ASSESSING METAPHOR COMPREHENSION AS A METASEMANTIC ABILITY IN STUDENTS FROM 9-TO-14 YEARS-OLD

Maria Antonietta Pinto

mariantonietta.pinto@uniroma1.it
University of Rome «La Sapienza» (Italy)

Sergio Melogno

sergiomelogno@tiscali.it
University of Rome «La Sapienza» (Italy)

Paolo Iliceto

paolo.iliceto@libero.it
www.laboratoriopsicometrico.it

Abstract: This article presents a test for assessing metaphor comprehension (MCT) in explicit linguistic form for subjects aged 9- to 14 years-old, i.e. in a transitional age where significant qualitative changes appear for this type of ability. Metaphor is viewed as a form of semantic conflict induced by the anomalous combination of the conventional meanings of its main constituents - tenor and vehicle - and metaphor comprehension is framed as a metasemantic ability based on the analysis of these meanings (Gombert 1990), that can have implications both for teaching and for learning strategies.The authors propose a functionalist piagetian frame, based on Piaget's latest equilibration model (1975), for analysing how this semantic conflict can be faced and solved by children in the developmental span considered. The test is paper-and-pencil, composed of 12 items subdivided into 2 groups of metaphors: Psycho-physical (PP) and Conceptual (C), mainly drawn and adapted to Italian language from international literature on metaphor comprehension. The sample is composed of 874 Italian children from 4th to 8th grade, with gender balance, of average social background. By means of Principal Components Analysis, with *oblimin* rotation, a two-factor solution emerged, that espouses the C/PP metaphors distinction. All corrected item-total correlation coefficients >.30 were representative and acceptable. Correlation coefficients between MCT and Standard progressive Matrices (SPM38) and some validated metalinguistic subtests were all significant at p <0.01 level, showing good convergent validity. *Cronbach's Alpha* coefficients for the two subscales were: .62 (C), .51 (PP) and .70 for MCT total score. Test-retest correlation coefficients for the two subscales were .79 (C), .69 (PP), and .83 for MCT total score. *Cohen's Kappa* coefficients

for interrater reliability are: .75 (4th-5th), .74 (6th), .67 (7th), .81 (8th). A two-factor ANOVA (gender and age) showed that the test is gender-sensitive in favour of females but that age is an even stronger factor, which underlines the developmental character of the test. Limitations and possible research developments are analysed in the Discussion.

Keywords: Metaphor comprehension, metasemantic abilities, late childhood, mid-adolescence, assessment, Italian language

1 – Introduction

The problematics of metaphor comprehension have been extensively studied by developmental psychologists during the past five decades, since the pioneering study by Asch and Nerlowe (1960), and explored in their cognitive, linguistic, metalinguistic and methodological aspects (for a review, see Winner 1997). Starting from the Eighties of last century, a convergent interest for metaphor understanding coming from separate disciplinary fields, such as Cognitive Linguistics (Lakoff & Johnson 1980), Clinical Psychology related to the Theory of Mind line of research (Happé 1993; 1995), Metalinguistics (Tunmer, Pratt & Herriman 1984; Gombert 1990), Language Teaching (Ortony, ed. 1979; Danesi 1986; Zhang & Hu 2009), Linguistic Anthropology (Duranti, ed. 2001) magnified its relevance in psychological, clinical, anthropological and educational areas.

In this paper we will present a test for assessing metaphor comprehension (MCT, *henceforth*) in an age range, 9- to 14 years-old, that covers late childhood up to mid-adolescence, in developmental terms, and the transition from elementary to Junior school, in educational terms. We will first outline the theoretical reasons that brought us to create a test for this precise developmental span and its implications for education. Then we will describe the methodology used for implementing this project: the standardisation sample, the rationale for the creation of the items, the administration and rating procedures and the essential psychometric characteristics of the test. Finally, we will discuss limitations and possible further developments.

1.1- Metaphor comprehension: a complex metasemantic ability for solving a peculiar semantic conflict

Metaphors, as other figures of speech, can be grasped on very intuitive grounds at any age and for various reasons. But, as scholars, we may be interested in studying more advanced levels of awareness by asking *to analyse the reasons why a given metaphor produces its characteristic semantic effect*. This leads to analyse the peculiar semantic combination underlying all metaphors, namely, the combination between the conventional meanings of the first part of the metaphor – canonically called 'tenor' and those of its second part – 'the vehicle' (Richards 1936), which normally pertain to independent semantic areas. The essence of

the surprise generated by new metaphors lies precisely in the strangeness of this semantic combination, which forces the interpreter to search for a possible 'common ground' (Richards 1936) for reconciling the respective meanings of tenor and vehicle. If, moreover, metaphors are presented in purely linguistic form and can be considered exclusively in linguistic form, the cognitive and linguistic processes elicited have a definite metalinguistic character. Specifically, the type of metalinguistic ability required is *metasemantic*, as focus is mainly on meanings and their reciprocal relations (Gombert 1990). This type of task is cognitively demanding both for the nature of the linguistic meanings that constitute its object and for the metalinguistic level at which subjects are solicited. This is the challenge underlying the creation of MCT, whose age range enables to measure significant qualitative changes in the way children process metaphors.

The concern for a relatively late achievement in metaphor comprehension can be questioned, since we know of studies reporting much earlier achievements (Billow 1975; Gentner 1988 & Siltanen 1986). For instance, since the mid Seventies of last century some scholars have explored the facilitating role of certain types of metaphors in relation to the cognitive characteristics of early ages. Apparently, when children were presented metaphors based on perceptual features, close to their experience and also to those metaphors they spontaneously produce, they were able to grasp metaphorical meanings as early as 4 years-old. (Winner 1988; Gentner 1988; Siltanen 1986; Melogno & Pinto 1996). During the Seventies and the Eighties of last century, other studies also produced evidence of early achievement when children were given the opportunity to respond in nonverbal form, for instance by choosing among figurative alternatives or precoded verbal responses or enacting metaphorical meanings with toys (Billow 1975; Vosniadou *et al.* 1984; Waggoner & Palermo 1989).

All these successfull outcomes, while they opened interesting perspectives on semantic processing in very young children, did not inform us on how these subjects coped with metaphors at a specific metalinguistic level (Cometa & Eson 1978), which is precisely the level that the whole process of schooling tends to stimulate in order to bring students toward an abstract, context-independent way of thinking, what M. Donaldson (1978) called «desembedded thought»[1]. To this end, metalinguistic abilities have a special role because they are highly involved in text comprehension, critical reading, paraphrasing skills, notetaking, all reflective abilities crucial to academic achievement. Beside their involvement in study skills in first language, metalinguistic abilities have been studied in relation to bilingualism as one of its possible beneficial effects (Reynolds 1991; Bialystok (ed) (1991); Hamers & Blanc 2000: Jessner 2006), and metalinguistic awareness, in general, has been viewed as the strongest predictor of good proficiency in second or third language acquisition (Lasagabaster 1997, 1998; Perales & Cenoz 2002; Jessner 2006).

---

[1] For the role of schooling under this respect, see also Olson, 1996; Olson & Bruner 1996.

Within the various forms of metalinguistic abilities - metagrammatical, metasemantic, metaphonological, metapragmatic, etc. (Tunmer *et al.* 1984; Gombert 1990) - those involved in metaphor comprehension come into play *not only for understanding special usages in literary texts*, but also for understanding ordinary teachers' communication in the classroom. It is a common instructional practice, in fact, to convey new notions and concepts through metaphors, often spontaneously created on the spot.

To give an instance of this practice, we will consider an item of the MCT, namely: "Intelligence is a skyscraper". If we take each word separately, "intelligence" pertains to the domain of mental functions and "skyscraper" to that of "architectural works", but their juxtaposition induces a semantic bridging where intelligence is likely to be seen in its incremental, multilayered and challenging features, on one hand, and, on the other, the skyscraper may be viewed as an expression of intelligence. All these aspects are suddenly visualised in a new light when the separate words, «intelligence» and «skyscraper», are juxtaposed for the first time. Under this respect, comprehending a new metaphor is an act of thought in a full sense, as Lakoff and Johnson have argued (Lakoff & Johnson 1980). At the same time, when we are asked to interpret a metaphor of the type 'A is B', as in the above instance, we must implicitly accept a sort of 'fiction game', since our interlocutor cannot state that 'A is B' on *literal* grounds and we must interpret his/her communicative intention. Therefore, understanding metaphors is partly also a form of 'sociocognitive decentering' in a piagetian sense, in relation to someone else's mind, to which we attribute the capability of creating 'intriguing', 'odd meanings'. Not surprisingly, children with severe communicative difficulties, as in the autistic spectrum troubles, show a specific deficit in elaborating metaphors produced by others, in spite of relatively normal cognitive and linguistic potentials (Thomas *et al*, 2010; Rundblad & Annaz 2010; Melogno *et al*. 2011). Furthermore, as all metaphors draw their meanings from a given culture and are acknowledged within its borders and might be nonsensical in other cultures or in a different historical period of the same culture, we may say that the ability to interpret metaphors is partly also an ability to play a sort of 'cultural game'. Before skyscrapers existed, the utterance «Intelligence is a skyscraper» would have been a nonsense.

We will propose to view *metaphor as a form of semantic conflict generated by the anomalous combination of the conventional meanings of tenor and vehicle*. The solution of this confllict requires the analysis of both terms and the selection of those relevant features likely to justify the metaphorical 'ground' that links tenor to vehicle in such an anomalous combination. To be able to analyse linguistic meanings and to select just those relevant to solve a semantic conflict is precisely a form of *metasemantic* processing, as anticipated (see above, p.58).

In constructing the assessment scale of MCT, we drew on Piaget's last model of equilibration (1975), where Piaget posited the existence of three types of what he called "mental regulations" for coping with cognitive tasks. We remind of the fact that, for Piaget, any advance in the process of knowing arises from the perturbation

of a given equilibrium both in the outer world and in mental representation. Therefore, the activation of cognitive processes comes from a conflictual source, as it typically appears in any conservation task, when children are confronted with the problem of identity of substance, weight, volume, lenght, etc..., that remain the same in spite of the trasformations of the form of the objects (Piaget & Inhelder 1941; Piaget *et al.* 1948).

According to Piaget's latest conception of equilibration, there can be three types of "mental regulations" for solving these cognitive conflicts, called "alfa", "beta" and "gamma". With *alfa regulations*, subjects just ignore perturbations, whether by denying the very existence of the conflict or by simply refusing to cope with it. With *beta regulations*, the conflict is acknowledged and partially faced, but only by means of local and gradual adjustments. With *gamma regulations*, subjects have an immediate sight of how conflicting data can be composed in a new equilibrium and, consequently, their solution is at the same time consistent and mentally anticipated. For instance, in conservation tasks, when they reach the concrete operational level, children are immediately able to say that "things are still the same", in spite of changes in form, because what has been lost in one dimension (e.g. lenght or height or width, etc.) can be exactly compensated by something that has increased in another dimension, or that "If you transform things the other way round, it will all come back as it was at the start".

If we try to transpose this piagetian schema to the solution of the peculiar type of semantic conflict generated by creative metaphors, the following possibilities can be envisaged.

- Alfa regulations: subjects ignore the essence of the semantic conflict or refuse to cope with it. They might refuse the metaphorical usage in global terms or deny the very legitimacy of the metaphor or else accept it only on exceptional grounds, which confirms that, according to them, a metaphor is normally something semantically inacceptable. A further variant of this level can be a response that is simply overvague or exclusively centred on tenor or on vehicle. For instance, for the item: "Intelligence is a skyscraper", responses at the alfa level can range from:

> I don't know" to: "No, you can't say it: skyscrapers have nothing to do with intelligence" (or the reverse), or: "Yes, in skyscrapers there may be intelligent people working in it", "In movies you sometimes see intelligent managers in skyscrapers", or else: "It (intelligence) never stops working", "The skyscraper can go up to the sky.

- *Beta regulations:* instead of thinking that tenor ("intelligence") and vehicle ("skyscraper") are irreconcilable or concilable only on physical grounds as in real life, or in movies, in the media, etc..., subjects try to analyse some relevant characteristic of each term and find some common feature. The *semantic conflict is apparently acknowledged, although faced tentatively:* responses tend to focus on features midway between physical and symbolic.

Ex.: "Yes, you can say this because intelligence can be very high"... or "it ("intelligence") can have many floors"..."many steps"..."with intelligence you can go upward...".

In fact, each of these responses can be interpreted whether as physical or as metaphorical, but the comparative process upon which the whole metaphor is built is not made explicit and therefore they have been considered as intuitive and approximative solutions.

Gamma regulations: the implicit comparison that the metaphor establishes between tenor and vehicle is made explicit and framed at its proper symbolic level, beyond physical analogies. By distinguishing both differences and similarities between tenor and vehicle and by coordinating them appropriately, the semantic conflict is at the same time acknowledged and recomposed in a consistent way. At this level, subjects underline the fact that intelligence, *like* skyscrapers but *not physically* inside skyscrapers, appears as:

a never-ending process", "something where the final step is difficult to see, and still there must be some", that "there are steps in intelligence just as in skyscrapers there are many floors", "that intelligence is a sort of challenge like a skyscraper for architects.

We are aware that, in previous studies, Piagetian theory has been used for interpreting the development of metaphor comprehension in children, but within a structuralist framework. For instance, Cometa & Eson (1978) exploring the relation between a metaphor comprehension task and a Piagetian class intersection task, found a positive correlation. On the other hand, the concept of equilibration belongs to the functionalist part of the theory, that Piaget particularly developed during the last decade of his scientific activity (Inhelder 1982)[2]. To our knowledge, the possibility of using the Piagetian schema of mental regulations as a key for interpreting metaphor comprehension in children has not been explored so far. We attempted to do it in the present study as we believe that it can illuminate the way children process metalinguistic tasks[3].

2 - Method
2.1 - Items development

The MCT was originally composed of 12 items subdivided into two groups of metaphors: the psycho-physical (P.P.) and the conceptual (C.), based on a different

[2] About the shift from a structuralist to a functionalist approach to the whole thought-language relation in Piaget's theory in its last decade, see Montangero (1989).

[3] In fact, this applies precisely to three Italian metalinguistic tests, respectively for 4- to 6 year-old children (TAM-1, Pinto, Candilera 2000), 9- to 14 year-old children (TAM-2, Pinto, Candilera, Iliceto 2003), late adolescence to adulthood (TAM-3, Pinto, Iliceto 2007). Currently, they are validated in Italian language but also exist in English translation as MAT-1, MAT-2 and MAT-3 (Pinto *et al.*,1999), and in Spanish, as THAM-1, THAM-2 and THAM-3 (Pinto et al. 2000), although not validated in these languages.

type of link between tenor and vehicle. In P.P. metaphors, tenor is a human being and vehicle an object or a non human animate being, whereas in the C group of metaphors, tenor is a general concept or a category of objects, and vehicle is a single object or, sometimes, also a category of objects. The items have been drawn mostly from seminal research articles on the development of metaphor processing, such as Winner, Rosentiel & Gardner's (1976), Johnson & Pascual Leone's (1989), or Nippold & Sullivan's (1987), Waggoner & Palermo's (1989), Silberstein's et al. (1982), whereas others are taken from spontaneous metaphorical sentences created by children of the age range considered for the test. All the items are in Italian language[4], mostly translated and partially adapted from the original English ones and some spontaneously produced by Italian children. Here is the list of the 12 items.

Psycho-physical (P.P.) metaphors
Trial item: This baby-sitter is a cup of chocolate.
1) The guardian of the prison is a rock (*ad. from* Winner, Rosentiel & Gardner 1976)
2) My sister is a butterfly (*ad. from* Johnson & Pascual Leone 1989)
3) This child is a train without the locomotive (*ad. from* Nippold & Sullivan 1987)
4) This man is a volcano (*ad. from* Silberstein *et al.* 1982)
5) Betty is a soap bubble (*ad. from* Waggoner & Palermo 1989)
6) This child is a doggy without the leash (*ad. from* Nippold & Sullivan 1987)

Conceptual (C.) metaphors
Trial item: Mind is a computer.
1) Family is an umbrella (*ad. from* Evans, Gamble 1988)
2) Flowers are a garden's calendar (*ad. from* Evans, Gamble 1988)
3) Friendship is a coat (*ad. from* Fonzi, Negro Sancipriano 1975)
4) Autumn leaves are old photographs (*ad. from* Silberstein *et al.* 1982)
5) Intelligence is a skyscraper (*from spontaneous production*)
6) Memory is a sieve (*from spontaneous production*)

2.3 - Administration

MCT is a paper-and-pencil test that can be administered whether individually in face-to-face interaction, with the examiner writing the responses on a special protocol, or in a group, for instance in a classroom, where children write the responses on their own protocol. Whatever the condition, individual or collective, before starting each part of the test the examiner presents two trial items with the following directions: "You are going to listen to some sentences where words

---

[4] Some years ago, Pafumi (2005) translated the MCT into Spanish and administered the Italian and the Spanish version to two subsamples of randomly chosen pupils, from 4th to 8th grade, of an Argentinean school where Italian is taught as a compulsory second language. Results showed the same developmental pattern in each subsample, beyond differences in the language of the test, although performances were better in the Spanish version, which is the first language of the subjects.

Pinto, Maria Antonietta; Melogno, Sergio; Iliceto, Paolo - Assessing metaphor comprehension as a metasemantic ability in students from 9- to 14 years-old

LINGVARVM ARENA - VOL. 2 - ANO 2011 - 57 - 77

are used in a slightly different way from usual; try to explain in what sense these sentences can be understood. For instance, if we say: "This baby-sitter is a cup of chocolate" (P.P. type of metaphors), in what sense can we understand it ?", or if we say: "Mind is a computer" (C. type of metaphors), in what sense can we understand it ?".

The examiner and the child, if face to face, start to analyse features of tenor and vehicle in search of a common semantic denominator. If the child remains silent, the examiner takes the initiative and stimulate him/her to continue on his/her own until a satisfactory solution is found on the basis of the above criteria and then recapitulates the relevant common features of tenor and vehicle that make the trial item plausible and understandable (E.g.: "Baby-sitter" and "cup of chocolate" are both "sweet and good", "psychologically and physically supportive", "energizing", "comforting", "rewarding", etc...). Then the examiner says: "Now you will do it yourself", and starts presenting the items one after the other. If the test is administered in a classroom, the presentation is the same as with the individual child, the only difference being that more children are likely to contribute the analysis of the trial items and point to different aspects; this generally enriches discussion and facilitates the examiner's recapitulation. The time required for individual administration ranges from 15 to 30 minutes, depending on age and mental characteristics of each child, whereas the administration to groups takes from a minimum of 40 to a maximum of 60 minutes, including the initial presentation. In order to see if the quality of the responses could vary as a function of the way of administering the test, individual or in group, a pilot study has been conducted with 120 children, 60 for each modality, with 15 children per age level, but no significant differences appeared.

2.4 - The assessment scale

Above, we indicated three fundamental levels of metaphor comprehension, in terms of levels of processing of a semantic conflict between tenor and vehicle. These will be considered as general criteria for building a 4-step scale for MCT assessment. Below, some examples for the aforementioned item "Intelligence is a skyscraper".

- **Level 0 (score: 0):** *the semantic conflict is whether denied or refused. Refusals and/ or denials can take the following forms.*
a) No responses: blanks in written protocols or "I don't know" responses in oral and written form.
b) "No, it's nonsense".
c) Metonymic responses: "A skyscraper is full of people who work in it and are intelligent".
d) Overvague responses or responses exclusively focused on vehicle or on tenor: "if you are intelligent you are grown up and tall", "It ("intelligence") never stops working", "It ("the skyscraper") can go up to the sky".

- Level 1 (score: 1): *the conflict is faced but tentatively, with a common ground based on features midway between physical and symbolic:*

Ex.: "Intelligence and skyscrapers are high and have many floors", "It ("intelligence") gets built   gradually", "It ("intelligence") is something which goes higher and higher.

- Level 2 (score: 2): *the conflict is explicitly acknowledged and the structure of the metaphor is justified; both psychological characteristics of the tenor and physical characteristics of the vehicle are mentioned but on a clearly distinct basis.*

Ex.:"With intelligence you can reach any goal and important positions in life, as a skyscraper that has many floors", "It may mean that there are different levels of intelligence in people, as there are floors in skyscrapers".

- Level 3 (score:3): *This is a refinement of the previous level, based on a deeper and more exhaustive analysis of the psychological characteristics of the tenor "intelligence", in relation to the vehicle "skyscraper".*

Ex.: "As in a staircase, the more you go upward the more you can understand, but only to a certain point, because there must be a threshold, some limitations, also depending on age".

"As you grow up, levels of intelligence also open up, because they cannot develop all at once,   and as in skyscrapers you see things gradually".

"When you grow up in intelligence you can look at things from an upper level and understand them better, as in skyscrapers you discover things and they look small."

The first two steps of the above scale (score 0 and 1) globally correspond to a "pre-metaphorical" level, since the essence of the semantic conflict is either ignored or responded elusively. Symmetrically, at the third and fourth step (score 2 and 3) responses are specifically metaphorical. For each subset of metaphors, P.P. and C., the total score results from the addition of the separate score of each item and, in turn, the total score of MCT results from the sum of the total P.P. and the total C. scores.

2.5 - Participants

The sample is composed of 874 children (410 males and 464 females) from different regions of Italy, Northern, Central and Southern (Lombardy, Latium, Abruzzi, Campania, Calabria), aged 9 to 14 years, an age range that corresponds to grades 4th to 8th in the Italian school system. Table 1 reports subjects' distribution, by age and gender. Socio-cultural background, based on instruction and profession indicators of both parents, was homogeneous and average, following the current standards for Italian cities of medium size[5].

---

[5] For instance, in each family, level in education ranged from 8th grade, which is the end of compulsory studies in Italy, and the degree, but for one parent only.

Table 1 – Characteristics of sample

School grade		Male	Female	Total
4th-5th grade	N	112	111	223
	%	27.3%	23.9%	25.5%
6th grade	N	94	119	213
	%	22.9%	25.6%	24.4%
7th grade	N	97	109	206
	%	23.7%	23.5%	23.6%
8th grade	N	107	125	232
	%	26.1%	26.9%	26.5%
Total	N	410	464	874
	%	100.0%	100.0%	100.0%

2.6 - Procedure

Seven graduate students, specifically trained, collected and independently rated the 874 protocols of the sample. Pilot studies, conducted with children of the same age range, offered a broad spectrum of examples, 3 for each of the 4 levels of the assessment scale and for each of the 12 items (144 examples in all). Every student tried his/her assessment on his/her own; all doubtful cases were gathered and discussed within the training group until a solution was found within one of the 4 levels of the above scale. For calculating interrater reliability, the protocols of 120 subjects (30 for each grade) were randomly selected and scored independently by two raters, one of the co-authors and one doctoral level psychologist having undergone the same type of training as the students having collected the data.

2.7 - Data analysis

To determine interrater reliability we used weighted Cohen's Kappa for the 4-way classification score that provides an index of agreement between evaluators. Descriptive statistics for item of MCT were provided: *mean* as measure of central tendency, *standard deviation* as measure of dispersion, *skewness* and *kurtosis* to test the normality of shape of the distribution. Principal components analysis (PCA) with *oblimin rotation* was performed. The number of factors was determined by *scree plot* and by *eigenvalues* greater than 1.0. Items were identified for each subscale by their factor loadings. Subscales were scored by adding up the score of items endorsed on each component. MCT total was composed by the sum of subscale scores. *Cronbach's Alpha* was used as a measure of internal consistency for each subscale and MCT total. To compare MCT total and subscale mean scores by grade and gender we used analyses of variance (ANOVA) with Duncan test for *post hoc* comparisons. Corrected item-total correlations, test–retest stability and convergent validity were evaluated using Pearson's r correlation coefficients. All statistical analyses were conducted using SPSS 15.0 for Windows.

3 - Results
3.1 - Factor structure

The shape of the distribution and factor analysability of the matrix were assessed through descriptive statistics, *skewness* and *kurtosis* between -1 and 1, all measures of sampling adequacy (MSA's) > .800, Bartlett's test of sphericity (Chi^2 = 969.5; p <.000) and the Kaiser-Mayer-Olkin (KMO) test = .832 (Table 2).

Table 2 – Descriptive statistics

	Mean	Std. Deviation	Skewness	Kurtosis	MSA
Intelligence is a skyscraper	,94	,83	,595	-,222	,815
My sister is a butterfly	1,24	,87	,351	-,499	,824
Memory is a colander	1,42	1,10	-,038	-,344	,827
Flowers are a garden's calendars	1,04	,94	,310	-,118	,867
Betty is a soap bubble	1,09	,83	,448	-,296	,855
This child is a train without the locomotive	1,39	,92	,102	-,825	,825
Autumn leaves are old photographs	1,18	,87	,217	-,748	,828
The guardian of the prison is a rock	1,82	,77	-,025	-,660	,805
Family is an umbrella	1,57	1,02	-,242	-,879	,843
Friendship is a coat	1,42	1,00	-,117	-,635	,833
This child is a doggy without the leash	1,37	,90	-,088	-,881	,813
This man is a volcano	1,57	1,05	-,381	-,097	,863

The correlations matrix of MCT items of the 874 participants was subjected to a Principal Components Analysis. Based on the results of the *scree plot* (Cattel & Vogelman 1997) a two-factor solution was extracted, easily interpretable after *oblimin rotation* with *eigenvalues* > 1.0 and accounting for 32.78% of the variance. The initial *eigenvalue* for the first component, which accounted for 17.69% of the variance after rotation, was 2.86. The *eigenvalue* for the second component (15.09% of the variance after rotation) was 1.07. These two components were correlated (r = .394). These two factors partially correspond to the two groups of metaphors we initially classified as C. and P.P., although unequally distributed: 5 *vs* 7 instead of 6 *vs* 6 in each part. The item that mostly contributes to the definition of the first factor, with highest factor loading (0.727), is the one analysed so far: "Intelligence is a skyscraper", whereas the item that mostly contribute to the definition of the second group is: "The guardian of the prison is a rock" (factor loading = 0.672), which, by the way, has become a classical quote in the literature on the development of metaphor comprehension.

Nevertheless, we must note that two of the items that had been viewed as metaphors of the C. type shifted towards the P.P. type after factor analysis, which calls for interpretation. The items in question are: "Family is an umbrella" and "Friendship is a coat", that, in general terms, apparently meet with the requirements of Conceptual metaphors, where tenor must be a general concept or category and vehicle a single object. But in a child's perspective it is also understandable that the concepts of "family" and of "friendship" become representable by means of concrete experiences with actual individuals belonging to these categories.

In other words, the idea of "family" is known through the experience of actual relatives just as the idea of "friendship" is known through the experience of actual friends. Note that these two concepts are the only ones having to do with human beings within the list of the 6 originary C. items, contrary to the other 4, that are based either on mental functions ("intelligence" and "memory") or on categories of physical objects ("autumn leaves" and "flowers"). Conversely, in the list of metaphors initially viewed as P.P., 3 items shifted towards the C. part after factor analysis: "My sister is a butterfly", "Betty is a soap bubble" and "This child is a train without the locomotive", where, on formal linguistic grounds, the tenor is represented by a human being and the vehicle by an object, following a typical P.P metaphorical connection. But if we look at the vehicle in these items we can note that all 3 refer to some escaping physical feature, difficult to touch and seize: a bubble soap, a moving train or a butterfly. Besides, their escaping character is precisely the idea that these metaphors intend to suggest in relation to the human individuals they refer to. On the other side, in the remaining initial list of P.P items, the vehicle is something more stable and tangible as a rock, a volcano or a doggy. In short, although the originary list of items had been based on the usual criteria for to the C./P.P. metaphor distinction, children processed some of these metaphors on semantic grounds partially different from those of adults, probably introducing some aspects of their actual experience.

In the definitive protocol, items are still grouped in two categories, following the results of factor analysis, and since more than a half of the items in each group still pertain to the initial classification, the corresponding P.P and C denominations have been kept. In practice, the items of the first factor have been considered as "prevailingly" pertaining to the P.P. type and those of the second factor as "prevailingly " pertaining to the C. type.

Construct validity has been assessed by using the corrected item-total correlation, i.e. the correlation of an item with the sum of the remaining items in the scale. All coefficients >.30 are representative and highly acceptable (Nunnally & Bernstein 1994) (Table 3).

Table 3 – Corrected Item-total correlations and factor loadings (N = 874)

	Factors		Item-total correlations
	C.	P.P.	
Intelligence is a skyscraper	**.727**	-.203	.328
My sister is a butterfly	**.664**	-.014	.405
Memory is a sieve	**.620**	-.048	.356
Flowers are a garden's calendars	**.499**	.067	.343
Betty is a soap bubble	**.373**	.270	.336
This child is a train without the locomotive	**.371**	.205	.380
Autumn leaves are old photographs	**.354**	.183	.316
The guardian of the prison is a rock	-.080	**.672**	.319
Family is an umbrella	-.047	**.627**	.317
Friendship is a coat	.033	**.573**	.339
This child is a doggy without the leash	.043	**.519**	.308
This man is a volcano	.239	**.347**	.339
Eigenvalues	2.86	1.07	
% Explained variance (after rotation)	17.69	15.09	
Boldfaced items with highest loadings on factors			

3.2 - Convergent validity

Convergent validity has been studied on a subsample of 642 children, 304 males and 338 females. Correlations have been calculated between MCT, 3 subtests of an Italian Metalinguistic battery, already validated for the same age range (TAM-2, Pinto, Candilera, Iliceto 2003, see footnote 2, p. 62), namely the subtests of *Synonymy*, *Acceptability* and *Ambiguity* and the SPM38 (Standard Progressive Matrices 1938, Raven 1982).

The metalinguistic tests have been selected for their metasemantic nature, i.e. for the fact that correct responses require reflection on meanings of single linguistic units, lexical and grammatical, something involved also in explaining the semantic grounds of metaphors in a plausible way. In all subtests of the TAM-2 assessment is based on a distinction between intuitive and explicit metalinguistic processing, respectively coded as L and ML. L stands for *Linguistic*, based on an intuitive knowledge of the rules of the language that ensures that a given response, in yes-or-no form, is likely to be the correct one. ML stands for *Metalinguistic*, and refers to the subsequent response to the same item, where subjects have to justify the grounds of their previous L response and therefore analyse language more reflectively.

On the other hand, the SPM38 have been chosen not only as an emblematic test of abstract non verbal intelligence but also because it is based on a type of reasoning that share some resemblance with the analysis of the structure of metaphors. On structural grounds, Raven's Progressive Matrices can be considered as a class intersection task. In order to find the appropriate figure that goes into the blank in each matrix requires to abstract the composition principles that underly the horizontal and the vertical axis of the matrix itself, and to combine these principles in a consistent way. Something similar takes place when we analyse the linguistic

meanings of tenor and vehicle in a metaphor, detect relevant analogies and differences and compose them in a plausible common ground.

In Table 4 we can visualise the correlations between all the L and ML scores of the 3 metalinguistic subtests and the 3 scores of the MCT, where, aside from the L part of the Synonymy subtest, all the values are positive and significant and the correlations between the SPM38 and the MCT are also positive and significant.

Table 4 - Correlations Matrix (Pearson's r) between MCT, TAM-2 and SPM38 (N= 642, * p<.05, ** p<0.01)

	C. metaphors	P.P. metaphors	MCT total
Synonymy L	.044	.066	.065
Synonymy ML	.106*	.105*	.125*
Acceptability L	.153*	.185*	.196*
Acceptability ML	.211**	.235**	.258**
Ambiguity L	.256**	.215**	.271**
Ambiguity ML	.243**	.254**	.286**
SPM38	.326**	.260**	.324**

3.3 - The role of gender and school grade

Being the MCT a developmental test, applied to a sample stratified by school grade and gender, we thought it relevant to study the influence of these two factors by means of a two-factor ANOVA, analysing main effects and interactions. No significant interaction emerged, contrary to what happened with gender and school grade as separate factors, that both generated significant differences in results. Girls obtained significantly higher scores than boys in C. metaphors and in the MCT total, as shown in Table 5, below. Nevertheless, in both cases, the *effect size* is under the lower limit: η^2 <.06 (Cohen 1988).

Table 5 - MCT results: ANOVA in relation to gender (N=874)

Gender	Mean	S.D.	$F_{(1,866)}$	p
C. Metaphors				
Boys	7.94	3.37	7.70	<.006
Girls	8.62	3.64		
P.P. Metaphors				
Boys	7.57	2.77	2.49	n.s.
Girls	7.91	2.76		
MCT Total				
Boys	15.50	5.25	7.28	<.007
Girls	16.53	5.52		

ANOVA results referred to school grade (Table 6) underlined the strength of this factor, with very significant differences in each score of the test and a high *effect size*, over the upper limit: C. metaphors: $\eta^2 > .14$; P.P. metaphors: $\eta^2 > .14$; MCT Total: $\eta^2 > .19$ (Cohen 1988). The post hoc Duncan test highlighted a regular incremental pattern, grade by grade and in each score of the test, as indicated by the letters a,b,c,d, next to the mean of each school grade. Therefore each advance in school grade brings a significant improvement in results, whatever the type of score, partial or total.

**Table 6** - MCT Results: ANOVA in relation to school grade (N=874)

School grade	Mean	S.D.	$F_{(3,866)}$	p
C. Metaphors				
4th-5th grade	6.22[a]	2.91	50.55	<.000
6th grade	8.16[b]	3.03		
7th grade	8.89[c]	3.58		
8th grade	9.90[d]	3.47		
P.P. Metaphors				
4th-5th grade	6.19[a]	2.72	48.11	<.000
6th grade	7.58[b]	2.62		
7th grade	8.14[c]	2.40		
8th grade	9.05[d]	2.49		
MCT Total				
4th-5th grade	12.41[a]	4.57	71.33	<.000
6th grade	15.74[b]	4.67		
7th grade	17.03[c]	5.13		
8th grade	18.94[d]	5.03		

[a, b, c, d] Significant differences of Duncan *post hoc* test (p <.05).
Different superscripts indicate significant difference.

This gradual growth is even clearer when we analyse how the percentages of each of the 4 qualitative levels of the assessment scale are distributed in relation to school grade. In Fig. 1, we can see the gradual decrease, with age, of the first two levels of response, that we called 'pre-metaphorical', and a symmetrical increase of the second and third level, where metaphors are processed on a symbolic plane, as required. Although the third level of response is visibly less frequent - as it is based on a more sophisticated analysis - the incremental pattern in relation to school grade is very similar to that of the second level.

Figura 1 - MCT total(%)
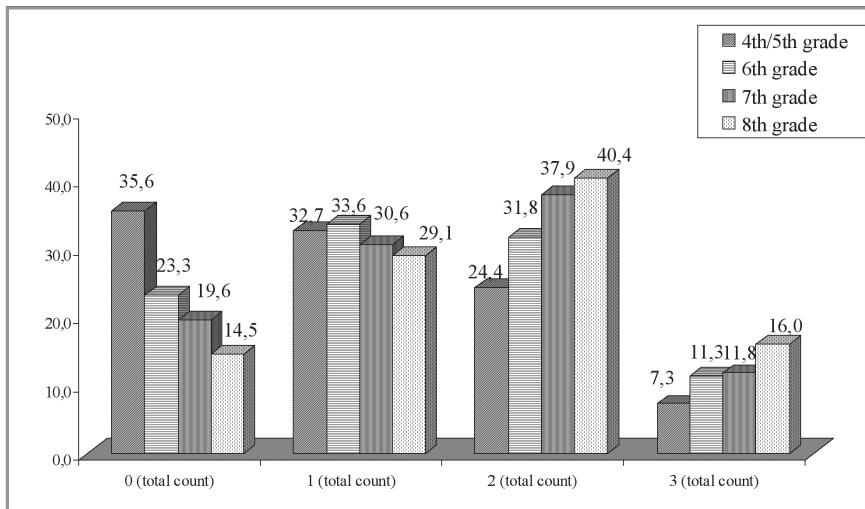
For each school grade and each type of score, standardized T-scores have been calculated as reference points into which empirical results can be converted at the corresponding age level. Within the total interval of the standardized T-scores (range: 0 - 100), 5 levels of metaphor processing have been defined, corresponding to the following 5 ranges: 0-30 as deficitary; 31-40 as low-average 41-60 as average; 61-70 as high-average and 71-100 as superior level.

3.4 - Reliability

Reliability has been studied as internal consistency by calculating the Cronbach's Alpha coefficients on the whole sample and by applying the test-retest method to a subsample of 240 subjects. Results, are shown in Table 7, below, and can be considered satisfactory.

Table 7 - Reliability coefficients for the MCT total and its parts. Cronbach's Alpha (N=874) and test- retest (N=240) coefficients.

Tasks	Cronbach's Alpha	Test-retest
	(N= 874)	(N=240)
C. Metaphors	0.62	0.78
P.P. Metaphors	0.51	0.69
MCT Total	0.70	0.83

3.5 - Interrater reliability

On a subsample of 120 protocols, subdivided into the 4 school grade groups and assessed by two independent raters, interrater agreement has been calculated by using Cohen's Kappa coefficient (1960). A very satisfactory agreement has been reached, as shown in Table 8, where coefficients and standard errors for each school grade are reported.

Table 8 - Interrater reliability (Cohen's Kappa coefficients and standard errors) for the 4 school grade groups.

4th-5th gr. (N=30)		6th gr. (N=30)		7th gr. (N=30)		8th gr. (N=30)	
K	k_{se}	K	k_{se}	K	k_{se}	K	k_{se}
.75	.027	.74	.028	.67	.030	.81	.024

4 - Discussion

The aim of this study was to present a developmental test for assessing a special type of metasemantic ability, namely the ability of analysing and justifying the semantic structure of metaphors on explicit grounds. The targeted age range, from late childhood to mid-adolescence, has been chosen as an interval where important changes take place in cognitive and linguistic abilities, towards a more abstract type of processing; these, in turn, are likely to bring about a significant transition from literal to symbolic responses in metaphor comprehension. By studying the role of age and gender, the developmental nature of the test clearly appeared as age proved to be the dominant factor, with a steady incremental pattern from pre-metaphorical to fully metaphorical responses in the age range considered.

Results from principal components analyses revealed two main dimensions, that partially correspond to the structural distinction between conceptual (C.) and psycho-physical (P.P.) metaphors in the initial list of items. In the discussion of the factor structure (see above, pp. 67-69), we provided a possible explanation for the discrepancies between the original classification of the items and the results of factor analysis. These discrepancies could point to modalities of semantic processing that children of a transitional age are elaborating along criteria different from those of adults. This phenomenon affects the processing of tenor in some of the C. metaphors - where tenor is a concept made out of human beings – and, symmetrically, it affects the processing of vehicle in some of the P.P metaphors, where vehicle is not as seizable and tangible as the other vehicles involved in this type of metaphors. In order to test this hypothesis, further research could devise two tenors of conceptual items with more abstract characteristics, such as to discourage from processing them on experiential grounds, and three vehicles of P.P. metaphors with more stable physical characteristics, easily representable by the child.

The study of convergent validity could be further explored and reinforced by administering the whole battery of the metalinguistic test adopted, the TAM-2 (Pinto et al. 2003), which comprises three more subtests and, above all, three total scores likely to highlight interesting correlations with the MCT scores. These are the total L score, i.e. the sum of the scores assessing metalinguistic processing at a first, intuitive level, the total ML score, i.e. the sum of the scores assessing metalinguistic processing at an explicit, analytical level and the total TAM-2 score, i.e. the sum of the total L and ML scores.

This study described the nature and the psychometric characteristics of the MCT, a test for assessing metaphor comprehension from 9- to 14 years-old. While further research can refine and strengthen some of its aspects, for instance through cross-linguistic research in other languages, the test offers itself as a measure of a peculiar type of metasemantic ability. As such, this instrument is of interest to developmental psychologists concerned with first language development, but also to L1 or L2 language teachers.

———————————————————————

References

Ash, S.; Nerlove, H. (1960), "The Development of Double Function Terms in Children: an Exploratory Study", in B. Kazlan, S. Wapners (Eds.), Perspectives in Psychological Theory New York: International University Press.

Bialystok, E. (ed). 1991. *Language processing in bilingual children*. Cambridge: Cambridge University Press.

Billow, R. M. 1975. A Cognitive Developmental Study of Metaphor Comprehension. *Developmental Psychology*. 11: 415-423.

Cattel, R.B.;Vogelman, S. 1977. A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research*. 12: 289-325.

Cohen, J. 1960. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*. 20: 37-46.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.

Cometa, M. S.; Eson, M. E. 1978. Logical operations and metaphors interpretation: A Piagetian Model. *Child Development.* 49: 649-659.

Danesi, M. 1986. The Role of Metaphor in Second Language Pedagogy. *Rassegna Italiana di Linguistica Applicata*. 18: 1-10.

Donaldson, M. 1978. *Children's minds*. Glasgow: Collins.

Duranti, A. (ed). 2001. *Linguistic Anthropology: A Reader*. Malden, MA: Blackwell.

Evans, M.A.; Gamble, L. 1988. Attribute Saliency and Metaphor Interpretation in

School-age Children. *Journal of Child Language.* 15, 435-449.

Fonzi, A.; Negro Sancipriano, E. 1975. *La magia delle parole. Alla riscoperta della metafora,* Torino: Einaudi.

Gentner, D. 1988. Metaphor as Structure-Mapping: The Relational Shift. *Child Development.* 59 : 47-59.

Gombert, J. E. 1990. *Le développement métalinguistique*. Paris: PUF.

Hamers, J.; Blanc, M. 2000. *Bilinguality and Bilingualism.* Cambridge: Cambridge University Press.

Happé, F. 1993. Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition.* 48:101-119.

Happé, F. 1995. Understanding minds and metaphors: insight form the study of figurative language in Autism. *Metaphor and Symbolic Activity.* 10: 275-95.

Inhelder, B. 1982. Sulle ultime ricerche di Jean Piaget. In: L. Camaioni, L. (a cura di), *La teoria di Jean Piaget. Recenti sviluppi e applicazioni.* Firenze: Giunti Barbera.

Jessner, U. 2006. *Linguistic awareness in Multilinguals. English as a Third Language.* Edimburgh: Edimburgh University Press.

Johnson, J.; Pascual Leone, J. 1989. Development Levels of Processing in Metaphor Interpretation. *Journal of Experimental Child Psychology.* 48: 1-31.

Lasagabaster, D.1997. *Creatividad y conciencia metalingüística: incidencia en el aprendizaje del ingles como L3.* Tésis doctoral – Universidad del País Vasco – Vitoria Gasteiz.

Lasagabaster, D. 1998. Metalnguistic awareness and the learning of English as an L3, *ATLANTIS.* XX (2): 69-79.

Lakoff, G.; Johnson, M. 1980. *Metaphors We Live by,* Chicago: University of Chicago Press.

Melogno, S.; D'Ardia, C.; Mazzoncini, B. ; Levi, G. 2011, La métaphore chez les enfants avec troubles envahissants du développement non spécifiés. Etude d'un cas clinique. *Neuropsychiatrie de l'enfance et de l'adolescence*. 1: 41-47.

Melogno S.; Pinto M. A. 1996. La comprensione di metafore in bambini dai 4 ai 6 anni. Una ricerca in contesto italiano. *Rassegna Italiana di Linguistica Applicata*. 3: 43-74.

Montangero, J. 1989. Introduction. In: J. Montangero and A. Tryphon (Eds), *Language and Cognition,* pp. 1-6: Geneva: Cahiers de la Fondation Archives Jean Piaget.

Nippold, M.A.; Sullivan, M. P. 1987. Verbal and Perceptual Analogical Reasoning and Proportional Metaphor Comprehension in Young Children. *Journal of Speech and Hearing Research.* 27: 197-205.

Nunnally, J.C.; Bernstein, I.H. 1994. *Psychometric Theory*. New York: McGraw-Hill.

Olson, D. R. 1996. Literate mentalities: Literacy, consciousness of language and modes of thought. In: D.R. Olson, N. Torrance (Eds.). *Modes of thought*, 141-151. Cambridge: Cambridge University Press.

Olson, D.R.; Bruner, J. 1996. Folk psychology and folk pedagogy. In: D.R. Olson, N. Torrance (Eds.). *The handbook of education and human development*,

Cambridge: Blackwell Publishers.

Ortony, A.(ed.).1979. *Metaphor and thought*, Cambridge: Cambridge University Press.

Pafumi, M. L. 2005. La comprensión de metáforas en alumnos bilingües español-italiano de 9 a 14 años. Un estudio translingüístico en Argentina. *Rivista di Psicolinguistica Applicata/Journal of Applied Psycholinguistics.* 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 Currriculum.* 15, I: 1-15.

Piaget, J.;Inhelder, B.;Szeminska. 1948. *La géométrie spontanée chez l'enfant.* Paris: PUF.

Piaget, J. 1975. *L'équilibration des structures cognitives.* Paris: PUF.

Piaget, J.; Inhelder, B. 1941. *Le développement des quantités physiques chez l'enfant.* Neuchâtel: Delachaux-Niestlé.

Pinto, M. A.; Candilera, G. 2000. *La valutazione del primo sviluppo metalinguistico. Il TAM-1 (Test di abilità metalinguistiche n. 1: 4 – 6 anni.,* Milano: Franco Angeli.

Pinto, M. A.; Candilera, G.; Iliceto, P. 2003. TAM-2. *Test di abilità metalinguistiche n. 2 (9-14 anni). La valutazione dello sviluppo metalinguistico tra scuola elementare e scuola media. Manuale di istruzioni.* Roma: Scione.

Pinto, M. A.; Iliceto, P. 2007. TAM-3.*(Test di abilità metalinguistiche n.3).* Fascia adolescente-adulta. Roma: Carocci.

Pinto, M. A.; Melogno, S.; Intaglietta, B. 2000. Il primo sviluppo metalinguistico e la comprensione di metafore in rapporto all'apprendimento precoce di una lingua straniera. *Rassegna Italiana di Linguistica Applicata.* 2-3: 75-106.

Pinto, M. A.; Titone ,R.; Trusso, F. 1999. *Metalinguistic awareness. Theory, development and measurement instruments.* Pisa-Roma: I.E.P.I.

Pinto, M. A.; Titone, R.; Gonzales Gil, L. 2000. *La consciencia metalingüística: Teoría, desarrollo e instrumentos de medición.* Pisa-Roma: I.E.P.I.

Raven, J.C. 1982. *Progressive Matrices 1938. Manuale di istruzioni.* Firenze: Organizzazioni Speciali.

Reynolds, A. G. 1991. The cognitive consequences of bilingualism. In: A.G. Reynolds (ed.). *Bilingualism, Multiculturalism and Second Language Learning: The McGill Conference in Honour of Wallace E. Lambert.* Hillsdale, New Jersey: Erlbaum, 145-182.

Richards, I.A. 1936. *The Philosophy of Rhetoric,* Oxford: Oxford University Press.

Rundblad, G.; Annaz, D. 2010. The atypical development of metaphor and metonymy comprehension in children with autism. *Autism.* 14: 29-46.

Silberstein, L.; Gardner, H.; Phelps, E.; Winner, E. 1982. Autumn leaves and old photographs: the development of metaphors preferences. *Journal of Experimental Child Psychology.* 34: 135-150.

Siltanen, S. A. 1986. "Butterflies are rainbows?": A developmental Investigation of Metaphor Comprehension, *Communication Education.* 35: 1-10.

Thomas, M.S.C.; Van Duuren, M.; Purser, H.; Meschal, D.; Ansari, D.; Karmiloff-Smith, A. 2010. The development of metaphorical language comprehension in typical development and in Williams syndrome. *Journal of Experimental*

Child Psychology. 106: 99-114.

Tunmer, W. E.; Pratt, C.; Herriman, M. L. 1984. *Metalinguistic Awareness in Children. Theory, Research and Implications.* Berlin: Springer.

Vosniadou, S.; Ortony, A.; Reynolds, R.E.; Wilson, P.T. 1984. Sources of difficulty in the young child's understanding of metaphorical language. *Child Development*. 55: 1588-1606.

Waggoner, J. E.; Palermo, D. S. 1989. Betty is a Bouncing Bubble: Children's Comprehension of Emotion-descriptive Metaphors. *Developmental Psychology*. 25: 152-163.

Winner, E. 1997. *The Point of Words. Children's Understanding of Metaphor and Irony.* Harvard: Harvard University Press.

Winner, E.; Rosentiel, A. K.; Gardner, H. 1976. The Development of Metaphoric Understanding. *Developmental Psychology.* 12: 289-297.

Winner, E. (1988), *The Point of the Words*, Cambridge (M.A.): Harvard University Press.

Zhang, F.; Hu, J. 2009. A Study of Metaphor and its Application in Language Learning and Teaching. *International Education Studies*. Vol.2, N. 2: 77-81.