40	30.69	11.46

Table 6. MKT-3 (Jessner, Pellegrini, Moroder, Hofer & Pinto, 2015). Skewness and Kurtosis. German-speaking sample vs Italian normative sample.

Subtests	Skewness (Ger - sample)	Kurtosis (Ger - sample)	Skewness (Italian sample)	Kurtosis (Italian sample)
Comprehension L	-.221	-1.072	.205	-.886
Comprehension ML	-.470	-.463	-.177	.516
Acceptability L	-.303	-.506	.480	-.648
Acceptability ML	.384	-.940	.054	.824
Figurative Language L	-3.945	14.824	-1.033	-.775
Figurative Language ML	-.354	-1.025	-.812	.214
Ger- THAM-3 total L	-.532	-.040	.134	-.555
Ger-THAM-3 total ML	-.021	-1.014	-.012	.664
Ger-THAM-3 total score	-.130	-.826	-.126	.479

Table 7. MKT-3 (Jessner, Pellegrini, Moroder, Hofer & Pinto, 2015). Inter-correlations matrix

	1	2	3	4	5	6	7	8	9
1) Comprehension L	1								
2) Comprehension ML	.691**	1							
3) Acceptability L	.400**	.580**	1						
4) Acceptability ML	.458**	.538**	.765**	1					
5) Fig Language L	.192	.337*	.321*	.240	1				
6) Fig Language ML	.535**	.674**	.647**	.791**	.419**	1			
7) MKT-3 total L	.590**	.688**	.966**	.773**	.451**	.719**	1		
8) MKT-3 total ML	.610**	.788**	.780**	.936**	.337*	.894**	.835**	1	
9) MKT-3 total score	.623**	.789**	.843**	.927**	.372*	.882**	.895**	.993**	1

* p <.05; ** p <.01; *** p <.001

Table 8. MKT-3 (Jessner, Pellegrini, Moroder, Hofer & Pinto, 2015). PCA (Principal Component Analysis): factor loadings and commonalities

Subtests	ML Component	Commonalities
Comprehension	.826	.682
Acceptability	.884	.781
Figurative Language	.937	.877

Collana Materiali e documenti

1. La plastica nell'arte e per l'arte. I polimeri come materiali di base e di restauro per i beni culturali
   a cura di Luigi Campanella, Alice Hansen, Ezio Martuscelli, Antonella Russo
2. Museo di Merceologia, Sapienza Università di Roma. Catalogo ragionato degli strumenti scientifici / Museum of Commodity Science, Sapienza University of Rome. Catalogue Raisonné of scientific instruments
   Małgorzata Biniecka, Patrizia Falconi, Raffaella Preti
3. Video didattico sull'uso interattivo del TAM-2
   Federica Micale, Irene Bracone, Maria Antonietta Pinto
4. Video didattico sull'uso interattivo del TAM-3
   Federica Micale e Maria Antonietta Pinto
5. Utilización interactiva del THAM-2
   Pilar Núñez Delgado y María Santamarina Sancho
6. Utilización interactiva del THAM-3
   Vídeo didáctico sobre un grupo de discusión
   Jon Ander Merino y David Lasagabaster
7. Utilisation interactive du THAM-3
   Vidéo didactique à partir d'items du THAM-3
   Isabelle Monette & Sonia El Euch
8. Tham-2 test de habilidades metalingüísticas nº 2 (9-14 años)
   Pilar Núñez Delgado y Maria Antonietta Pinto
9. The "MATEL" Project: Research Results
   edited by Maria Antonietta Pinto
10. Metalinguistic Excercises as Classroom Activities
    edited by Maria Antonietta Pinto

Materiali e documenti

ISBN 978-88-98533-72-5

www.editricesapienza.it